



UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

Leandro Luiz de Lima

**Teorema da função implícita e da função
inversa e abordagens de funções reais de
várias variáveis no ensino médio**

**Rio Claro
2021**



UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

Teorema da função implícita e da função inversa e abordagens de funções reais de várias variáveis no ensino médio

Leandro Luiz de Lima

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação – Mestrado Profissional em Matemática em Rede Nacional, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Rio Claro.

Orientadora
Profa. Dra. Érika Capelato
FCL / UNESP / Araraquara (SP)

Rio Claro
2021

L732t Lima, Leandro Luiz de
Teorema da Função Implícita e da Função Inversa e abordagens de
funções reais de várias variáveis no ensino médio / Leandro Luiz de
Lima. -- Rio Claro, 2021
74 f. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp),
Instituto de Geociências e Ciências Exatas, Rio Claro
Orientadora: Érika Capelato

1. Teorema da Função Inversa. 2. Teorema da Função Implícita. 3.
Função real de várias variáveis. 4. Matemática. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Geociências e Ciências Exatas, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

TERMO DE APROVAÇÃO

Leandro Luiz de Lima

TEOREMA DA FUNÇÃO IMPLÍCITA E DA FUNÇÃO INVERSA E
ABORDAGENS DE FUNÇÕES REAIS DE VÁRIAS VARIÁVEIS NO
ENSINO MÉDIO

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação – Mestrado Profissional em Matemática em Rede Nacional, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, pela seguinte banca examinadora:

Profª. Dra. Érika Capelato
FCL / UNESP / Araraquara (SP)
Orientadora

Prof. Dr. Ricardo de Sá Teles
UTFPR / Londrina (PR)

Prof(a). Dra. Alessandra Aparecida Verri
CCET / UFSCAR / São Carlos (SP)

Rio Claro, 14 de setembro de 2021

*Aos meus pais, Sandra e José Luiz,
por me darem uma vida a sonhar,
e a minha querida esposa, Aline,
que partilha dos meus sonhos e divide os seus.*

AGRADECIMENTOS

A presente dissertação de mestrado ocorreu graças ao precioso apoio de diferentes pessoas. Em primeiro, Meus pais: Sandra Regina e José Luiz que me permitiram crescer com o melhor que os pais tem a oferecer a seus filhos, amor, carinho e a tão valiosa educação.

Minha esposa Aline que se tornou uma fiel companheira para todos os momentos que dividia minha atenção com este projeto e firmava meu caminho.

Não posso deixar de agradecer a minha orientadora, Professora Érika Capelato, por toda a paciência e empenho com que sempre me orientou. Sem me desmotivar, corrigir me quando necessário e fazer valer a pena cada momento do meu esforço nesta jornada.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001, à qual agradeço.

Desejo igualmente agraciar a todos os meus colegas do Mestrado cujo apoio e amizade estiveram presentes em todos os momentos.

Por fim quero agradecer a meus amigos, em diferentes momentos marcaram a minha superação com o apoio que tenho destes.

- Fernão, você se lembra do que disse, há muito tempo, acerca de amar o bando o bastante para voltar a ele e ajudá-lo a aprender?

— Claro que me lembro!

— Não compreendo como você consegue amar um punhado de pássaros que acabam de tentar matá-lo.

— Oh! Chico! Não é isso que você ama! Você não ama o ódio e o inferno, é claro.

Você tem de treinar até ver a verdadeira gaivota, o que há de bom em cada uma delas, e ajudá-las a ver isso nelas próprias. Para mim, o amor é isso. Quando você conseguir compreender e pôr isso em prática, até achará divertido.

Richard Bach

Resumo

Neste trabalho estudamos o teorema da função inversa e da função implícita. Para que pudéssemos apresentar o enunciado, demonstração e aplicações destes resultados, exploramos os seguintes conceitos: determinantes, topologia do espaço euclidiano e funções reais de várias variáveis. Finalmente, propomos atividades para que o estudo de funções reais de várias variáveis seja introduzido na sala de aula do ensino médio.

Palavras-chave: Teorema da Função Inversa, Teorema da Função Implícita, Função real de várias variáveis.

Abstract

In this work we study the theorem of the inverse function and the implicit function. In order to present the statement, demonstration and applications of these results, we explored the following concepts: determinants, Euclidean space topology and real functions of n variables. Finally, we propose activities for the study of multi-variable real function to be introduced in the high school classroom.

Keywords: Inverse Function Theorem, Implicit Function Theorem, Multi-variable real function.

Lista de Figuras

1.1	Estudo local por Newton	22
2.1	Projeção	36
2.2	Curva definida implicitamente no aberto $I \times J$	50
4.1	Imagem no software Paint	64
4.2	Bandeira do Brasil, Colorida	65
4.3	Bandeira do Brasil, escala cinza pela média ponderada	66
4.4	Retângulo com área máxima	67
4.5	Retângulo justaposto com área máxima	68

Sumário

1	Introdução	21
2	Resultados preliminares	25
2.1	Determinantes	25
2.1.1	Cálculo de determinantes	27
2.1.2	Teorema da unicidade do determinante	29
2.1.3	Expansão para determinantes. Menor principal e cofator	30
2.1.4	Existência da função determinante	33
2.2	Topologia do espaço euclidiano	35
2.2.1	O espaço euclidiano n -dimensional	35
2.2.2	Produto interno e norma	36
2.2.3	Bolas e conjuntos limitados	38
2.2.4	Sequência em \mathbb{R}^n	39
2.2.5	Aplicações contínuas	40
2.2.6	Conjuntos abertos	41
2.2.7	Conjuntos fechados	41
2.2.8	Conjuntos compactos	42
2.2.9	Conexidade	43
2.3	Funções reais de n variáveis	44
2.3.1	Derivadas Parciais	44
2.3.2	Derivadas direcionais	45
2.3.3	Funções diferenciáveis	46
2.3.4	Funções de classe C^k	48
3	Teorema da Função Implícita e Teorema da Função Inversa	55
3.1	Teorema da Função Implícita	55
3.2	Teorema da Função Inversa	58
3.3	Aplicações	60
4	Propostas de atividade para trabalhar funções reais de várias variáveis no Ensino Médio	63
4.1	Atividade 1: Equilibrando as finanças	63
4.2	Atividade 2: Computação gráfica: converter imagem colorida em monocromática	64
4.3	Atividade 3: Maximização da área de um quadrilátero	67
4.4	Atividade 4: Maximização da área de um quadrilátero justaposto	68
5	Considerações Finais	71

1 Introdução

De acordo com Krantz e Parkz (1951) o teorema da função implícita, junto com o teorema da função inversa, é um dos mais importantes e mais antigos paradigmas da matemática moderna.

Em seu livro, os autores escrevem que os primeiros trabalhos em álgebra, começando com *Al-jabr w'al mugâbala* de Mohammed ben Musa Al-Khowârizmî (8 d.c), apresentavam problemas e soluções por exemplos numéricos. A noção de uma função, seja implícita ou explícita, não fazia sentido em tal contexto. Foi só por volta de 1600 que a ideia de usar letras para denotar incógnitas e coeficientes foi introduzida por François Viète. Os métodos algébricos de Viète foram adotados por René Descartes e combinados com a inspiração do próprio sistema de coordenadas de Descartes. Este avanço, fundamental em 1637, finalmente trouxe a matemática ao ponto em que a noção de uma função pudesse fazer sentido. A abordagem das funções era mostrar como elas se comportam e não provar que elas existem.

Segundo Krantz e Parkz (1951), o trabalho de Newton, que descrevemos a seguir, pode ser um dos primeiros exemplos de análise do comportamento de uma função definida implicitamente. No trabalho “De Analysisi per Equationes Infinitas de 1669”, Newton aborda a questão de expressar a solução da equação

$$Y^3 + a^2Y - 2a^3 + aXY - x^3 = 0 \quad (1.1)$$

como uma série em x que será válida próximo a $x = 0$ e que dará a raiz $a \neq 0$ quando $x = 0$.

O procedimento é o que hoje conhecemos como “Método de Newton” para encontrar raízes, e a solução em série é apresentada como uma extensão do método numérico. Em 1670 Newton refina o seu método, o qual atualmente é conhecido como o polígono de Newton.

Para introduzirmos o método do polígono de Newton vamos considerar a seguinte equação:

$$y^3 + x^2y^2 - xy + x^4 = 0 \quad (1.2)$$

próxima de $x = 0$. O esboço desta equação é apresentado em azul na Figura 1.1. Suponha que exista uma solução da equação (1.2) da forma $y = y(x)$ que tem seu gráfico passando pela origem e que é definida, pelo menos, para valores de x pequenos. Em particular, teremos $y(0) = 0$.

A ideia do polígono de Newton é supor que podemos escrever

$$y(x) = x^\alpha \tilde{y}(x)$$

com $\tilde{y}(x)$ uma função contínua que não é nula quando $x = 0$. O número α é um parâmetro que deve ser escolhido apropriadamente para que $\tilde{y}(0)$ possa ser determinado.

Substituindo $y = x^\alpha \tilde{y}$ na equação (1.2) temos

$$x^{3\alpha} \tilde{y}^3 + x^{2\alpha+2} \tilde{y}^2 - x^{\alpha+1} \tilde{y} + x^4 = 0 \quad (1.3)$$

Para sermos capaz de determinar $\tilde{y}(0)$ usando a equação (1.3), deve haver dois ou mais monômios na equação (1.3) que têm a mesma potência de x e todos os outros monômios devem ter uma potência maior de x .

Assim, se usarmos $\alpha = 3$ na equação (1.3) teremos

$$x^9 \tilde{y}^3 + x^8 \tilde{y}^2 - x^4 \tilde{y} + x^4 = 0 \quad (1.4)$$

Dividindo a equação (1.4) por x^4 , obtemos:

$$x^5 \tilde{y}^3 + x^4 \tilde{y}^2 - \tilde{y} + 1 = 0 \quad (1.5)$$

Fazendo $x = 0$ na equação (1.5), obtemos $\tilde{y}(0) = 1$. Isto nos diz que a curva que descreve a equação (1.2) contém uma curva, próxima de $x = 0$, dada por

$$y = x^3$$

Esta curva é ilustrada em vermelho na Figura 1.1

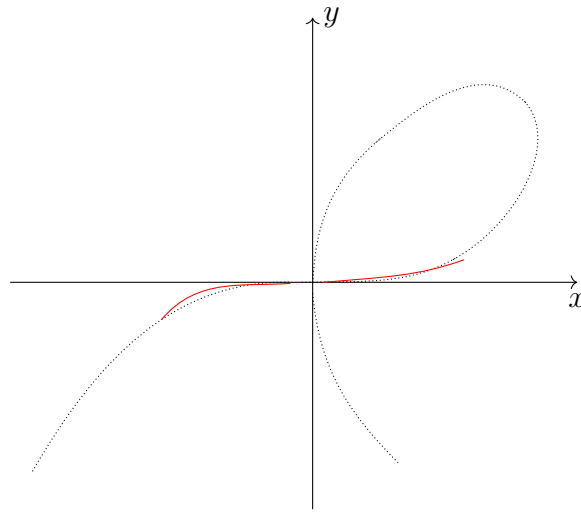


Figura 1.1: Estudo local por Newton

No ano de 1770 Lagrange provou o que pode ser o primeiro teorema da função implícita, mas em sua forma intimamente relacionada como um teorema da função inversa. O resultado é conhecido como Teorema da Inversão de Lagrange.

Já entre os anos 1876 - 1877 foi enunciado e demonstrado o Teorema da Função Implícita de variável real por Ulisse Dini.¹ A primeira formulação do teorema da função implícita ocorreu para funções de duas variáveis reais e a hipótese da matriz jacobiana ser *não singular*, então generalizada para diversas variáveis.

¹Ulisse Dini (1845-1918) passou toda sua carreira acadêmica na Universidade de Pisa. Ainda que seu trabalho mais importante em matemática fosse sobre a teoria das funções de variáveis reais, ele fez muitas contribuições para a geometria diferencial, séries de potências e funções analíticas.

Atualmente o teorema da função implícita possui demonstrações alternativas e muitas generalizações, é uma relíquia da matemática moderna, uma ferramenta importantíssima para a resolução de equações.

Assim, motivados por este importante resultado da matemática, o propósito deste trabalho é estudar o Teorema da Função Implícita e o Teorema da Função Inversa.

Além desta introdução, este trabalho é organizado da seguinte forma. No capítulo 2 fazemos um estudo dos seguintes conceitos preliminares: determinantes, topologia do espaço euclidiano e funções reais de n variáveis. No capítulo 3 enunciamos, demonstramos e fazemos aplicações do Teorema da Função Implícita e do Teorema da Função Inversa. No capítulo 4, propomos atividades que motivem o estudo de funções reais de várias variáveis no Ensino Médio. Finalmente, no capítulo 5, apresentamos as considerações finais deste trabalho.

2 Resultados preliminares

Neste capítulo apresentaremos os seguintes conceitos e resultados: função determinante, tópicos de topologia do espaço euclidiano, funções reais, derivadas e funções diferenciáveis, teorema do valor médio, teorema do valor intermediário. Eles são pré-requisitos para a compreensão e demonstração do *Teorema da Função Implícita* e do *Teorema da Função Inversa*.

As referências utilizadas neste capítulo são Apostol (1969), De Lima (2015), Lima (2018) e Royden e Fitzpatrick (1988).

2.1 Determinantes

Se $A = (a_{ij})$ é uma matriz $n \times n$ de entradas reais ou complexas, denotamos suas linhas por A_1, \dots, A_n . Então a i -ésima linha de A é um vetor no espaço n -dimensional dado por

$$A_i = (a_{i1}, a_{i2}, \dots, a_{in})$$

Consideremos o determinante como uma função de n linhas, A_1, \dots, A_n , e denotemos o seu valor por $d(A_1, \dots, A_n)$ ou por dA . A seguir daremos a definição axiomática do determinante como uma função.

Definição 2.1. *Uma função d de valor real ou complexo, definida para cada n -upla ordenada de vetores A_1, \dots, A_n no espaço n -dimensional, é chamada de função determinante de ordem n se ela satisfizer os seguintes axiomas, para todas as escolhas de vetores A_1, \dots, A_n e C neste espaço:*

Axioma 2.2. *Homogeneidade em cada linha: Se a k -ésima linha, A_k , é multiplicada por um escalar t , então o determinante também será multiplicado por t :*

$$d(\dots, tA_k, \dots) = td(\dots, A_k, \dots).$$

Axioma 2.3. *Aditividade em cada linha: Para cada k temos*

$$d(A_1, \dots, A_k + C, \dots, A_n) = d(A_1, \dots, A_k, \dots, A_n) + d(A_1, \dots, C, \dots, A_n).$$

Axioma 2.4. *O determinante é nulo se quaisquer duas linhas são iguais:*

$$d(A_1, \dots, A_n) = 0 \text{ se } A_i = A_j \text{ para qualquer } i \text{ e } j \text{ com } i \neq j$$

Axioma 2.5. *Determinante da matriz identidade é igual a 1:*

$$d(I_1, \dots, I_n) = 1, \text{ onde } I_k \text{ é o } k\text{-ésimo vetor de coordenada unitária.}$$

Os Axiomas (2.1) e (2.3) afirmam que o determinante de uma matriz é uma função linear de cada uma de suas linhas. Frequentemente dizemos que o determinante é uma função multilinear de suas linhas. Utilizando repetidas vezes a multilinearidade, podemos escrever:

$$d\left(\sum_{k=1}^p t_k C_k, A_2, \dots, A_n\right) = \sum_{k=1}^p t_k d(C_k, A_2, \dots, A_n),$$

onde t_1, \dots, t_p são escalares e C_1, \dots, C_p são vetores quaisquer do espaço de dimensão n .

Algumas vezes, uma versão mais fraca do Axioma (2.4) é usada:

Axioma 2.6. *O determinante é nulo se duas linhas adjacentes são iguais:*

$$d(A_1, \dots, A_n) = 0 \text{ se } A_k = A_{k+1} \text{ para algum } k = 1, 2, \dots, n-1.$$

Um fato notável, é que existe uma única função d que satisfaz os Axiomas 2.1, 2.3, 2.5 e 2.6. Este resultado será demonstrado no Teorema 2.10.

O próximo teorema fornece propriedades do determinante que serão deduzidas apenas dos Axiomas 2.1, 2.3 e 2.6.

Teorema 2.7. *Uma função determinante satisfazendo os Axiomas 2.1, 2.3 e 2.6 tem as seguintes propriedades:*

1. *O determinante é nulo se alguma linha é 0:*

$$d(A_1, \dots, A_n) = 0 \text{ se } A_k = 0 \text{ para algum } k.$$

2. *O determinante troca o sinal se duas linhas adjacentes são trocadas:*

$$d(\dots, A_k, A_{k+1}, \dots) = -d(\dots, A_{k+1}, A_k, \dots).$$

3. *O determinante troca o sinal se duas linhas quaisquer A_i e A_j com $i \neq j$ são trocadas.*

4. *O determinante é nulo se quaisquer duas linhas são iguais:*

$$d(A_1, \dots, A_n) = 0 \text{ se } A_i = A_j \text{ para algum } i \text{ e } j \text{ com } i \neq j.$$

5. *O determinante é nulo se suas linhas são dependentes.*

Demonstração. Para provar a primeira afirmação, basta tomar $t = 0$ no Axioma (2.1).

Para demonstrar a segunda afirmação. Seja B uma matriz com as mesmas linhas da matriz A , exceto para a linha k e a $k+1$. Ambas as linhas B_k e B_{k+1} iguais a $A_k + A_{k+1}$. Então pelo Axioma (2.6) $dB = 0$. Assim podemos escrever

$$d(\dots, A_k + A_{k+1}, A_k + A_{k+1}, \dots) = 0. \quad (2.1)$$

Aplicando a propriedade da adição na linha k e na linha $k+1$, reescrevemos a equação (2.1) como segue:

$$d(\dots, A_k, A_k, \dots) + d(\dots, A_k, A_{k+1}, \dots) + d(\dots, A_{k+1}, A_k, \dots) + d(\dots, A_{k+1}, A_{k+1}, \dots) = 0.$$

O primeiro e quarto termo são iguais a zero pelo Axioma (2.6). Então o primeiro e segundo termos são valores opostos um do outro.

Para demonstrar a terceira afirmação, vamos assumir que $i < j$. Podemos trocar a linha A_i e A_j realizando um número ímpar de trocas entre linhas adjacentes. Primeiro trocamos a linha A_j sucessivamente com linhas adjacentes anteriores $A_{j-1}, A_{j-2}, \dots, A_i$. Isto necessita $j - i$ trocas. Agora trocamos a linha A_i sucessivamente com as linhas adjacentes posteriores $A_{i+1}, A_{i+2}, \dots, A_{j-1}$. Desta vez são necessárias outras $j - i - 1$ trocas. Cada troca resulta na mudança do sinal da determinante, realizamos $(j - i) + (j - i - 1) = 2(j - i) - 1$ trocas ao todo (um número ímpar), com um número ímpar de trocas a determinante altera seu sinal.

Para demonstrar a quarta afirmação, basta tomar a matriz B obtida através da troca entre as duas linhas iguais da matriz A , A_i e A_j . Como $A_i = A_j$, temos que $B = A$ e então $dB = dA$. Pelo item 3 $dB = -dA$ então $dA = 0$.

Para demonstrar a última afirmação, suponha que existam escalares c_1, \dots, c_n , pelo menos algum diferente de zero, de modo que $A_1 = \sum_{k=2}^n c_k A_k = 0$. Então qualquer linha A_k com $c_k \neq 0$ pode ser escrito como uma combinação linear das outras linhas. Por simplicidade, suponha que A_1 é uma combinação linear de outros, assim podemos escrever $A_1 = \sum_{k=2}^n t_k A_k$. Pela linearidade da primeira linha temos

$$d(A_1, A_2, \dots, A_n) = d\left(\sum_{k=2}^n t_k A_k, A_2, \dots, A_n\right) = \sum_{k=2}^n t_k d(A_k, A_2, \dots, A_n).$$

porém cada termo $d(A_k, A_2, \dots, A_n)$ na última soma é zero, pois A_k é igual a pelo menos um de A_2, \dots, A_n . Portanto toda a soma vale zero. Agora, se a linha A_i é uma combinação linear das outras linhas, argumentamos do mesmo modo, utilizando linearidade na i -ésima linha. \square

2.1.1 Cálculo de determinantes

Com as ferramentas disponíveis até o momento podemos calcular a determinante em alguns casos usando apenas os axiomas e as propriedades do Teorema 2.7, assumindo que a função determinante exista. Observe que nos exemplos abaixo não utilizaremos o Axioma 2.5.

Exemplo 2.8. Determinante de uma matriz triangular superior. Uma matriz da forma

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

é chamada de *matriz triangular superior*. Todos seus elementos abaixo da diagonal principal são zeros. Vamos demonstrar que o determinante desta matriz é igual o produto dos elementos de sua diagonal

$$dU = u_{11}u_{22} \cdots u_{nn}.$$

Demonstração. Primeiro provaremos que o $dU = 0$, se algum elemento da diagonal $u_{ii} = 0$. Caso o último elemento u_n seja zero, então a última linha é O e $dU = 0$ pelo primeiro item do Teorema 2.7. Suponhamos então que algum elemento anterior da diagonal seja zero. Por simplicidade, seja $a_{22} = 0$. Então os $n - 1$ vetores linhas U_2, \dots, U_n , tem os dois primeiros componentes igual a zero. Portanto, estes vetores abrangem um subespaço de dimensão máxima $n - 2$. Por consequência estas $n - 1$ linhas são dependentes. Pelo último item do Teorema 2.7, $dU = 0$. Da mesma forma afirmamos que $dU = 0$ se qualquer elemento da diagonal é zero.

Agora vamos tratar o caso geral. Primeiro escrevemos a primeira linha U_1 como uma soma de dois vetores,

$$U_1 = V_1 + V'_1$$

onde $V_1 = [u_{11}, 0, \dots, 0]$ e $V'_1 = [0, u_{12}, u_{13}, \dots, u_{1n}]$. Pela linearidade da primeira linha, temos

$$dU = d(V_1, U_2, \dots, U_n) + d(V'_1, U_2, \dots, U_n).$$

Mas $d(V'_1, U_2, \dots, U_n) = 0$ visto que este é o determinante de uma matriz triangular com um elemento na diagonal igual a zero. Então, temos

$$dU = d(V_1, U_2, \dots, U_n). \quad (2.2)$$

Agora tratamos o vetor da linha U_2 de maneira similar, expressando como uma soma,

$$U_2 = V_2 + V'_2,$$

sendo $V_2 = [0, u_{22}, 0, \dots, 0]$ e $V'_2 = [0, 0, u_{23}, \dots, u_{2n}]$.

Trabalhando com o lado direito da equação (2.2) e aplicando linearidade na segunda linha, obtém-se

$$dU = d(V_1, V_2, U_3, \dots, U_m), \quad (2.3)$$

pois $(V_1, V'_2, \dots, V_n) = 0$. Repetindo estes argumentos para cada uma das linhas posteriores lado direito da equação (2.3), finalmente obtemos

$$dU = u_{11}u_{22} \cdots u_{nn},$$

como afirmado. □

Exemplo 2.9. Cálculo pelo método de Gauss-Jordan. O processo de eliminação de Gauss-Jordan para resolução de sistemas de equações lineares, também é um dos melhores métodos para calcular o determinante. Recordamos que o método consiste em aplicar três operações na linha da matriz.

1. Troca na posição de duas linhas;
2. Multiplicar todos os elementos de uma linha por um escalar diferente de zero;
3. Adicionar a uma linha o produto de outra linha por um escalar.

Aplicando estas três operações repetidas vezes de forma sistemática, é possível transformar qualquer matriz quadrada A em uma matriz triangular superior U cujo determinante sabemos calcular. É fácil determinar a relação entre a matriz A e a matriz U . Para cada vez que aplicar a operação de trocar posições, o determinante troca seu sinal. Para cada escalar que multiplica uma linha, o determinante também é multiplicado por c . Da última operação, adicionar uma linha em outra não altera o determinante. Concluimos que se as operações de trocar linhas p vezes e de multiplicar por escalares c_1, \dots, c_q , diferentes de zero, então o determinante será determinado pela equação abaixo:

$$dA = (-1)^p(c_1c_2 \cdots c_q)^{-1}dU. \quad (2.4)$$

Note que esta fórmula é consequência dos três primeiros Axiomas. Sua prova não depende do Axioma 2.5.

2.1.2 Teorema da unicidade do determinante

No Exemplo 2.8 da seção anterior mostramos que os Axiomas 2.1, 2.3 e 2.4 implicam na fórmula $dU = u_{11}u_{22} \cdots u_{nn}dI$. Combinando este fato com a equação (2.4) vemos que para toda matriz quadrada A de ordem $n \times n$ existe um escalar c , que depende de A , tal que

$$d(A_1, \dots, A_n) = c d(I_1, \dots, I_n). \quad (2.5)$$

Além disso, esta fórmula é uma consequência dos Axiomas 2.1, 2.3 e 2.4. A partir disto, podemos demonstrar facilmente que não existe mais de uma função determinante.

Teorema 2.10. (*TEOREMA DA UNICIDADE PARA DETERMINANTES*) *Seja d uma função satisfazendo os quatro axiomas para uma função determinante de ordem n , e seja f outra função satisfazendo os Axiomas 2.1, 2.3 e 2.4. Então para cada escolha dos vetores A_1, A_2, \dots, A_n no espaço de dimensão n , temos*

$$f(A_1, \dots, A_n) = d(A_1, \dots, A_n)f(I_1, \dots, I_n) \quad (2.6)$$

Além disso, se f também satisfaz o Axioma 2.5, teremos

$$f(A_1, \dots, A_n) = d(A_1, \dots, A_n).$$

Demonstração. Seja $g(A_1, \dots, A_n) = f(A_1, \dots, A_n) - d(A_1, \dots, A_n)f(I_1, \dots, I_n)$. Provaremos que $g(A_1, \dots, A_n) = 0$ para cada escolha de A_1, \dots, A_n . Uma vez que d e f satisfazem os Axiomas 2.1, 2.3 e 2.4, o mesmo vale para g . Então g também satisfaz a equação (2.5) que foi deduzida utilizando os três primeiros axiomas. Consequentemente, podemos escrever

$$g(A_1, \dots, A_n) = cg(I_1, \dots, I_n), \quad (2.7)$$

onde c é um escalar que depende da matriz A . Tomando $A = I$ na definição de g , note que d satisfaz o Axioma 2.5, encontramos

$$g(I_1, \dots, I_n) = f(I_1, \dots, I_n) - f(I_1, \dots, I_n) = 0$$

Portanto, a equação (2.7) torna-se $g(A_1, \dots, A_n) = 0$, completando a demonstração. \square

2.1.3 Expansão para determinantes. Menor principal e cofator

Nesta seção vamos explorar as propriedades de linearidade e o Teorema da Unicidade para demonstrar que se o determinante existir ele pode ser calculado pela fórmula que expressa cada determinante de ordem n como uma combinação linear de determinantes de ordem $n - 1$. A fórmula geral irá propor um método para provar a existência de funções determinantes por indução.

Cada uma das linhas de uma matriz A de ordem $n \times n$ podem ser expressas como uma combinação linear de n vetores de coordenadas unitárias I_1, \dots, I_n . Por exemplo, a primeira linha A_1 pode ser escrita como:

$$A_1 = \sum_{j=1}^n a_{1j} I_j.$$

uma vez que os determinantes são lineares na primeira linha, temos

$$d(A_1, A_2, \dots, A_n) = d\left(\sum_{j=1}^n a_{1j} I_j, A_2, \dots, A_n\right) = \sum_{j=1}^n a_{1j} d(I_j, A_2, \dots, A_n). \quad (2.8)$$

Portanto, para calcular a dA é suficiente calcular $d(I_j, A_2, \dots, A_n)$ para cada coordenada do vetor I_j .

Vamos utilizar a notação A'_{1j} para representar a matriz obtida de A substituindo a primeira linha A_1 pelo vetor unitário I_j . Exemplo, se $n = 3$ Existem três matrizes:

$$A'_{11} = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad A'_{12} = \begin{bmatrix} 0 & 1 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad A'_{13} = \begin{bmatrix} 0 & 0 & 1 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Observe que $dA'_{1j} = d(I_j, A_2, \dots, A_n)$. A equação (2.8) pode ser escrita na forma

$$dA = \sum_{j=1}^n a_{1j} \det A'_{1j}. \quad (2.9)$$

Esta é a chamada *Fórmula da expansão*; e expressa o determinante de A como uma combinação linear dos elementos em suas primeiras linhas.

O argumento usado para obter a equação (2.9) pode ser aplicado à k -ésima. O resultado será a fórmula da expansão em termos dos elementos da k -ésima linha.

Teorema 2.11. (*EXPANSÃO POR COFATORES*) *Seja A'_{kj} a representação obtida da matriz A substituindo a k -ésima linha, A_k , pelo do vetor de coordenada de unidade I_j . Então temos a fórmula da expansão*

$$dA = \sum_{j=1}^n a_{kj} dA'_{kj}. \quad (2.10)$$

a qual expressa o determinante de A como combinação linear de elementos da k -ésima linha. O número dA'_{kj} é chamado de cofator de entrada a_{kj} .

No próximo teorema vamos demonstrar que cada cofator, diferenciando-se por um positivo e outro negativo, é igual a determinante de uma matriz de ordem $n - 1$. Estas matrizes reduzidas são chamadas de *menores*.

Definição 2.12. Dada uma matriz quadrada A de ordem $n \geq 2$, a matriz quadrada de ordem $n - 1$ obtida pela exclusão da k -ésima linha e j -ésima coluna de A é chamada k, j menor de A e denotada por A_{kj} .

Exemplo 2.13. Uma matriz $A = (a_{kj})$ de ordem três tem nove menores. A seguir apresentamos o 1,1 menor, o 1, 2 menor e o 1,3 menor de A :

$$A_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad A_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad A_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

O determinante da matriz A_{11} , de ordem dois, é definido através da diferença do produto da diagonal principal pelo produto da diagonal secundária. Ou seja,

$$dA_{11} = a_{22}a_{33} - a_{23}a_{32}.$$

Já para a matriz A , de ordem três, o determinante é definido pela *regra de Sarrus*¹:

$$dA = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \quad (2.11)$$

Usando os determinantes dos menores de A é fácil verificar que a equação (2.11) pode ser substituída por:

$$dA = a_{11}dA_{11} - a_{12}dA_{12} + a_{13}dA_{13}. \quad (2.12)$$

O próximo teorema estenderá esta fórmula para os casos $n \times n$ para qualquer $n \geq 2$.

Teorema 2.14. *Expansão pelo menor da k -ésima linha:* Para qualquer matriz A do tipo $n \times n$, $n \geq 2$, o cofator de a_{kj} relaciona-se com o menor A_{kj} pela fórmula

$$dA'_{kj} = (-1)^{k+j}dA_{kj}. \quad (2.13)$$

Portanto, a expansão da dA em termos dos elementos da k -ésima linha é dada por:

$$dA = \sum_{j=1}^n (-1)^{k+j} a_{kj} dA_{kj}. \quad (2.14)$$

Demonstração. Vamos iniciar a demonstração através do caso especial $k = j = 1$. A matriz A'_{11} tem a forma:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & & & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

Aplicando operações elementares de linha, podemos tornar cada elemento abaixo de 1 na primeira coluna igual a zero, deixando todas as outras como estão. Por exemplo, multiplicando a primeira linha por a_{21} e subtraindo da segunda linha obtemos $(0, a_{22}, a_{23}, \dots, a_{2n})$. Por um sucessão destas operações em cada linha, obteremos uma nova matriz a qual denotaremos por A'_{11} , a qual tem a forma

¹O nome refere-se ao matemático francês Pierre Frederic Sarrus (1798-1861).

$$A_{11}^0 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{32} & a_{33} & \dots & a_{3n} \\ 0 & & & & \vdots \\ 0 & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}.$$

Em vista que as operações de linhas não alteram o determinante, temos

$$dA_{11}^0 = dA'_{11} \quad (2.15)$$

Mas dA'_{11} é uma matriz em blocos, então, temos $dA_{11}^0 = dA_{11}$, onde A_{11} é 1, 1 menor de A ,

$$A_{11} = \begin{bmatrix} a_{22} & \dots & a_{2n} \\ a_{32} & \dots & a_{3n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Portanto, $dA'_{11} = dA_{11}$ ficando provado 2.13 para $k = j = 1$.

Vamos considerar o próximo caso especial $k = 1$, j arbitrário e provar que

$$dA'_{1j} = (-1)^{j-1} dA_{1j}. \quad (2.16)$$

Uma vez demonstrado a equação (2.16) a fórmula geral, dada pela equação (2.13) fica demonstrada, uma vez que a matriz A'_{kj} pode ser transformada em uma matriz da forma B'_{1j} por $k - 1$ trocas sucessivas de linhas adjacentes. O determinante muda o sinal a cada troca, assim, temos

$$dA'_{kj} = (-1)^{k-1} dB'_{1j}, \quad (2.17)$$

sendo B uma matriz $n \times n$ cuja primeira linha é I_j e cujo 1, j menor $B_{1,j}$ é A_{kj} . Da equação (2.16), temos

$$\det B'_{1j} = (-1)^{j-1} \det B_{1j} = (-1)^{j-1} dA_{kj},$$

então de (2.17) temos

$$\det A'_{kj} = (-1)^{k-1} (-1)^{j-1} dA_{kj} = (-1)^{k+j} dA_{kj}.$$

Portanto se provarmos (2.16) fica provado (2.13).

Vamos demonstrar a equação (2.16). A matriz A'_{1j} tem a forma

$$A'_{1j} = \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \\ a_{21} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix}.$$

Através das operações elementares nas linhas, passamos a ter uma coluna de zeros abaixo do 1 e transformamos a matriz em

$$A_{1j}^0 = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{21} & \dots & a_{2,j-1} & 0 & a_{2,j+1} & \dots & a_{2n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & 0 & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}.$$

Como antes, a determinante é inalterada, então $dA_{1j}^0 = \det A'_{1j}$. O $1, j$ menor de $A_{1,j}$ tem a forma

$$A_{1j} = \begin{bmatrix} a_{21} & \dots & a_{2,j-1} & a_{2,j+1} & \dots & a_{2n} \\ \vdots & & \vdots & & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}.$$

Agora consideramos dA_{1j}^0 como uma função de $n - 1$ linhas de A_{1j} , digamos $dA_{1j}^0 = f(A_{1j})$. A função f satisfaz os três primeiros axiomas da função determinante de ordem $n - 1$. Portanto, pelo Teorema da Unicidade, podemos escrever

$$f(A_{1j}) = f(J)dA_{1j}, \tag{2.18}$$

onde J é a matriz identidade de ordem $n - 1$. Portanto para demonstrar (2.16) devemos mostrar que $f(J) = (-1)^{j-1}$. Agora $f(J)$ é, pela definição, o determinante da matriz

$$C = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & & 0 & 0 & 0 & & 0 \\ \vdots & \diagdown & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \diagdown & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \leftarrow j\text{-ésima linha}$$

\uparrow
 $j\text{-ésima coluna}$

Os elementos ao longo das linhas inclinadas são todos 1. Os outros elementos não visíveis, são todos 0. Através da troca sucessiva da primeira linha de C com as linhas $2, 3, \dots, j$ chegamos na matriz identidade $I, n \times n$ após $j - 1$ trocas. Como a determinante altera seu sinal para cada troca, temos $C = (-1)^{j-1}$. Portanto $f(J) = (-1)^{j-1}$ o que demonstra (2.16) e conseqüentemente (2.13). \square

2.1.4 Existência da função determinante

Nesta seção, usamos a indução em n , o tamanho da matriz, para provar que existem funções determinantes de cada ordem. Para $n = 2$ já mostramos que a função determinante existe. Também podemos dispensar o caso de $n = 1$ pela definição $d[a_{11}] = a_{11}$.

Assumindo que a função determinante exista para o caso $n - 1$, um candidato lógico para a função determinante de ordem n seria uma das fórmulas de expansão do Teorema 2.14, por exemplo, a expansão nos termos dos menores da primeira linha. Entretanto, é mais fácil verificar os axiomas se usarmos uma fórmula diferente, mas análoga, expressa em termos do menor da primeira coluna.

Teorema 2.15. *Assuma que a determinante de ordem $n - 1$ existe. Para qualquer matriz $A = (a_{jk})$ do tipo $n \times n$, seja f uma função definida pela fórmula*

$$f(A_1, \dots, A_n) = \sum_{j=1}^n (-1)^{j+1} a_{j1} dA_{j1}. \quad (2.19)$$

Então f satisfaz todos os quatro axiomas para uma função determinante de ordem n . Portanto, por indução, determinantes de ordem n existem para qualquer n .

Demonstração. Considere cada termo da soma em (2.19) como um função das linhas de A e escrevemos

$$f_j(A_1, \dots, A_n) = (-1)^{j+1} a_{j1} dA_{j1}.$$

Se verificarmos que cada f_j satisfaz os Axiomas (2.1) e (2.3), o mesmo será verdade para f .

Considere a possibilidade de multiplicar a primeira linha de A por um escalar t . O menor de A_{11} não é alterado, desde que não envolva a primeira linha. O coeficiente a_{11} é multiplicado por t , então temos

$$f_1(tA_1, A_2, \dots, A_n) = ta_{11} dA_{11} = tf_1(A_1, A_2, \dots, A_n).$$

Se $j > 1$ a primeira linha de cada menor A_{j1} por ser multiplicada por t e o coeficiente a_{j1} não é afetado, então temos

$$f_j(tA_1, A_2, \dots, A_n) = tf_j(A_1, A_2, \dots, A_n).$$

Portanto, cada f_j é homogêneo na primeira linha.

Se a k -ésima linha de A é multiplicado por t , onde $k > 1$, o menor A_{k1} não se altera, mas a_{k1} é multiplicado por t , então f_k é homogêneo na k -ésima linha. Se $j \neq k$, o coeficiente a_{j1} não se altera, mas alguma linha de A_{j1} será multiplicado por t . Então todo f_j é homogêneo na k -ésima linha.

Um argumento similar mostra que para cada f_j é adicionado em toda linha, logo f satisfaz os Axiomas (2.1) e (2.3). Provaremos a seguir que f satisfaz o Axioma (2.6).

Para verificar que f satisfaz o Axioma (2.6), assumimos que duas linhas adjacentes são iguais, digamos $A_k = A_{k+1}$. Então, com exceção dos menores A_{k1} e $A_{k+1,1}$, cada menor A_{j1} tem duas linhas iguais, logo $dA_{j1} = 0$, portanto a soma em (2.19) consiste apenas nos dois termos correspondentes a $j = k$ e $j = k + 1$,

$$f(A_1, \dots, A_n) = (-1)^{k+1} a_{k1} + (-1)^{k+2} a_{k+1,1} dA_{k+1,k}. \quad (2.20)$$

Mas $A_{k1} = A_{k+1,1}$ e $a_{k1} = a_{k+1,1}$, pois $A_k = A_{k+1}$. Portanto os dois termos em (2.20) diferenciam-se apenas pelo sinal, logo $f(A_1, \dots, A_n) = 0$. Concluimos que f satisfaz o Axioma (2.6).

Por fim, vamos verificar que f satisfaz o Axioma (2.5). Quando tivermos $A = I$, consequentemente temos $a_{11} = 1$ e $a_{j1} = 0$ para $j > 1$. Afinal, A_{11} é a matriz identidade de ordem $n - 1$, logo cada termo em (2.19) é zero, exceto o primeiro que é igual a 1. Portanto $f(I_1, \dots, I_n) = 1$ o que satisfaz o Axioma (2.5). \square

Na demonstração anterior, poderíamos também usar uma função f definida em termos da k -ésima coluna menor de A_{jk} em vez dos menores da primeira coluna A_{j1} . De fato, se tomarmos

$$f(A_1, \dots, A_n) = \sum_{j=1}^n (-1)^{j+k} a_{jk} dA_{jk}, \quad (2.21)$$

exatamente o mesmo tipo de prova mostra que este f satisfaz todos os quatro axiomas para a função determinante. Uma vez que as funções determinantes são únicas, a fórmula da expansão em (2.21) e aquela em (2.14) são todas iguais a dA .

2.2 Topologia do espaço euclidiano

2.2.1 O espaço euclidiano n -dimensional

Seja n um número natural. O espaço *euclidiano n -dimensional* é o produto cartesiano de n fatores iguais a \mathbb{R} ;

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

Os pontos de \mathbb{R}^n são todas as n -listas $x = (x_1, \dots, x_n)$ cujas coordenadas x_1, \dots, x_n são número reais. Dados $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ em \mathbb{R}^n tem-se $x = y$ se, e somente se, $x_1 = y_1, \dots, x_n = y_n$.

$\mathbb{R}^1 = \mathbb{R}$ é a reta, isto é, o conjunto dos números reais. \mathbb{R}^2 é o plano, ou seja, o conjunto dos pares ordenados $z = (x, y)$ de números reais. \mathbb{R}^3 é o espaço euclidiano tridimensional da geometria analítica, cujos pontos são os ternos ordenados $p = (x, y, z)$.

Dados $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ em \mathbb{R}^n e um número real α , definimos a soma $x + y$ e o produto $\alpha \cdot x$ pondo

$$\begin{aligned} x + y &= (x_1 + y_1, \dots, x_n + y_n), \\ \alpha \cdot x &= (\alpha x_1, \dots, \alpha x_n). \end{aligned}$$

Estas operações fazem de \mathbb{R}^n um espaço vetorial de dimensão n sobre o corpo dos reais, no qual o elemento neutro para a adição é $0 = (0, \dots, 0)$ e o simétrico de $x = (x_1, \dots, x_n)$ é $-x = (-x_1, \dots, -x_n)$.

Os elementos de \mathbb{R}^n serão às vezes chamados *pontos* e às vezes *vetores*. Geometricamente considerar $x \in \mathbb{R}^n$ como vetor significa imaginar a seta que tem origem no ponto 0 e extremidade em x .

No espaço vetorial chama-se de *base canônica* ou *base natural* $\{e_1, \dots, e_n\}$, formada pelos vetores

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

que têm uma coordenada igual a 1 e as outras nulas. Dado $x = (x_1, \dots, x_n)$ em \mathbb{R}^n , tem-se

$$x = (x \cdot e_1 + \dots + x_n \cdot e_n).$$

A base canônica do espaço euclidiano permite estabelecer uma bijeção natural entre o conjunto $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ das aplicações (ou transformações) lineares $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ e o conjunto $M(n \times m)$ das matrizes reais (A_{ij}) com n linhas e m colunas.

2.2.2 Produto interno e norma

Um produto interno num espaço vetorial real E é uma regra que faz corresponder a cada par de vetores $x, y \in E$ um número real, indicado por $\langle x, y \rangle$, de tal modo que, para qualquer $x, x', y \in E$ e $\alpha \in \mathbb{R}$, se tenham:

- I. $\langle x, y \rangle = \langle y, x \rangle$;
- II. $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$;
- III. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle = \langle x, \alpha y \rangle$;
- IV. $x \neq 0 \Rightarrow \langle x, x \rangle > 0$.

Isto se exprime dizendo que um produto interno é uma função real simétrica, bilinear, positiva definida, $E \times E \rightarrow \mathbb{R}$.

O *produto interno canônico* do espaço euclidiano \mathbb{R}^n é dado por

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n,$$

onde $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$.

Salvo menção explícita em contrário, o produto interno canônico é o único que consideramos no espaço euclidiano \mathbb{R}^n .

Dado $x \in \mathbb{R}^n$, escrevemos $|x| = \sqrt{\langle x, x \rangle}$, ou seja,

$$|x| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Tem-se $|x|^2 = \langle x, x \rangle$, de modo que $|x| = 0 \Leftrightarrow x = 0$ e $|x| > 0 \Leftrightarrow x \neq 0$.

O número $|x|$ chama-se a *norma euclidiana* ou o *comprimento do vetor* $x \in \mathbb{R}^n$.

Dois vetores $x, y \in \mathbb{R}^n$ são ortogonais quando $\langle x, y \rangle = 0$. Evidentemente, o vetor 0 é ortogonal a todos os vetores de \mathbb{R}^n . Também e_i é ortogonal a e_j se $i \neq j$.

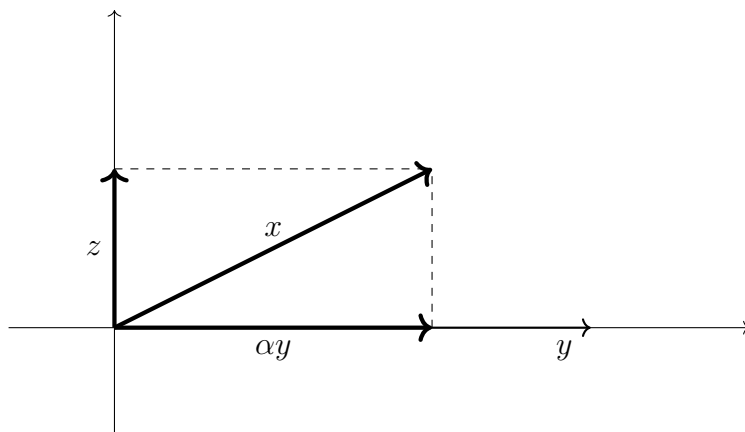


Figura 2.1: Projeção

Outro caso menos trivial de ortogonalidade é:

Proposição 2.16. *Dado dois vetores $x, y \in \mathbb{R}^n$, com $y \neq 0$ e $\alpha = \frac{\langle x, y \rangle}{|y|^2}$, podemos observar que o vetor $z = x - \alpha y$ é ortogonal a y .*

Demonstração. Tomando $z = x - \alpha y$

$$\begin{aligned}\langle z, y \rangle &= \langle x - \alpha y, y \rangle = \langle x, y \rangle - \alpha \langle y, y \rangle = \langle x, y \rangle - \alpha |y|^2 \Rightarrow \\ \langle z, y \rangle &= \langle x, y \rangle - \langle x, y \rangle = 0.\end{aligned}$$

□

A Figura 2.1 ilustra que o vetor αy é a projeção ortogonal de x sobre a reta que contém y .

Utilizaremos a proposição 2.16 para demonstrar a desigualdade fundamental da Geometria Euclidiana.

Teorema 2.17. (*Desigualdade de Cauchy-Schwarz*). Para quaisquer $x, y \in \mathbb{R}^n$, tem-se $|\langle x, y \rangle| \leq |x| \cdot |y|$. Vale a igualdade se, e somente se, um dos vetores x, y é um múltiplo escalar do outro.

Demonstração. Isto é óbvio se $y = 0$. Se, porém, for $y \neq 0$, colocamos $\alpha = \frac{\langle x, y \rangle}{|y|^2}$. Como acabamos de ver, o vetor $z = x - \alpha y$ é ortogonal a y . Segue daí que

$$|x|^2 = \langle z + \alpha y, z + \alpha y \rangle = |z|^2 + \alpha^2 |y|^2, \text{ donde}$$

$$|x|^2 \geq \alpha^2 |y|^2 = \frac{\langle x, y \rangle^2}{|y|^4} |y|^2, \text{ ou seja:}$$

$|x|^2 |y|^2 \geq \langle x, y \rangle^2$. Vale a igualdade se, e somente se, $z = 0$, ou seja, $x = \alpha y$. □

A norma euclidiana $|x| = \sqrt{\langle x, x \rangle}$ goza das seguintes propriedades, onde $x, y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ e $|\alpha|$ significa o valor absoluto do número real α :

- N1. $|x + y| \leq |x| + |y|$;
- N2. $|\alpha \cdot x| = |\alpha| \cdot |x|$;
- N3. $x \neq 0 \Rightarrow |x| > 0$.

As duas últimas são evidentes e a primeira propriedade decorre da desigualdade de Cauchy-Schwarz.

De modo geral, uma norma num espaço vetorial E é qualquer função real $|\cdot|: E \rightarrow \mathbb{R}$ que cumpra as condições N1, N2 e N3 acima estipuladas.

Há uma infinidade de normas que se podem considerar no espaço euclidiano \mathbb{R}^n . A norma euclidiana é motivada pela fórmula do comprimento de um vetor no plano em coordenadas cartesianas, que se prova com o Teorema de Pitágoras. Há outras duas normas que são de manipulação formal mais simples. Elas são:

$$|x|_M = \max\{|x_1|, \dots, |x_n|\} \quad \text{Norma do Máximo}$$

$$|x|_S = |x_1| + \dots + |x_n| \quad \text{Norma da Soma}$$

É fácil mostrar que as condições N1, N2 e N3 são satisfeitas e que para todo $x \in \mathbb{R}^n$

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

onde $|\cdot|$ é a norma euclidiana.

Generalizando as desigualdades que relacionam estas três normas usuais no \mathbb{R}^n , podemos dizer que: duas normas arbitrárias $|\cdot|$ e $||\cdot||$ em \mathbb{R}^n são *equivalentes* quando existirem constantes $a > 0$ e $b > 0$ tais que

$$|x| \leq a ||x|| \text{ e } ||x|| \leq b |x| \text{ para todo } x \in \mathbb{R}^n$$

Uma norma num espaço vetorial E dá origem a uma noção de distância em E . Dados $x, y \in E$, a distância de x a y é definida por

$$d(x, y) = |x - y|.$$

As condições N1, N2 e N3, que a norma satisfaz, implicam imediatamente que a distância desfruta das propriedades a seguir, para quaisquer $x, y, z \in E$:

$$\begin{aligned} d1. \quad d(x, z) &\leq d(x, y) + d(y, z) \\ d2. \quad d(x, y) &= d(y, x) \\ d3. \quad x \neq y &\Rightarrow d(x, y) > 0 \end{aligned}$$

A primeira propriedade é chamada de *desigualdade triangular*, um lado do triângulo não excede a soma dos outros dois lados.

Teorema 2.18 (Lima, 2018, p.19). *Duas normas quaisquer no espaço \mathbb{R}^n são equivalentes.*

2.2.3 Bolas e conjuntos limitados

Uma norma em \mathbb{R}^n permite que sejam definidas algumas noções geométricas básicas, que passamos a apresentar.

Definição 2.19. (*Bola aberta*) *Sejam a um ponto de \mathbb{R}^n e r um número positivo. O conjunto de todos os pontos $x \in \mathbb{R}^n$ tal que*

$$|x - a| < r$$

é chamado uma bola aberta de centro a e raio r . Denotaremos este conjunto por $B(a)$ ou $B(a; r)$.

A bola $B(a; r)$ consiste de todos os pontos cuja distância, partindo de a é menor que r . Em \mathbb{R}^1 este é um intervalo aberto com centro em a . Em \mathbb{R}^2 é um disco circular, em \mathbb{R}^3 é uma esfera de centro a e raio r .

Definição 2.20. (*Bola fechada*) *Sejam a um ponto de \mathbb{R}^n e r um número positivo. O conjunto de todos os pontos $x \in \mathbb{R}^n$ tal que*

$$|x - a| \leq r$$

é chamado bola fechada de centro a e raio r . Denotaremos este conjunto por $B[a]$ ou $B[a; r]$.

Definição 2.21. (*Esfera*) *Sejam a um ponto de \mathbb{R}^n e r um número positivo. O conjunto de todos os pontos $x \in \mathbb{R}^n$ tal que*

$$|x - a| = r$$

é chamado esfera de centro a e raio r . Denotaremos este conjunto por $S[a]$ ou $S[a; r]$.

Observe que $B[a; r] = B(a; r) \cap S[a; r]$.

Definição 2.22. *Sejam $x, y \in X \subset \mathbb{R}^n$. O segmento de reta de extremos x, y é o conjunto*

$$[x, y] = \{(1 - t)x + ty; 0 \leq t \leq 1\}.$$

Se o subconjunto $X \subset \mathbb{R}^n$ contém o segmento de reta com extremos $x, y \in X$ ele será dito convexo.

O próximo teorema diz que as bolas relativas a diferentes normas em \mathbb{R}^n têm em comum o fato de serem *convexas*.

Teorema 2.23. *Toda bola $B \subset \mathbb{R}^n$ é convexa.*

Demonstração. Seja $B = B(a; r)$. Se $x, y \in B$ então $|x - a| < r$ e $|y - a| < r$. Para qualquer $t \in [0, 1]$, temos:

$$\begin{aligned} |(1-t)x + ty - a| &= |(1-t)(x-a) + t(y-a)| \\ &\leq (1-t)|x-a| + t|y-a| \\ &< r \end{aligned}$$

□

Definição 2.24. *Um subconjunto $X \subset \mathbb{R}^n$ é dito limitado quando existe um número real $c > 0$ tal que $|x| \leq c$, pra todo $x \in X$. Isto equivale a dizer que X está contido na bola fechada de centro na origem e raio c .*

2.2.4 Sequência em \mathbb{R}^n

Uma sequência em \mathbb{R}^n é uma aplicação $f : \mathbb{N} \rightarrow \mathbb{R}^n$, que associa a cada número natural k um ponto $x_k \in \mathbb{R}^n$. As notações para uma sequência são (x_1, \dots, x_k, \dots) , $(x_k)_{k \in \mathbb{N}}$ ou simplesmente (x_k) .

Para cada $i = 1, \dots, n$, indicamos com x_{ki} a i -ésima coordenada de x_k . Assim $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$. Dar uma sequência em \mathbb{R}^n equivale a dar n sequências de números reais $(x_{k1})_{k \in \mathbb{N}}, \dots, (x_{kn})_{k \in \mathbb{N}}$.

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

Em particular, se a sequência $(x_k)_{k \in \mathbb{N}}$ é limitada então, para todo $i = 1, \dots, n$ a sequência $(x_{ki})_{k \in \mathbb{N}}$ das i -ésimas coordenadas de x_k também é limitada, pois $|x_{ki}| \leq |x_k|$. A recíproca também é verdadeira (Lima, 2004, p. 9).

Uma *subsequência* de $(x_k)_{k \in \mathbb{N}}$ é uma restrição desta sequência a um subconjunto infinito $\mathbb{N}' = \{k_1 < \dots < k_m < \dots\} \subset \mathbb{N}$. As notações $(x_k)_{k \in \mathbb{N}'}, (x_{k_m})_{m \in \mathbb{N}}$ ou $(x_{k_1}, \dots, x_{k_m})$ são usadas para indicar uma subsequência.

Dizemos que o ponto $a \in \mathbb{R}^n$ é o *limite* da sequência (x_k) quando, para todo $\epsilon > 0$ dado arbitrariamente, é possível obter $k_0 \in \mathbb{N}$ tal que $k > k_0$ implica $|x_k - a| < \epsilon$.

Em outras palavras: $k > k_0 \Rightarrow x_k \in B(a; \epsilon)$. Escrevemos $\lim_{k \rightarrow \infty} x_k = a$, $\lim_{k \in \mathbb{N}} x_k = a$ ou, simplesmente, $\lim x_k = a$.

De acordo com esta definição, temos $\lim x_k = a$ se, e somente se, $\lim |x_k - a| = 0$.

Uma sequência (x_k) em \mathbb{R}^n é *convergente* quando existe $a = \lim x_k$. Da observação acima, toda sequência convergente é limitada. É também natural que qualquer subsequência de uma sequência convergente é também convergente e tem o mesmo limite.

A definição de limite de uma sequência em \mathbb{R}^n faz uso de uma norma. As desigualdades que relacionam as três normas no espaço euclidiano nos dão: $|x_k - a|_M \leq |x_k - a| \leq |x_k - a|_S \leq n \cdot |x_k - a|_M$. Segue daí que a afirmação $\lim x_k = a$ independe de qual das três normas usuais estamos considerando.

Na demonstração do teorema abaixo usamos em \mathbb{R}^n , por conveniência, a norma do máximo. Porém, pelo o que acabamos de escrever, o fato nele enunciado é válido seja qual for a norma, dentre as três usuais.

Teorema 2.25. *A sequência (x_k) em \mathbb{R}^n converge para o ponto $a = (a_1, \dots, a_n)$ se, e somente se, para cada $i = 1, \dots, n$, tem-se $\lim_{k \rightarrow \infty} x_{ki} = a_i$, isto é, cada coordenada de x_k converge para a coordenada correspondente de a .*

Demonstração. Como $|x_{ki} - a_i| \leq |x_k - a|$, vemos que $\lim_{k \rightarrow \infty} x_k = a \Rightarrow \lim_{k \rightarrow \infty} x_{ki} = a_i$ para cada $i = 1, \dots, n$.

Reciprocamente, se $\lim_{k \rightarrow \infty} x_{ki} = a_i$ para todo $i = 1, \dots, n$ então, dado $\epsilon > 0$, existem $k_1, \dots, k_n \in \mathbb{N}$ tais que $k > k_i \Rightarrow |x_{ki} - a_i| < \epsilon$. Tomando $k_0 = \max\{k_1, \dots, k_n\}$. Então $k > k_0 \Rightarrow |x_k - a| = \max_i |x_{ki} - a_i| < \epsilon$. Logo $\lim x_k = a$. \square

Teorema 2.26. *[Bolzano-Weierstrass (Lima, 2018, p.17)] Toda sequência limitada em \mathbb{R}^n possui uma subsequência convergente.*

2.2.5 Aplicações contínuas

Definição 2.27. *Seja $f : X \rightarrow \mathbb{R}^n$ uma aplicação definida no conjunto $X \subset \mathbb{R}^m$. Diz-se que f é contínua no ponto $a \in X$ quando, para qualquer $\epsilon > 0$ dado, se pode obter $\delta > 0$ tal que todo ponto $x \in X$ cuja distância ao ponto a seja menor do que δ é transformado por f num ponto $f(x)$ que dista de $f(a)$ menor que ϵ . Em linguagem simbólica*

$$\forall \epsilon > 0 \exists \delta > 0; x \in X, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Em termos de bolas, a continuidade de f no ponto a se exprime assim: para toda bola aberta B' de centro $f(a)$ em \mathbb{R}^n existe uma bola aberta B de centro a em \mathbb{R}^m tal que $f(B \cap X) \subset B'$.

Se $f : X \rightarrow \mathbb{R}^n$ é contínua em todos os pontos do conjunto X , dizemos que f é uma *aplicação contínua*.

Embora a definição de continuidade de uma aplicação $f : X \rightarrow \mathbb{R}^n$ com $X \subset \mathbb{R}^m$ faça uso de uma norma em \mathbb{R}^m e outra em \mathbb{R}^n (ambas indicadas acima com a mesma notação). Segue da definição de normas equivalentes e do Teorema 2.18 que a continuidade (ou descontinuidade) de f persiste se alterarmos uma destas normas ou ambas.

A seguir daremos outra definição que independe da norma: dado $X \subset \mathbb{R}^m$, dizemos que a aplicação $f : X \rightarrow \mathbb{R}^n$ é *Lipschitziana* quando existe $k > 0$ (constante de lipschitz de f) tal que, para quaisquer $x, y \in X$ temos

$$|f(x) - f(y)| \leq k|x - y|$$

Toda aplicação Lipschitziana é contínua: dado $\epsilon > 0$ basta tomar $\delta = \frac{\epsilon}{k}$.

Toda transformação linear $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é Lipschitziana. Com efeito, seja $c = \max\{|T(e_1)|, \dots, |T(e_n)|\}$. Então para todo $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ temos

$$|T(x)| = |T(\sum x_i e_i)| = |\sum x_i T e_i| \leq \sum |x_i| |T e_i| \leq c \sum |x_i|$$

Tomando em \mathbb{R}^n a norma da soma, temos $|T(x)| \leq c|x|$ para todo $x \in \mathbb{R}^n$. Então, para $x, y \in \mathbb{R}^n$ arbitrários, vale:

$$|T(x) - T(y)| = |T(x - y)| \leq c|x - y|$$

Assim, T cumpre a condição de lipschitz, com uma constante igual à maior das normas dos vetores-coluna de sua matriz. Em particular, T é contínua.

Teorema 2.28. *Dados $X \subset \mathbb{R}^m, Y \subset \mathbb{R}^n, f : X \rightarrow \mathbb{R}^n$ contínua no ponto $a \in X$, com $f(X) \subset Y, g : Y \rightarrow \mathbb{R}^p$ contínua no ponto $b = f(a)$, então $g \circ f : X \rightarrow \mathbb{R}^p$ é contínua no ponto a .*

Demonstração. Dado arbitrariamente $\epsilon > 0$ existe, em virtude da continuidade de g , um número $\mu > 0$ tal que $y \in Y, |y - f(a)| < \mu \Rightarrow |g(y) - g(f(a))| < \epsilon$. Por sua vez, a partir de μ , a continuidade de f nos fornece $\delta > 0$ tal que $x \in X, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \mu$. Segue que $x \in X, |x - a| < \delta \Rightarrow |g(f(x)) - g(f(a))| < \epsilon$, logo $g \circ f$ é contínua no ponto a . \square

2.2.6 Conjuntos abertos

Definição 2.29. *Seja S um subconjunto de \mathbb{R}^n , e assumamos que $a \in S$. Então a é chamado ponto interior de S se existe uma bola aberta com centro em a que tem todos os seus pontos pertencentes a S .*

Em outras palavras, todo ponto interior a de S pode ser cercado por uma bola $B(a)$ tal que $B(a) \subseteq S$. O conjunto de todos os pontos interiores de S é chamado interior de S e denotado por $\text{int}S$.

Definição 2.30. *Um conjunto S em \mathbb{R}^n é aberto se todos os seus pontos são pontos interiores. Em outras palavras, S é aberto se e somente se $S = \text{int}S$.*

Teorema 2.31. *Seja $f : X \rightarrow \mathbb{R}^n$ uma aplicação definida no conjunto $X \subset \mathbb{R}^n$. A fim de que f seja contínua, é necessário e suficiente que a imagem inversa de $f^{-1}(A)$ de todo aberto $A \subset \mathbb{R}^n$ seja um conjunto aberto em X .*

Demonstração. (Necessário.) Se f é contínua e $A \subset \mathbb{R}^n$ é aberto tomemos um ponto $a \in f^{-1}(A)$. Então $f(a) \in A$. Pela definição de aberto, existe $\epsilon > 0$ tal que $B(f(a); \epsilon) \subset A$. Sendo f contínua existe $\delta > 0$ tal que $x \in X, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$. Isto quer dizer que $f(B(a; \delta) \cap X) \subset B(f(a), \epsilon) \subset A$, donde $B(a; \delta) \cap X \subset f^{-1}(A)$. Logo $f^{-1}(A)$ é aberto em X .

(Suficiente.) Se a imagem inversa por f de todo aberto de \mathbb{R}^n é aberto em X ; então, dados $a \in X$ e $\epsilon > 0$, como $B(f(a); \epsilon)$ é aberto, concluímos que $A = \{x \in X; |f(x) - f(a)| < \epsilon\}$ é aberto em X . Evidentemente que $a \in A$. Logo existe $\delta > 0$ de modo que $B(a; \delta) \cap X \subset A$. Isto quer dizer que $x \in X, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$, ou seja, que f é contínua no ponto a . Como $a \in X$ é qualquer, f é contínua. \square

Observação 2.32. O resultado continuaria válido se substituíssemos, no enunciado do Teorema 2.31, a expressão “todo aberto $A \subset \mathbb{R}^n$ ” por “todo conjunto $A \subset f(X)$, aberto em $f(X)$ ”.

Com efeito, um aberto em $f(X)$ tem a forma $A' \cap f(X)$ onde A' é aberto em \mathbb{R}^n e $f^{-1}(A' \cap f(X)) = f^{-1}(A')$. Logo, este novo enunciado, reduz-se ao que foi demonstrado.

2.2.7 Conjuntos fechados

Dizemos que ponto $a \in \mathbb{R}^n$ é aderente ao conjunto $X \subset \mathbb{R}^n$ quando existe uma seqüência de pontos $x_k \in X$ tais que $\lim x_k = a$.

Chamamos *fecho* do conjunto $X \subset \mathbb{R}^n$ ao conjunto \overline{X} formado por todos os pontos aderentes a X .

Definição 2.33. Um conjunto $X \subset \mathbb{R}^n$ chama-se fechado quando contém todos seus pontos aderentes, $X = \overline{X}$. Isto é, quando o limite de toda sequência convergente de pontos de $F \subset X$ é ainda um ponto de X .

O teorema abaixo resume as principais propriedades do fecho de um conjunto

Teorema 2.34. (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 $\mathbb{R}^n - F$ é aberto. Equivalentemente: $A \subset \mathbb{R}^n$ é aberto se, e somente se, $\mathbb{R}^n - A$ é fechado.

(c) O fecho de qualquer conjunto $X \subset \mathbb{R}^n$ é fechado. Em outras palavras: para todo $X \subset \mathbb{R}^n$ tem-se $\overline{\overline{X}} = \overline{X}$

Demonstração. (a) Se o ponto a é aderente a X então $a = \lim x_k$, com $x_k \in X$ para todo $k \in \mathbb{N}$. Portanto qualquer bola $B(a; r)$ contém pontos de X , a saber, todos os x_k com k suficientemente grande. Reciprocamente, se toda bola de centro a contém pontos de X , podemos escolher, para cada $k \in \mathbb{N}$, um ponto $x_k \in X$ que esteja na bola $B(a; \frac{1}{k})$, isto é, $|x_k - a| < \frac{1}{k}$. Então $\lim x_k = a$, logo a é aderente a X .

(b) Seja $A = \mathbb{R}^n - F$. Para todo $b \in A$ existe $r > 0$ tal que $B(b; r) \cap F = \emptyset$. Afirmamos que $B(b; r) \subset A$. Com efeito, se $y \in B(b; r)$ então $B(b; r)$ é um aberto contendo y e disjunto de F , logo $y \in \mathbb{R}^n - F = A$. Reciprocamente, se $X \subset \mathbb{R}^n$ é tal que $A = \mathbb{R}^n - X$ é um conjunto aberto, então $y \notin X \Rightarrow y \in A \Rightarrow B(y; r) \subset A$ para algum $r > 0 \Rightarrow B(y; r) \cap X = \emptyset \Rightarrow y$ não é aderente a X . Assim todo ponto aderente a X deve pertencer a X e conseqüentemente X é fechado.

(c) Se $x \in \mathbb{R}^n - \overline{X}$ (isto é, x não é aderente a X) então, pelo Teorema 2.34, existe uma bola $B = B(x; r)$ que não contém pontos de X , ou seja, $X \subset \mathbb{R}^n - B$. Logo $\overline{X} \subset \mathbb{R}^n - B$. Mas, pelo Teorema 2.34 acima, $\mathbb{R}^n - B$ é fechado; portanto $\overline{X} \subset \mathbb{R}^n - B$ ou, equivalentemente, $B \subset \mathbb{R}^n - \overline{X}$. Assim, todo ponto $x \in \mathbb{R}^n - \overline{X}$ é um ponto interior e $\mathbb{R}^n - \overline{X}$ é aberto. Segue-se que \overline{X} é fechado. \square

2.2.8 Conjuntos compactos

Dizemos que um conjunto $K \subset \mathbb{R}^n$ é compacto quando ele for fechado e limitado.

Em virtude do Teorema 2.26, um conjunto $K \subset \mathbb{R}^n$ é compacto se, e somente se, toda sequência de pontos $x_k \in K$ possui uma subsequência que converge para um ponto de K .

As propriedades a seguir decorrem imediatamente da definição:

1. K_1, \dots, K_p compactos em $\mathbb{R}^n \Rightarrow K_1 \cup \dots \cup K_p$ compacto.
2. A interseção de uma família qualquer de compactos $K_\lambda \subset \mathbb{R}^n$ é um conjunto compacto.
3. Se $K \subset \mathbb{R}^m$ e $L \subset \mathbb{R}^n$ são compactos então o produto cartesiano $K \times L \subset \mathbb{R}^{n+m}$ é compacto.
4. Propriedade de Cantor (Lima, 2018, p.45): Dada uma sequência decrescente $K_1 \supset \dots \supset K_k \supset \dots$ de compactos não vazios, a interseção $K = \bigcap_{k=1}^{\infty} K_k$ é compacta e não é vazia.

Teorema 2.35. *Seja $f : X \rightarrow \mathbb{R}^n$ contínua no conjunto $X \subset \mathbb{R}^m$. Para todo subconjunto compacto $K \subset X$, sua imagem $f(K)$ é compacta.*

Demonstração. Mostremos primeiro que $f(K)$ é fechado em \mathbb{R}^n . Seja $y \in \mathbb{R}^n$ aderente a $f(K)$. Então $y = \lim f(x_k), x_k \in K$ para todo $n \in \mathbb{N}$. Pela compacidade de K , uma subsequência (x_{k_i}) converge para um certo ponto $x \in K$. Segue que $y = \lim_{i \rightarrow \infty} f(x_{k_i}) = f(\lim x_{k_i}) = f(x)$, donde $y \in f(K)$. Agora mostremos que $f(K)$ é limitado. De fato, se não fosse, poderíamos obter um ponto $x_k \in K$ para cada $k \in \mathbb{N}$ de modo que $|f(x_k)| > k$. Então a sequência $(f(x_k))$ não admitiria subsequências convergentes. Mas (x_k) tem uma subsequência convergente, com $\lim_{i \rightarrow \infty} x_{k_i} = x \in K$. A continuidade de f nos fornece então $f(x) = f(\lim x_{k_i}) = \lim f(x_{k_i})$, uma contradição. \square

Corolário 2.36. *(Weierstrass) Toda função real contínua $f : K \rightarrow \mathbb{R}$, definida num compacto $K \subset \mathbb{R}^m$, atinge seu máximo e seu mínimo em K , isto é, existem pontos $x_0, x_1 \in K$ tais que $f(x_0) \leq f(x) \leq f(x_1)$ para qualquer $x \in K$.*

Com efeito $f(K) \subset \mathbb{R}$ é compacto, logo $y_0 = \inf f(K)$ e $y_1 = \sup f(K)$ pertencem a $f(K)$, isto é, existem pontos $x_0, x_1 \in K$ tais que $f(x_0) = y_0$ e $f(x_1) = y_1$. Então $f(x_0) \leq f(x) \leq f(x_1)$ para todo $x \in K$.

Corolário 2.37. *Seja $K \subset \mathbb{R}^m$ compacto. Toda aplicação contínua $f : K \rightarrow \mathbb{R}^n$ é fechada, isto é, $F \subset K \Rightarrow f(F) \subset \mathbb{R}^n$ fechado.*

2.2.9 Conexidade

Definição 2.38. *Uma cisão de um subconjunto $X \subset \mathbb{R}^n$ é uma decomposição $X = A \cup B$, onde $A \cap B = \emptyset$ e os conjuntos A, B são ambos abertos em X .*

Todo conjunto $X \subset \mathbb{R}^n$ admite pelo menos a cisão trivial $X = X \cup \emptyset$. Um exemplo de cisão não-trivial é $\mathbb{R} - \{0\} = (-\infty, 0) \cup (0, +\infty)$.

Um conjunto $X \subset \mathbb{R}^n$ chama-se *conexo* quando não admite outra cisão além da trivial. Assim, se X é conexo, $X = A \cup B$ com A e B disjuntos e abertos em X então $A = \emptyset$ ou $B = \emptyset$.

Quando existir uma cisão não-trivial $X = A \cup B$, diremos que X é desconexo.

Dada uma cisão $X = A \cup B$, temos $A = X - B$ e $B = X - A$, isto é, cada um dos conjuntos A e B é o complementar do outro em relação a X . Logo, cada um deles é aberto e fechado em relação a X . Poderíamos, portanto, definir uma cisão de X como uma decomposição $X = A \cup B$, onde A e B são disjuntos e fechados em X . Poderíamos também definir definir um conjunto conexo X como aquele cujos únicos subconjuntos simultaneamente abertos e fechados em X são \emptyset e o próprio X . Com efeito, dado $A \subset X$ aberto-fechado, $X = A \cup (X - A)$ é uma cisão.

Teorema 2.39. *A imagem de um conjunto conexo por uma aplicação contínua é um conjunto conexo.*

Demonstração. Sejam $X \subset \mathbb{R}^m$ conexo e $f : X \rightarrow \mathbb{R}^n$ contínua. Para toda cisão $f(X) = A \cup B$, o Teorema 2.31 (juntamente com a observação que a ele segue) garante que $X = f^{-1}(A) \cup f^{-1}(B)$ é uma cisão do conjunto conexo X . Logo $f^{-1}(A) = \emptyset$, (donde $A = \emptyset$), ou $f^{-1}(B) = \emptyset$, (donde $B = \emptyset$). \square

Teorema 2.40. *Um subconjunto $X \subset \mathbb{R}$ é conexo se, e somente se, é um intervalo.*

Demonstração. Todo intervalo de \mathbb{R} é, obviamente, um conjunto convexo e, portanto, é conexo.

Reciprocamente, suponhamos que $X \subset \mathbb{R}$ seja conexo e não vazio. Se X contém apenas um ponto, então X é um intervalo (degenerado). Caso contrário, sejam $a, b \in X$, tais que $a < b$. Então, dado c entre a e b , devemos ter $c \in X$. De fato, se não fosse assim, fazendo-se $A = X \cap (-\infty, c)$ e $B = X \cap (c, +\infty)$, a decomposição $X = A \cup B$ seria uma cisão não trivial de X , pois A e B são abertos não vazios de X , observe que $a \in A$ e $b \in B$. Logo X é um intervalo. \square

Corolário 2.41. *A imagem de um conjunto conexo $X \subset \mathbb{R}^n$ por uma função contínua real $f : X \rightarrow \mathbb{R}$ é um intervalo.*

A reformulação do Corolário acima constitui um clássico da topologia.

Teorema 2.42 (Teorema do Valor Intermediário). *Seja $f : X \rightarrow \mathbb{R}$ uma função contínua, definida num conjunto conexo $X \subset \mathbb{R}^n$. Se existem $a, b \in X$ e $d \in \mathbb{R}$ tais que $f(a) < d < f(b)$ então existe $c \in X$ de modo que $f(c) = d$.*

2.3 Funções reais de n variáveis

2.3.1 Derivadas Parciais

Ao se estudar funções reais de n variáveis, ou seja, definidas em subconjuntos do espaço \mathbb{R}^n , buscamos, para estas funções, uma noção de derivada que tenha propriedades análogas às de derivadas da função definida num intervalo. A ideia que se apresenta naturalmente é a de “derivada parcial”, que vamos expor.

Para efeito de derivação, onde se compara o acréscimo $f(a+h) - f(a)$ da função f com o acréscimo h dado ao ponto a , o domínio mais adequado para uma função é um subconjunto aberto $U \subset \mathbb{R}^n$ pois, neste caso, dado $a \in U$, tem-se ainda $a+h \in U$ para todo acréscimo suficientemente pequeno h .

Seja $f : U \rightarrow \mathbb{R}$ uma função real, definida num subconjunto aberto $U \subset \mathbb{R}^n$. Dado o ponto $a \in U$, a i -ésima derivada parcial de f no ponto a (onde $1 \leq i \leq n$) é o limite

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t},$$

quando este limite existir.

Dados o ponto $a \in U$ e o inteiro $i \in [1, n]$, a imagem do caminho $\lambda : \mathbb{R} \rightarrow \mathbb{R}^n$, $\lambda(t) = a + te_i$, é o que se chama “a reta que passa por a e é paralela ao i -ésimo eixo”. Como U é aberto existe $\epsilon > 0$ tal que $-\epsilon < t < \epsilon \Rightarrow \lambda(t) = a + te_i \in U$. A i -ésima derivada parcial de f no ponto a é a derivada, no ponto $t = 0$, da função $f \circ \lambda : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$, ou seja,

$$\frac{\partial f}{\partial x_i}(a) = (f \circ \lambda)'(0)$$

Podemos dizer que f , quando restrita ao segmento de reta aberto $J = (a - \epsilon e_i, a + \epsilon e_i)$, torna-se uma função real, $f(a + te_i)$, da variável real t e $\frac{\partial f}{\partial x_i}(a)$ é a derivada dessa função no ponto $t = 0$.

O comportamento da i -ésima derivada parcial $\frac{\partial f}{\partial x_i}$ ao longo de um segmento de reta contido no domínio de f e paralelo ao i -ésimo eixo dá informações sobre o crescimento de f ao longo de tal segmento.

O cálculo prático da i -ésima derivada parcial de uma função real $f(x_1, \dots, x_n)$ é realizado considerando as outras variáveis como constantes, exceto a i -ésima, e aplicando as regras habituais de derivação relativas a esta variável.

O comportamento da i -ésima derivada parcial $\frac{\partial f}{\partial x_i}$ ao longo de um segmento de reta contido no domínio de f e paralelo ao i -ésimo eixo nos fornece informações sobre o crescimento de f ao longo de tal segmento.

Dizemos que uma função $f : U \rightarrow \mathbb{R}$ não depende da i -ésima variável quando dados $a = (a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n)$ e $b = (a_1, \dots, a_{i-1}, y, a_{i+1}, \dots, a_n) \in U$, teremos $f(a) = f(b)$. Em outras palavras: se $a, b \in U$ então $f(a) = f(b)$.

O conjunto $U \subset \mathbb{R}^n$ será chamado de i -convexo quando $a, b \in U, a + te_i \Rightarrow [a, b] \subset U$.

Observação 2.43. Sejam $U \subset \mathbb{R}^n$ um aberto i -convexo e $f : U \rightarrow \mathbb{R}^n$ uma função tal que $\frac{\partial f}{\partial x_i}(x) = 0$ para todo $x \in U$. Então f é independente da i -ésima variável.

Com efeito, na observação acima, se $a, b \in U$, com $b = a + te_i$, então a função $\xi : [0, t] \rightarrow \mathbb{R}$, definida por $\xi(s) = f(a + se_i)$, possui derivada $\xi'(s) = \frac{\partial f}{\partial x_i}(a + se_i) = 0$ para todo $s \in [0, t]$, logo é constante, e assim $f(a) = \xi(0) = \xi(t) = f(b)$.

No plano, diz-se horizontalmente, ou verticalmente, convexo, em vez de 1-convexo ou 2-convexo, respectivamente.

Exemplo 2.44. Seja $X = \{(x, 0) \in \mathbb{R}^2; x \geq 0\}$ o semieixo positivo fechado das abscissas. O aberto $U = \mathbb{R}^2 - X$ é horizontalmente, mas não verticalmente convexo. A função $f : U \rightarrow \mathbb{R}$, definida por $f(x, y) = x^2$ quando $x > 0, y > 0$ e $f(x, y) = 0$ quando $x \leq 0$ ou $y \leq 0$, possui derivada parcial $\frac{\partial f}{\partial y} = 0$ em todos os pontos de U , mas f não é independente da segunda variável, y , pois se tomarmos $x > 0, y > 0$ teremos $f(x, y) = x^2 > 0$ e $f(x, -y) = 0$.

2.3.2 Derivadas direcionais

Como a derivada parcial fornece apenas informações sobre a função ao longo da reta paralela ao eixo quando desacompanhadas de hipóteses adicionais, tentamos estender a noção de derivada a outras direções além dessas. Assim o conceito de derivada direcional é alcançado.

Definição 2.45. Sejam $f : U \rightarrow \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n, a \in U$ e $v \in \mathbb{R}^n$. A derivada direcional de f no ponto a , segundo o vetor v é, por definição o limite

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

quando este limite existir.

Observe que as derivadas parciais tornam-se casos particulares das derivadas direcionais:

$\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i}(a)$ = derivada direcional de f no ponto a , segundo o vetor e_i .

A derivada direcional $\frac{\partial f}{\partial v}(a)$ é a derivada, no ponto $t = 0$, da função composta $f \circ \lambda : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$, onde $\lambda : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$ é o caminho retilíneo, $\lambda(t) = a + tv$, para o qual se tem, $\lambda(0) = a$ e $\lambda'(t) = v$ para todo t . Aqui $\epsilon > 0$ é escolhido tão pequeno de forma que a imagem de λ esteja totalmente contida em U .

2.3.3 Funções diferenciáveis

A noção de função diferenciável, que apresentaremos, é devida a Fréchet² e Stolz³. Ela constitui a extensão adequada do conceito de função derivável de uma só variável para funções de n variáveis.

Definição 2.46. Dada $f : U \rightarrow \mathbb{R}$, com $U \subset \mathbb{R}^n$, seja $a \in U$. Diremos que a função f é diferenciável no ponto a quando existirem as derivadas parciais $\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)$ e, além disso, para todo vetor $v = (\alpha_1, \dots, \alpha_n)$ tal que $a + v \in U$, tivermos

$$f(a + v) = f(a) + \frac{\partial f}{\partial x_1}(a) \cdot \alpha_1 + \dots + \frac{\partial f}{\partial x_n}(a) \cdot \alpha_n + r(v),$$

onde $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$.

O “resto” $r(v)$ é definido como sendo igual a $f(a + v) - f(a) - \sum \left(\frac{\partial f}{\partial x_i} \right) (a) \cdot \alpha_i$, na igualdade acima. Esta definição pode ser dada para qualquer função que possua derivadas parciais. A essência da definição de diferenciabilidade é que, tomando $r(v)$ desta maneira, tem-se $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$. Esta é a condição crucial, que deve ser verificada sempre que quisermos provar que uma função é diferenciável.

De $\lim_{v \rightarrow 0} \left(\frac{r(v)}{|v|} \right) = 0$, concluímos que $\lim_{v \rightarrow 0} r(v) = 0$ pois $r(v) = \left(\frac{r(v)}{|v|} \right) \cdot |v|$. Daí resulta que toda função diferenciável num ponto é contínua nesse ponto.

Em certas ocasiões, é preferível usar $p = p(v)$, em vez de $r(v)$, definida para os valores de v tais que $a + v \in U$, do seguinte modo: $p(v) = \frac{r(v)}{|v|}$ se $v \neq 0$ e $p(0) = 0$. Então f é diferenciável no ponto $a \in U$ se, e somente se, possui derivadas parciais nesse ponto e, para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ tal que $a + v \in U$, vale

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

²Maurice René Fréchet(1878-1973) foi um matemático francês eleito membro da *Académie des sciences* em 1958.

³Otto Stolz (1842 — 1905) foi um matemático austríaco conhecido pelas contribuições em análise matemática.

Portanto, f é diferenciável no ponto a se, e somente se, a função real $p = p(v)$, definida pela igualdade acima (se $v \neq 0$) e por $p(0) = 0$, é contínua no ponto $v = 0$.

Mostraremos agora que f admite derivada direcional segundo qualquer vetor $v = (\alpha_1, \dots, \alpha_n)$, e vale a fórmula

$$\frac{\partial f}{\partial v}(a) = \frac{\partial f}{\partial x_1} \cdot \alpha_1 + \dots + \frac{\partial f}{\partial x_n}(a) \cdot \alpha_n. \quad (2.22)$$

Com efeito, para todo t suficientemente pequeno, temos $a + tv \in U$. Pela definição de diferenciabilidade, (as derivadas parciais sendo consideradas no ponto a) temos:

$$f(a + tv) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot t\alpha_i + p(tv) \cdot |t| \cdot |v|$$

donde

$$\frac{f(a + tv) - f(a)}{t} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \alpha_i \pm p(tv) \cdot |v|$$

como $\lim_{t \rightarrow 0} p(tv) = 0$, segue a equação (2.22).

Um propriedade relevante das funções diferenciáveis é dada pela regra da cadeia.

Teorema 2.47. (Regra da cadeia) *Sejam $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ abertos, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ tal que $f(U) \subset V$ e cada função coordenada $f_k : U \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$. Seja ainda $g : V \rightarrow \mathbb{R}$ uma função diferenciável no ponto $b = f(a)$. Então a função composta $g \circ f : U \rightarrow \mathbb{R}$ é diferenciável no ponto a e suas derivadas parciais são*

$$\frac{\partial (g \circ f)}{\partial x_i}(a) = \sum_{k=1}^n \frac{\partial g}{\partial y_k}(b) \cdot \frac{\partial f_k}{\partial x_i}(a).$$

Demonstração. Seja U_0 o conjunto dos vetores $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ tais que $a + v \in U$. Para $v \in U_0$ e $k = 1, \dots, n$, temos

$$f_k(a + v) = f_k(a) + \sum_{i=1}^m \frac{\partial f_k}{\partial x_i}(b) \cdot \alpha_i + p_k \cdot |v|. \quad (2.23)$$

onde cada $p_k = p_k(v)$ é uma função definida em U_0 , contínua no ponto 0, que se anula quando $v = 0$. Consideramos a aplicação $w = (\beta_1, \dots, \beta_n) : U_0 \rightarrow \mathbb{R}^n$, contínua no ponto 0, cujas funções coordenada são definidas por

$$\beta_k(v) = \sum_{i=1}^m \frac{\partial f_k}{\partial x_i} \alpha_i + p_k \cdot |v|. \quad (2.24)$$

Adotando, por exemplo, a norma da soma, temos $\frac{|\alpha_i|}{|v|} \leq 1$ se $v \neq 0$, logo cada $\frac{|\beta_k|}{|v|}$,

e portanto a função $\frac{|w|}{|v|}$, é limitada numa vizinhança do ponto $v = 0$. Em virtude das equações (2.23) e (2.24) e da diferenciabilidade de g no ponto $b = f(a)$ podemos afirmar que vale para todo $v \in U_0$:

$$g \circ f(a + v) = g(b + w) = g(b) + \sum_{i=1}^n \frac{\partial g}{\partial y_k} \cdot \beta_i + \sigma \cdot |w|$$

onde $\sigma = \sigma(v)$ é uma função real contínua no ponto 0, que se anula no ponto $v = 0$, pois w também se anula nesse ponto. Usando a definição de β_k , obtemos

$$\begin{aligned} g \circ f(a + v) &= g \circ f(a) + \sum_{i=1}^n \frac{\partial g}{\partial y_k} \left[\sum_{i=1}^m \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + p_k \cdot |v| \right] + \sigma \cdot |w| \\ &= g \circ f(a) + \sum_{i=1}^m A_i \cdot \alpha_i + R \end{aligned}$$

onde $A_i = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \cdot \frac{\partial f_k}{\partial x_i}$ e $R = \sum_{k=1}^m \frac{\partial g}{\partial y_k} \cdot p_k \cdot |v| + \sigma \cdot |w|$.

Daí

$$\frac{R}{|v|} = \sum_{k=1}^m \frac{\partial g}{\partial y_k} p_k + \sigma \cdot \frac{|w|}{|v|}.$$

Quando $v \rightarrow 0$, sabemos que cada função $p_k \rightarrow 0$, que o quociente $\frac{|w|}{|v|}$ é limitado e que

$\lim_{v \rightarrow 0} \sigma = 0$. Segue-se que $\lim_{v \rightarrow 0} \left(\frac{R}{|v|} \right) = 0$. Isto mostra que $g \circ f$ é diferenciável no ponto a e suas derivadas parciais são os números A_i . \square

2.3.4 Funções de classe C^k

Seja $f : U \rightarrow \mathbb{R}$ uma função que possui as n derivadas parciais em todos os pontos do aberto $U \subset \mathbb{R}^n$. Ficam então definidas as n funções

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

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

De forma mais geral, dizemos que uma função $f : U \rightarrow \mathbb{R}$ é de classe C^k quando ela possuir derivadas parciais em todos os pontos de U e as funções

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \rightarrow \mathbb{R}$$

forem de classe C^{k-1} , aqui k é um inteiro maior que 0.

Para completar a definição indutiva, diremos que uma função $f : U \rightarrow \mathbb{R}$ é de classe C^0 quando ela for contínua. Usaremos a notação $f \in C^k$. Escrevemos também $f \in C^\infty$, e diremos que f é de classe C^∞ , quando $f \in C^k$ para todo $k \geq 0$.

Uma aplicação $f : U \rightarrow \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, diz-se de classe C^1 quando cada uma de suas funções coordenada $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ é de classe C^1 .

Muitas propriedades importantes das funções de classe C^1 resultam de serem elas diferenciáveis.

Teorema 2.48. *Toda função $f : U \rightarrow \mathbb{R}$ de classe C^1 é diferenciável.*

Demonstração. Por simplicidade, supomos $U \subset \mathbb{R}^2$. O caso geral se trata analogamente, apenas com uma notação mais elaborada. Fixemos $c = (a, b) \in U$ e tomemos $v = (h, k)$ tal que $c + v \in U$. Seja

$$r(v) = r(h, k) = f(a + h, b + k) - f(a, b) - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial y} \cdot k,$$

onde as derivadas são calculadas no ponto $c = (a, b)$. Podemos escrever

$$r(v) = f(a + h, b + k) - f(a, b + k) + f(a, b + k) - f(a, b) - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial y} \cdot k.$$

Pelo Teorema do Valor Médio para funções de uma variável real, existem $\theta_1, \theta_2 \in (0, 1)$ tais que

$$r(v) = f(a + \theta_1, b + k) \cdot h + \frac{\partial f}{\partial y}(a, b + \theta_2 k) \cdot k - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial y} \cdot k,$$

logo

$$\begin{aligned} \frac{r(v)}{|v|} &= \left[\frac{\partial f}{\partial x}(a + \theta_1, b + k) - \frac{\partial f}{\partial x}(a, b) \right] \frac{h}{\sqrt{h^2 + k^2}} \\ &\quad + \left[\frac{\partial f}{\partial y}(a, b + \theta_2 k) - \frac{\partial f}{\partial y}(a, b) \right] \frac{k}{\sqrt{h^2 + k^2}} \end{aligned}$$

Quando $v \rightarrow 0$ os termos dentro dos colchetes acima tendem a zero, pela continuidade das derivadas $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$. Além disso, os termos fora dos colchetes têm valor absoluto ≤ 1 . Portanto $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$ e então f é diferenciável. \square

Corolário 2.49. *Toda função de classe C^1 é contínua.*

Definição 2.50. *O gradiente de uma função diferenciável $f : U \rightarrow \mathbb{R}$ no ponto a é o vetor*

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

Se v é qualquer vetor de \mathbb{R}^n , a derivada direcional de f no ponto a , na direção v é, por definição

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

Outra notação para $\text{grad } f$ é ∇f . Estas definições permitem enunciar o seguinte corolário:

Corolário 2.51. *(TEOREMA DO VALOR MÉDIO) Dada $f : U \rightarrow \mathbb{R}$ 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*

$$\begin{aligned} f(a + v) - f(a) &= \frac{\partial f}{\partial x_i}(a + \theta v) = \langle \text{grad } f(a + \theta v), v \rangle \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + \theta v) \cdot \alpha_i \end{aligned}$$

onde $v = (\alpha_1, \dots, \alpha_n)$.

Demonstração. Com efeito, considerando o caminho retilíneo $\lambda : [0, 1] \rightarrow U$, dado por $\lambda(t) = f(a + tv)$, vemos que $f(a + v) - f(a) = (f \circ \lambda)(1) - (f \circ \lambda)(0)$. 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.

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

□

Para a demonstração do Teorema da Função implícita, parte central deste trabalho, será necessário a demonstração do lema a seguir.

Lema 2.52. *Sejam $I \subset \mathbb{R}^m, J \subset \mathbb{R}^k$ compacto, $f : I \times J \rightarrow \mathbb{R}^p$ contínua e $c \in \mathbb{R}^p$. Se $f^{-1}(c)$ é o gráfico de uma aplicação $\xi : I \rightarrow J$, (isto é, para cada $x \in I$ existe um único $y = \xi(x) \in J$ com $f(x, \xi(x)) = c$) então ξ é contínua.*

A afirmação de que $f^{-1}(c) \cap (I \times J)$ é o gráfico de uma função significa que, olhando para uma função $\xi : I \rightarrow J$, para cada $x \in I$, existe um único $y \in J$ com $f(x, y) = c$. Colocando-se $\xi(x) = y$; a função $\xi : I \rightarrow J$ diz-se “definida implicitamente” pela equação $f(x, y) = c$ no aberto $I \times J$ (ver Figura 2.2).

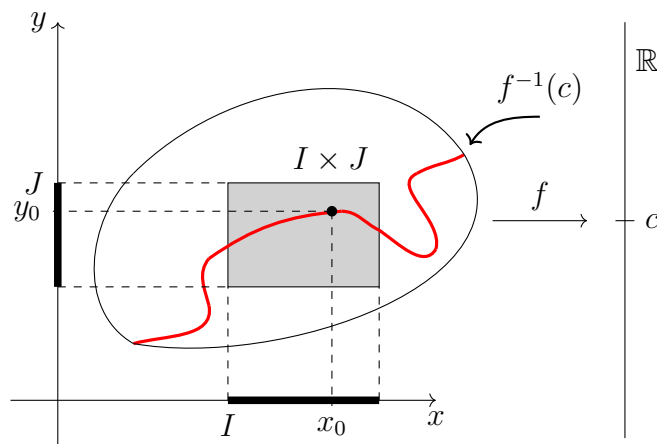


Figura 2.2: Curva definida implicitamente no aberto $I \times J$.

Demonstração. Dado $x_0 \in I$, seja $y_0 = \xi(x_0)$. Tomamos uma sequência de pontos $x_n \in I$, com $\lim x_n = x_0$, e queremos provar que $\lim \xi(x_n) = y_0$. Como a sequência $(\xi(x_n))$ é limitada (pois $\xi(x_n) \in J$ para todo n), basta provar que toda subsequência $\xi(x'_n)$, convergente em \mathbb{R}^k , tem limite y_0 . Ora, se for $\lim \xi(x'_n) = y$, deve ser $y \in J$ pois J é fechado. Como $f(x'_n, \xi(x'_n)) = c$ para todo n , temos $f(x_0, y) = \lim f(x'_n, \xi(x'_n)) = c$. Pela unicidade de y_0 , isto obriga $y = y_0$ e conclui a demonstração. □

Para o próximo resultado vamos considerar U um conjunto aberto de \mathbb{R}^n , uma função $F : U \rightarrow \mathbb{R}^m$ e $p \in U$. Escrevemos $F(p) = (F_1(p), \dots, F_m(p))$ com $F_i : U \rightarrow \mathbb{R}$ a i -ésima coordenada de F com $i = 1, \dots, m$. Dizemos que F é diferenciável no ponto p quando cada uma de suas funções coordenadas é diferenciável neste ponto.

Se este é o caso então, para todo $v = (\alpha_1, \dots, \alpha_n)$ tal que $p + v \in U$ e para cada $i = 1, \dots, m$, temos

$$F_i(p + v) - F_i(p) = \sum_{j=1}^n \frac{\partial F_i}{\partial x_j}(p) \cdot \alpha_j + r_i(v)$$

com $\lim_{v \rightarrow 0} \frac{r_i(v)}{|v|} = 0$.

A matriz Jacobiana de F no ponto p é:

$$JF(p) = \left[\frac{\partial F_i}{\partial x_j}(p) \right] \in M(n \times m)$$

A transformação linear $F'(p) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, cuja matriz em relação às bases canônicas de \mathbb{R}^n e \mathbb{R}^m é $JF(p)$ chama-se derivada da aplicação F no ponto p .

Lema 2.53. *Seja $f : U \rightarrow \mathbb{R}^n$, com $f \in C^1$, $U \subset \mathbb{R}^n$ aberto, e $p \in U$ satisfazendo $dJf(p) \neq 0$. Então, f restrita a alguma bola $B(p; r)$, com $r > 0$, é injetiva.*

Demonstração. Como f é de classe C^1 e a função $d : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$ é contínua e $dJf(p) = d\left(\frac{\partial f}{\partial x_j}(p)\right) \neq 0$, existe $r > 0$ tal que $d\left(\frac{\partial f_i}{\partial x_j}(\xi_{ij})\right)$ não seja zero, para todo ξ_{ij} em $B(p; r)$, onde $1 \leq i, j \leq n$. Sejam a e b distintos em $B(p; r)$. Pelo Corolário 2.51, para cada componente F_i , encontramos c_i no segmento \overline{ab} , dentro de $B(p; r)$, de modo que $f_i(b) - f_i(a) = \langle \nabla f_i(c_i), b - a \rangle$. Consequentemente,

$$\begin{pmatrix} f_1(b) - f_1(a) \\ \vdots \\ f_n(b) - f_n(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(c_1) & \cdots & \frac{\partial f_1}{\partial x_n}(c_1) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(c_n) & \cdots & \frac{\partial f_n}{\partial x_n}(c_n) \end{pmatrix} \begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{pmatrix}$$

Como $d\left(\frac{\partial f_i}{\partial x_j}(c_i)\right) \neq 0$ e $b - a \neq 0$, deduz-se que $f(b) \neq f(a)$. □

Antes de enunciar os próximos resultados precisamos das seguintes definições:

Definição 2.54. *Um homeomorfismo do conjunto $X \subset \mathbb{R}^m$ sobre um conjunto $Y \subset \mathbb{R}^n$ é uma bijeção contínua $f : X \rightarrow Y$ cuja inversa $f^{-1} : Y \rightarrow X$ também é contínua.*

Definição 2.55. *Sejam U contido em \mathbb{R}^m , V contido em \mathbb{R}^n , com U e V abertos. Uma aplicação $F : U \rightarrow V$ é chamada de difeomorfismo entre U e V quando é uma bijeção diferenciável, cuja inversa $g = f^{-1} : V \rightarrow U$ também é diferenciável.*

Se $Id_V : U \rightarrow V$ é um difeomorfismo, com $g = f^{-1} : V \rightarrow U$, então $g \circ f = id_U$ e $f \circ g = id_V$ resulta, pela Regra da Cadeia, que $g'(f(x)) \cdot f'(x) = Id_{\mathbb{R}^m}$ e $f'(x) \cdot g'(f(x)) = Id_{\mathbb{R}^n}$ para todo $x \in U$, portanto $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é um isomorfismo cujo inverso é $g'(f(x)) : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Em particular, $m = n$, ou seja, dois abertos em espaços euclidianos de dimensões diferentes não podem ser difeomorfos.

Vale ressaltar que nem todo homeomorfismo diferenciável é um difeomorfismo, isto é, tem inverso diferenciável. Um simples exemplo deste fato é a função $f : \mathbb{R} \rightarrow \mathbb{R}$, com $f(x) = x^3$. Como $f'(0) = 0$, a função inversa de f , $g(x) = x^{\frac{1}{3}}$, não é diferenciável no ponto $0 = f(0)$.

Uma aplicação diferenciável $f : U \rightarrow \mathbb{R}^m$, definida no aberto $U \subset \mathbb{R}^m$, chama-se 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-se daí que se $f : U \rightarrow \mathbb{R}^m$ é um difeomorfismo local então $f(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é um isomorfismo, para todo $x \in U$. O Teorema da Aplicação Inversa, que provaremos a frente, diz que quando $f \in C^1$ vale a recíproca: se $f'(x)$ é um isomorfismo para todo $x \in U$ então f é um difeomorfismo local.

Decorre da definição acima que um difeomorfismo local $f : U \rightarrow \mathbb{R}^m$ é uma aplicação aberta, isto é, a imagem $f(A)$ de qualquer aberto $A \subset U$ é um subconjunto aberto de \mathbb{R}^m . Com efeito, se tomarmos para cada $x \in A$ uma bola aberta $B_x \subset A$, com centro x , tal 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 $f(A) = f(\cup B_x) = \cup f(B_x) = \cup V_x$ é uma reunião de abertos, logo é um aberto.

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

Teorema 2.56. *Se o difeomorfismo $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^m$ é de classe C^k ($k \geq 1$) então seu inverso $g = f^{-1} : V \rightarrow U$ também é de classe C^k .*

Demonstração. Vamos utilizar indução em k . Para todo $y = f(x) \in V$, temos $g'(y) = [f'(x)]^{-1} = [f'(f^{-1}(y))]^{-1}$, portanto a aplicação $g : V \rightarrow \mathcal{L}(\mathbb{R}^m) = \mathbb{R}^{m^2}$ se exprime como a composta

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

onde Inv leva todo operador invertível $X : \mathbb{R}^m \rightarrow \mathbb{R}^m$ no seu inverso X^{-1} , $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m)$, f' leva todo ponto $x \in U$ na derivada (invertível) $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ e $f^{-1} : V \rightarrow U$ é aplicação inversa de f . Sabemos que $\text{Inv} \in C^\infty$. Portanto, se $f \in C^k$ então $f' \in C^{k-1}$ e, pela hipótese de indução, $f^{-1} \in C^{k-1}$, logo $g' \in C^{k-1}$, como composta de três aplicações de classe C^{k-1} . Por definição, isto significa que $g \in C^k$. \square

Teorema 2.57. *Seja $F : U \rightarrow \mathbb{R}^n$ de classe C^1 no aberto $U \subset \mathbb{R}^m$. Se, para algum $a \in U$, a derivada $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é injetiva então existem $\delta > 0$ e $c > 0$ tais que $B = B(a, \delta) \subset U$ e, para quaisquer $x, y \in B$ tem-se $|f(x) - f(y)| \geq c|x - y|$. Em particular, a restrição $F|_B$ é injetiva.*

Demonstração. A função $u \mapsto |f'(a) \cdot u|$ é positiva em todos os pontos u da esfera unitária S^{m-1} , a qual é compacta. Pelo Corolário 2.36, existe $c > 0$ tal que $|f'(a) \cdot u| \geq 2c$ para todo $u \in S^{m-1}$. Por linearidade, segue-se que $|f'(a) \cdot v| \geq 2c \cdot |v|$ para todo $v \in \mathbb{R}^m$. Para todo $x \in U$, escrevemos

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

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

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

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

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

Observamos que a aplicação r , acima definida, é de classe C^1 , com $r(a) = 0$. Pela continuidade de r' existe $\delta > 0$ tal que $|x - a| < \delta \Rightarrow x \in U$ e $|r'(x)| < c$. A desigualdade do valor médio, aplicada a r no conjunto convexo $B = B(x; \delta)$ nos assegura que se $x, y \in B$ então $|r(x) - r(y)| \leq c|x - y|$. Consequentemente, $x, y \in B \Rightarrow |f(x) - f(y)| \geq 2c|x - y| - c|x - y|$, ou seja, $|f(x) - f(y)| \geq c|x - y|$, como queríamos demonstrar. \square

Teorema 2.58. (DIFERENCIABILIDADE DO HOMEOMORFISMO INVERSO) *Seja $f : U \rightarrow V$ um homeomorfismo diferenciável entre os abertos $U, V \subset \mathbb{R}^m$. Se para algum $x \in U$, a derivada $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é um operador invertível então o homeomorfismo inverso $g = f^{-1} : V \rightarrow U$ é diferenciável no ponto $f(x)$, com $g'(f(x)) = [f'(x)]^{-1}$.*

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

$$\begin{aligned} w &= f(x + v) - f(x) = f'(x) \cdot v + r(v) \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0 \text{ e} \\ v &= g(f(x + v)) - g(f(x)) = g(y + w) - g(y). \end{aligned}$$

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

$$g(y + w) - g(y) = f'(x)^{-1} \cdot w + s(w) \tag{2.25}$$

e mostremos que $\lim_{w \rightarrow 0} \frac{s(w)}{|w|} = 0$. Entretanto na igualdade 2.25 com as expressões de v e w acima obtidas, vem:

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

ou seja,

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

donde

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

isto é:

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

Quando $w \rightarrow 0$, tem-se $v \rightarrow 0$ pela continuidade da g e $\frac{r(v)}{|v|} \rightarrow 0$. Além disso, pelo Teorema 2.57, existem $\delta > 0$ e $c > 0$ tais que $|v| < \delta$ implica

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

Assim, $\lim_{w \rightarrow 0} \frac{s(w)}{|w|} = 0$. \square

Vamos utilizar o lema a seguir para concluir a demonstração do Teorema da Aplicação Inversa.

Lema 2.59. *Sejam $U \subset \mathbb{R}^m$ aberto e $g : U \rightarrow \mathbb{R}^n$ diferenciável no ponto $a \in U$, com $g'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ sobrejetiva. Se a é um ponto de mínimo local de $|g(x)|, x \in U$, então $g'(a) = 0$.*

Demonstração. Se a é um ponto de mínimo local para $|g(x)|$, será também um ponto de mínimo local para a função $\varphi : U \rightarrow \mathbb{R}$, definida por $\varphi(x) = |g(x)|^2 = \langle g(x), g(x) \rangle$, logo $\varphi'(a) = 0$. Mas, como $\varphi'(a) \cdot v = 2 \langle g'(a) \cdot v, g'(a) \rangle$, isto significa que $g(a)$ é ortogonal à imagem de $g'(a)$, a qual é \mathbb{R}^n . Logo $g(a) = 0$. \square

3 Teorema da Função Implícita e Teorema da Função Inversa

Estes teoremas são de grande relevância na área de análise matemática para funções de n variáveis e aplicações diversas envolvendo equações diferenciais e equações algébricas não lineares. Utilizamos como referências principais deste capítulo Lima (2018) e De Oliveira (2014).

3.1 Teorema da Função Implícita

Neste momento, apresentamos a demonstração do Teorema da Função Implícita em \mathbb{R}^2 , seguindo Lima (2018), a fim de fixar ideias. Posteriormente, apresentamos o Teorema da Função implícita no \mathbb{R}^n .

Dada $f : U \rightarrow \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^2$, e fixado $c \in \mathbb{R}$, dizemos que a equação $f(x, y) = c$ define y implicitamente como função de x quando existe uma função $\xi : I \rightarrow \mathbb{R}$, definida num intervalo $I \subset \mathbb{R}$, tal que $f(x, y) = c \Leftrightarrow y = \xi(x)$. Isto quer dizer que $f^{-1}(c)$ é o gráfico da função ξ .

Exemplo 3.1. É mais comum que uma equação do tipo $f(x, y) = c$ defina y em função de x , ou x em função de y , apenas localmente. Por exemplo, seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x, y) = x^2 + y^2$ e tomemos $c = 1$. A equação $x^2 + y^2 = 1$ não define y como função de x (nem x como função de y). Por exemplo, para cada $x \in (-1, 1)$ existem dois valores de y tais que $x^2 + y^2 = 1$. Mas se tomarmos $U_1 = \{(x, y) \in \mathbb{R}^2; y > 0\}$ e $U_2 = \{(x, y) \in \mathbb{R}^2; y < 0\}$, $U_3 = \{(x, y) \in \mathbb{R}^2; x > 0\}$ e $U_4 = \{(x, y) \in \mathbb{R}^2; x < 0\}$ então a equação $x^2 + y^2 = 1$ equivale a $y = \sqrt{1 - x^2}$, quando $(x, y) \in U_1$, equivale a $y = -\sqrt{1 - x^2}$, quando $(x, y) \in U_2$, a $x = \sqrt{1 - y^2}$, quando $(x, y) \in U_3$ e a $x = -\sqrt{1 - y^2}$, quando $(x, y) \in U_4$. Como o círculo $S^1 = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 = 1\}$, conjunto de todas as soluções (x, y) da equação $x^2 + y^2 = 1$, está contido na reunião dos quatro abertos U_1, U_2, U_3 e U_4 , dizemos que a equação $x^2 + y^2 = 1$ define localmente y como função de x , ou x como função de y . Isto quer dizer que cada solução (x_0, y_0) desta equação está contida em algum aberto U_i tal que $f^{-1}(1) \cap U_i$ é o gráfico de uma função $x = \xi(y)$, ou $y = \xi(x)$.

Vale observar que uma equação do tipo $f(x, y) = c$ pode não definir nenhuma função, basta que c não esteja no conjunto imagem da função. Por exemplo, $x^2 + y^2 + 1 = 0$ não possui soluções reais (x, y) . Logo, não define y como função de x nem x como função de y . Mesmo que $f(x, y) = c$ possua soluções, elas podem não definir funções. É o caso de $x^2 + y^2 = 0$; a única solução é $(0, 0)$, que obviamente não é gráfico de uma função definida num intervalo não-degenerado. Outro exemplo é o seguinte: $(0, 0)$ é

solução de $x^2 + y^2 = 0$, mas para nenhum aberto V contendo a origem, a intersecção $f^{-1}(0) \cap V$ é o gráfico de uma função $y = \xi(x)$ ou $x = \zeta(y)$, pois tal intersecção contém sempre dois segmentos de reta de inclinação ± 1 que se cortam na origem.

Depois destes exemplos, enunciamos o teorema:

Teorema 3.2. (TEOREMA DA FUNÇÃO IMPLÍCITA) *Sejam $f : U \rightarrow \mathbb{R}$ uma função de classe C^k ($k \geq 1$), definida num aberto $U \subset \mathbb{R}^2$, e $(x_0, y_0) \in U$ tal que $f(x_0, y_0) = c$, $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Então existe um retângulo aberto $I \times J$, de centro (x_0, y_0) , tal que $f^{-1}(c) \cap (I \times J)$ é o gráfico de uma função $\xi : I \rightarrow J$, de classe C^k . Tem-se $\xi'(x) = -\frac{\partial f/\partial x}{\partial f/\partial y}$ estas derivadas sendo calculadas no ponto $(x, \xi(x))$.*

A afirmação $f^{-1}(c) \cap (I \times J)$ é o gráfico de uma função $\xi : I \rightarrow J$ significa que para cada $x \in I$, existe um único $y \in J$ com $f(x, y) = c$. Põe-se $y = \xi(x)$; a função $\xi : I \rightarrow J$ diz-se “definida implicitamente”, no aberto $I \times J$, pela equação $f(x, y) = c$.

Demonstração. Para fixar ideias, suponhamos $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $\frac{\partial f}{\partial y}$ é contínua, existem $\delta > 0$ e $\epsilon > 0$ tais que, pondo $I = (x_0 - \delta, x_0 + \delta)$ e $J = (y_0 - \epsilon, y_0 + \epsilon)$, temos $I \times \bar{J} \subset U$ e $\frac{\partial f}{\partial y}(x, y) > 0$ para todo $(x, y) \in I \times \bar{J}$. Então, para todo $x \in I$, a função $y \mapsto f(x, y)$ é estritamente crescente no intervalo \bar{J} . Em particular, como $f(x_0, y_0) = c$, temos $f(x_0, y_0 - \epsilon) < c$ e $f(x_0, y_0 + \epsilon) > c$. Pela continuidade de f , podemos supor δ suficientemente pequeno de modo que, para todo $x \in I$, tenhamos $f(x, y_0 - \epsilon) < c$ e $f(x, y_0 + \epsilon) > c$. Em virtude do Teorema 2.42 existe, para cada $x \in I$, um único $y = \xi(x) \in \bar{J}$ tal que $f(x, y) = c$. Tem-se obrigatoriamente $y \in J$, portanto $f^{-1}(c) \cap (I \times \bar{J}) = f^{-1}(c) \cap (I \times J)$ é o gráfico de uma função $\xi : I \rightarrow J$. Vamos mostrar que ξ é de classe C^k , ou seja, que existe $\xi'(x)$ para todo $x \in I$ e que $\xi' : I \rightarrow \mathbb{R}$ é de classe C^{k-1} .

Escrevendo $k = \xi(x+h) - \xi(x)$, temos $\xi(x+h) = \xi(x) + k$, logo $f(x+h, \xi(x) + k) = f(x, \xi(x)) = c$. Pelo Corolário 2.51, existe θ , com $0 < \theta < 1$, tal que

$$\begin{aligned} 0 &= f(x+h, \xi(x) + k) - f(x, \xi(x)) = \\ &= \frac{\partial f}{\partial x}(x + \theta h, \xi(x) + \theta k) \cdot h + \frac{\partial f}{\partial y}(x + \theta h, \xi(x) + \theta k) \cdot k \end{aligned}$$

Assim,

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

Segundo o Lema 2.52, ξ é contínua. Isto significa que $\lim_{h \rightarrow 0} k = 0$. A continuidade das derivadas parciais de f nos dá portanto

$$\xi'(x) = \lim_{h \rightarrow 0} \frac{\xi(x+h) - \xi(x)}{h} = -\frac{\frac{\partial f}{\partial x}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}.$$

Se $f \in C^1$, sendo $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ e ξ contínuas, esta fórmula mostra que ξ' é contínua. Logo, $\xi \in C^1$. Se $f \in C^2$ então $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ e ξ são de classe C^1 . A fórmula que dá ξ' mostra

então que ξ' é também de classe C^1 , isto é, $\xi \in C^2$. E assim por diante: se $f \in C^k$, então $\xi \in C^k$. \square

Trataremos agora de funções implícitas com um número qualquer de variáveis. No teorema abaixo representaremos os pontos de \mathbb{R}^{n+1} por pares (x, y) , onde $x \in \mathbb{R}^n$ e $y \in \mathbb{R}$.

Teorema 3.3 (TEOREMA DA FUNÇÃO IMPLÍCITA). *Seja $f : U \rightarrow \mathbb{R}$ uma função de classe C^k ($k \geq 1$), definida num aberto $U \subset \mathbb{R}^{n+1}$. Se um ponto $p = (x_0, y_0) \in U$ é tal que $f(p) = c$ e $\frac{\partial f}{\partial y}(p) \neq 0$, então existem uma bola $B = B(x_0, \delta) \subset \mathbb{R}^n$ e um intervalo $J = (y_0 - \epsilon, y_0 + \epsilon)$ tais que $f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi : B \rightarrow J$, de classe C^k . Para todo $x \in B$, tem-se*

$$\frac{\partial \xi}{\partial x_i}(x) = -\frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}, \quad (i = 1, \dots, n).$$

A função $y = \xi(x)$ diz-se “definida implicitamente pela equação $f(x, y) = c$ ”. A afirmação de que $f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função significa que, para cada $x \in B$ existe um único $y = \xi(x) \in J$ tal que $f(x, y) = c$. Evidentemente, $\xi(x_0) = y_0$.

Demonstração. Não há diferença essencial entre o caso geral e o já demonstrado, em que $n = 1$. No entanto, vamos repetir a prova. Para fixar ideias, seja $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $\frac{\partial f}{\partial y}$ é contínua, existem $\delta > 0, \epsilon > 0$ tais que, pondo $B = B(x_0; \delta)$ e $J = (y_0 - \epsilon, y_0 + \epsilon)$ temos $B \times \bar{J} \subset U$ e $\frac{\partial f}{\partial y} > 0$ para todo ponto $(x, y) \in B \times \bar{J}$. Então, para todo $x \in B$, a função $y \rightarrow f(x, y)$ é estritamente crescente no intervalo $\bar{J} = [y_0 - \epsilon, y_0 + \epsilon]$. Como $f(x_0, y_0) = c$, segue-se que $f(x_0, y_0 - \epsilon) < c$ e $f(x_0, y_0 + \epsilon) > c$. Pela continuidade de f , podemos supor δ tão pequeno que, para todo $x \in B$, tenhamos $f(x, y_0 - \epsilon) < c$ e $f(x, y_0 + \epsilon) > c$. Em virtude do Teorema 2.42, existe, para cada $x \in B$, um único $y = \xi(x) \in \bar{J}$ tal que $f(x, y) = c$. Tem-se obrigatoriamente $y \in J$, logo $f^{-1}(c) \cap (B \times \bar{J}) = f^{-1}(c) \cap (B \times J)$ é o gráfico de um função $\xi : B \rightarrow J$ a qual, pelo Lema 2.52 é contínua. Mostraremos agora que, em todo ponto $x \in B$, existem as derivadas parciais de ξ . Com efeito, pondo $k = k(t) = \xi(x + te_i) - \xi(x)$, temos $\xi(x + te_i) = \xi(x) + k$, logo $f(x + te_i, \xi(x) + k) = f(x, \xi(x)) = c, \forall t \in (-\delta, \delta)$.

Pelo Teorema 2.51, para todo $t \in (-\delta, \delta)$ existe $\theta = \theta(t) \in (0, 1)$ tal que

$$\begin{aligned} 0 &= f(x + te_i, \xi(x) + k) - f(x, \xi(x)) = \\ &= \frac{\partial f}{\partial x_i}(x + \theta te_i, \xi(x) + \theta k) \cdot t + \frac{\partial f}{\partial y}(x + \theta te_i, \xi(x) + \theta k) \cdot k. \end{aligned}$$

Logo

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

Pela continuidade de ξ ,

$$\lim_{t \rightarrow 0} k(t) = 0$$

A continuidade das derivadas parciais de f nos dá então

$$\frac{\partial \xi}{\partial x_i}(x) = \lim_{t \rightarrow 0} \frac{\xi(x + te_i) - \xi(x)}{t} = - \frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}.$$

Sendo $f \in C^1$, resulta desta fórmula que as derivadas parciais de ξ são contínuas, logo $\xi \in C^1$. Se for $f \in C^2$, então suas derivadas parciais são de classe C^1 . Como já temos $\xi \in C^1$, resulta ainda da fórmula acima que as derivadas parciais de ξ são de classe C^1 , logo $\xi \in C^2$. E assim por diante: $f \in C^k \Rightarrow \xi \in C^k$. \square

Considerando o aberto $V = B \times J \subset \mathbb{R}^{n+1}$, o teorema acima diz que, nas condições das hipóteses, tem-se

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

Em outras palavras, $f^{-1}(c) \cap V$ é o gráfico da função $\xi : B \rightarrow J$.

Observação 3.4. Evidentemente, não há nada de especial quanto a última coordenada, exceto simplificar a escrita na demonstração. Se, para algum inteiro $i \in [1, n+1]$, tivermos $\frac{\partial f}{\partial x_i}(z_0) \neq 0$ onde $z_0 \in U$ e $f(z_0) = c$, existirá um aberto $V \ni z_0$, tal que, para $z \in V$, a equação $f(z) = c$ definirá $x_i = \xi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$ como função das outras n coordenadas e $f^{-1}(c) \cap V$ será o gráfico dessa função ξ , de classe C^k . De modo geral, se $\text{grad} f(z_0) \neq 0$ e $f(z_0) = c$ então existe $V \ni z_0$ aberto tal que $f^{-1}(c) \cap V$ é o gráfico de uma função real de n variáveis, de classe C^k .

3.2 Teorema da Função Inversa

Fizemos a demonstração deste importante teorema da Análise Real como LIMA (2004). Este teorema permite estabelecer a existência de uma função inversa, f^{-1} , para uma função f continuamente diferenciável, embora possa ser necessário fazer uma boa restrição para o domínio, ou seja, estabelecer uma relação localmente na vizinhança de um ponto da função f .

Para realizar a demonstração deste teorema vamos nos preocupar com três pontos importantes. O primeiro ponto da demonstração é assegurar que a função seja localmente injetiva, isso é possível através do Teorema 2.57. O segundo ponto é garantir que a inversa na vizinhança também seja diferenciável, utilizaremos o Teorema 2.58 da Diferenciabilidade do Homeomorfismo Inverso. Por fim, precisamos garantir que a relação seja sobrejetiva.

Teorema 3.5. (TEOREMA DA FUNÇÃO INVERSA) *Seja $f : U \rightarrow \mathbb{R}^m$ de classe C^k ($k \geq 1$) no aberto $U \subset \mathbb{R}^m$. Se $a \in U$ é tal que $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é invertível então existe uma bola aberta $B = B(a; \delta) \subset U$ tal que a restrição $f|_B$ é um difeomorfismo sobre um aberto $V \ni f(a)$.*

Demonstração. Diminuindo δ , se necessário, no Teorema 2.57 podemos admitir que $\bar{B} = \bar{B}[a; \delta] \subset U$ e que f é injetiva no conjunto compacto \bar{B} , logo é um homeomorfismo de B sobre $f(B)$. Além disso, como $f'(x)$ depende continuamente de x e todo operador linear suficientemente próximo de um invertível é também invertível, podemos supor

que, para todo $x \in B$, a derivada $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é um isomorfismo. Pelo Teorema 2.58, basta então mostrar que $f(B) \subset \mathbb{R}^m$ é aberto. Seja então $q = f(p)$, $p \in B$. Chamando de $S = S[a, \delta]$ a esfera que é a fronteira de \overline{B} , a injetividade de $f|_{\overline{B}}$ assegura que $q \notin f(S)$, logo existe $\epsilon > 0$ tal que $|f(x) - q| \geq 2\epsilon$ para todo $x \in S$, pois $f(S)$ é compacto. Afirmamos que $B(q; \epsilon) \subset f(B)$. Com efeito, se $y \in B(q; \epsilon)$, então, pondo $g(x) = f(x) - y$, o mínimo de $|g(x)|$, quando x varia no compacto \overline{B} , não é atingido num ponto $x \in S$ pois $x \in S \Rightarrow |f(x) - y| \geq \epsilon$ enquanto $|f(p) - y| = |q - y| < \epsilon$, com $p \in B$. Assim, o mínimo de $|f(x) - y|$, $x \in \overline{B}$ é atingido num ponto $x_0 \in B$. Pelo Lema 2.59, isto implica que esse mínimo é zero, portanto $y = f(x_0)$, donde $y \in f(B)$, ou seja, $B(q; \epsilon) \subset f(B)$. \square

Outra demonstração do Teorema da Função Inversa

Para esta demonstração utilizamos como referência De Oliveira (2014). A demonstração ocorre como consequência do Teorema da Função Implícita.

Teorema 3.6. (TEOREMA DA FUNÇÃO INVERSA) *Seja $F : U \rightarrow \mathbb{R}^n$, uma função de classe C^1 , definida num aberto $U \subset \mathbb{R}^n$, e $p \in U$ tal que $JF(p)$ é invertível. Então, existe um conjunto aberto X contendo p , um aberto Y contendo $F(p)$, e uma função $G : Y \rightarrow X$ de classe C^1 que satisfaz $F(G(y)) = y$, para todo $y \in Y$, e $G(F(x)) = x$, para todo $x \in X$. Além disso,*

$$JG(y) = JF(G(y))^{-1}, \text{ para todo } y \in Y.$$

Demonstração. Primeiramente vamos mostrar que nestas condições, a função é bijetora e terá inversa, em seguida, mostramos como conseguir a fórmula de diferenciação.

Escolhendo U , se necessário, pelo Lema 2.53, podemos assumir que F é injetiva. A função

$$\xi(x, y) = F(x) - y$$

para (x, y) em $U \times \mathbb{R}^n$, é de classe C^1 , com $\xi(p, F(p)) = 0$, e satisfaz $\frac{\partial \xi}{\partial x}(p, F(p)) = JF(p)$. Do Teorema 3.3 segue que existe um aberto Y contendo $F(p)$ e uma função $G : Y \rightarrow U$ de classe C^1 tal que $\xi(G(y), y) = F(G(y)) - y = 0$, para todo $y \in Y$. Isto é,

$$F(G(y)) = y, \text{ para todo } y \in Y.$$

Portanto, o conjunto Y está contido na imagem de F . Como F é contínua e injetiva, o conjunto $X = F^{-1}(Y)$ é aberto, contém p e F é uma bijeção de X em Y .

A identidade $F(G(y)) = y$, para todo $y \in Y$, implica que G aplica Y em X . Como F é bijetiva de X em Y , a função G é bijetiva de Y em X .

Para a fórmula da diferenciação: Se escrevermos $F(x) = (F_1(x), \dots, F_n(x))$ e $G(y) = (G_1(y), \dots, G_n(y))$. Derivando $(G_1(F(x)), \dots, G_n(F(x)))$ obtemos

$$\sum_{k=1}^n \frac{\partial G_i}{\partial y_k} \frac{\partial F_k}{\partial x_j} = \frac{\partial x_i}{\partial x_j} = \begin{cases} 1, & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$$

\square

3.3 Aplicações

Nesta seção apresentamos alguns exemplos de aplicações. Vamos começar com um exemplo de aplicação do Teorema da Função Implícita.

Exemplo 3.7. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$. Para todo $(x, y) \in \mathbb{R}^2$, temos $\frac{\partial f}{\partial x}(x, y) = 2x$ e $\frac{\partial f}{\partial y}(x, y) = 2y$. A equação $x^2 + y^2 = c$ define o conjunto vazio quando $c < 0$. O Teorema 3.2 não se aplica, pois não existe o ponto (x_0, y_0) tal que $f(x_0, y_0) = c$. Quando $c = 0$ a equação $x^2 + y^2 = 0$ é satisfeita apenas quando $x = y = 0$. Agora existe (x_0, y_0) mas $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Quando $c > 0$, a equação $x^2 + y^2 = c$ define a circunferência de centro na origem e raio \sqrt{c} , a qual não é gráfico de função alguma do tipo $y = \xi(x)$ nem $x = \zeta(y)$, pois há retas verticais e horizontais que a cortam em dois pontos. Mas se considerarmos os abertos

$$V_1 = \{(x, y) \in \mathbb{R}^2; y > 0\}, \quad V_2 = \{(x, y) \in \mathbb{R}^2; y < 0\}$$

$$V_3 = \{(x, y) \in \mathbb{R}^2; x > 0\}, \quad V_4 = \{(x, y) \in \mathbb{R}^2; x < 0\}$$

veremos que $f^{-1}(c) \cap V_1$ e $f^{-1}(c) \cap V_2$ são gráficos das funções $\xi_1, \xi_2 : (-1, 1) \rightarrow \mathbb{R}$ dadas por $\xi_1(x) = \sqrt{1 - x^2}$, $\xi_2(x) = -\sqrt{1 - x^2}$, enquanto $f^{-1}(c) \cap V_3$ e $f^{-1}(c) \cap V_4$ são gráficos de $\xi_3, \xi_4 : (-1, 1) \rightarrow \mathbb{R}$ dadas por $\xi_3(y) = \sqrt{1 - y^2}$, $\xi_4(y) = -\sqrt{1 - y^2}$. Assim, em V_1 e V_2 a equação $x^2 + y^2 = c$ (com $c > 0$) define implicitamente y como função de x enquanto em V_3 e V_4 define x como função de y . Evidentemente, salvo na vizinhança dos quatro pontos $(\pm\sqrt{c}, 0), (0, \pm\sqrt{c})$ tem-se a opção de tomar y como função de x ou x como função de y .

O próximo exemplo é uma aplicação do Teorema da Função Inversa.

Exemplo 3.8. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$f(x, y) = (e^x \cos y, e^x \sin y)$$

Evidentemente f é de classe C^∞ . Sua matriz jacobiana é

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

Como o $dJf(x, y) = e^x \neq 0$, logo $f'(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ é um isomorfismo, para todo $(x, y) \in \mathbb{R}^2$. Pelo Teorema da Função Inversa f é um difeomorfismo local (não é global pois $f(x, y + 2\pi) = f(x, y)$), ou seja, f admite inversa local em qualquer ponto de \mathbb{R}^2 .

Exemplo 3.9. Dadas as matrizes $\mathbf{x}, \mathbf{m} \in M(n \times n)$, dizemos que \mathbf{x} é uma raiz quadrada de \mathbf{m} quando $\mathbf{x}^2 = \mathbf{m}$. Nem toda matriz \mathbf{m} possui raiz quadrada: como $d(\mathbf{x}^2) = (d\mathbf{x})^2$, uma condição necessária é que $d\mathbf{m} \geq 0$. Mas esta condição não é suficiente pois é fácil ver que, embora a matriz

$$\mathbf{m} = \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix}$$

tenha determinante positivo, não existe $\mathbf{x} \in M(2 \times 2)$ tal que $\mathbf{x}^2 = \mathbf{m}$. O Teorema 3.5 pode ser usado para mostrar que toda matriz próxima da identidade \mathbf{I}_n tem raiz quadrada. Com efeito, consideremos a aplicação $f : M(n \times n) \rightarrow M(n \times n)$, $f(\mathbf{x}) = \mathbf{x}^2$, de classe C^∞ . Sua derivada num ponto $\mathbf{x} \in M(n \times n)$ é a transformação linear $f'(\mathbf{x}) : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$, dada por $f'(\mathbf{x}) \cdot \mathbf{m} = \mathbf{m} \cdot \mathbf{x} + \mathbf{x} \cdot \mathbf{m}$. Em particular, para $\mathbf{x} = \mathbf{I}_n$, tem-se $f'(\mathbf{I}_n) \cdot \mathbf{m} = 2\mathbf{m}$, logo $f'(\mathbf{I}_n) : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ é um isomorfismo. Segue do Teorema 3.5 que existe um aberto U em $M(n \times n)$, contendo a matriz identidade, restrita ao qual f é um difeomorfismo sobre o aberto $V = f(U)$. Assim, para toda matriz $\mathbf{y} \in V$, existe uma única matriz $\mathbf{x} = \sqrt{\mathbf{y}} \in U$ tal que $\mathbf{x}^2 = \mathbf{y}$. Além disso, a aplicação $f^{-1} : V \rightarrow U$, $y \mapsto \sqrt{y}$ é de classe C^∞ .

4 Propostas de atividade para trabalhar funções reais de várias variáveis no Ensino Médio

As propostas de atividades que seguem foram pensadas de forma a motivar os alunos, aproximando as ferramentas existente com reflexões cotidianas que valorizem o conhecimento científico. Com essas propostas visamos proporcionar ao professor problemas em que eles possam, além de trabalhar a sua resolução, explorá-los para a formação cidadã do aluno, o que acreditamos ser uma boa forma de mostrar a importância da matemática e torná-la prazerosa.

4.1 Atividade 1: Equilibrando as finanças

Esta atividade tem o propósito de apresentar aos estudantes uma função real de várias variáveis.

Propor aos alunos para, em grupos, criarem uma planilha para gerenciar a vida financeira de uma família, logo levantarão hipóteses sobre ganhos, gastos e economias. Socializar e discutir as ideias enriquecerá a atividade. Após esta etapa, propor ao aluno anotar gastos que ele tem em sua residência, nominalmente e não monetário. Aparecerão gastos que esquecemos no nosso cotidiano como: bebidas, lanches, roupas, jogos, etc.

Retomar a planilha realizada por eles, colocando estes gastos no planejamento da família e descobrir se houve déficit orçamentário, ou não. Será que algum planejamento prevê a criação de uma poupança, seja para a aposentadoria, ou outro motivo que o cidadão considere importante?

Um primeiro passo pode ser compreender a expressão entre o(s) ganho(s), gastos e valor poupado de uma pessoa economicamente ativa. Para esta atividade vamos considerar a obtenção da renda (ou ganhos) para o mês de forma fixa. No entanto, sabemos que a renda pode variar com outros valores como, comissões, férias, décimo terceiro e possíveis horas extras. Além disso, para pessoas que são empreendedoras, e possuem seu próprio negócio, esta renda pode variar. Sendo assim, x denotará o salário, y_n o n -ésimo gasto e z o valor poupado. Assim, depara-se com uma equação do tipo:

$$x - (y_1 + y_2 + \dots + y_n) - z = 0$$

ou ainda

$$x - y - z = 0 \tag{4.1}$$

onde

$$y = \sum_{n=1}^n y_n$$

Discutir a equação (4.1) com os alunos auxiliará na compreensão matemática de função real de várias variáveis e na aprendizagem de conceitos de educação financeira, permitindo planejar-se melhor para o futuro.

Desta forma podemos introduzir a noção de função real de várias variáveis para os alunos da seguinte forma:

$$z = x - y = x - \sum_{n=1}^n y_n = x - (y_1 + y_2 + \dots + y_n)$$

ou ainda

$$z = f(x, y_1, y_2, \dots, y_n) = x - (y_1 + y_2 + \dots + y_n)$$

onde $z = f(x, y_1, y_2, \dots, y_n)$ representa a poupança em função do salário mensal e dos gastos mensais.

4.2 Atividade 2: Computação gráfica: converter imagem colorida em monocromática

Esta atividade também tem o propósito de apresentar aos estudantes função real de várias variáveis.

Primeiramente, precisamos entender o funcionamento do sistema RGB. Sua nomenclatura vem das siglas da cor vermelha, verde e azul, em inglês: *red*, *green* e *blue*. Essas três cores são as primárias para os aparelhos eletrônicos e a combinação na tela enviam a informação para nosso cérebro tons diferentes.

No *software Paint*, incluso como um acessório no sistema operacional do Windows, da Microsoft; e utilizado para a criação de desenhos simples e também para a edição de imagens, é possível encontrar em “editar cores” o sistema RGB, veja Figura 4.1.

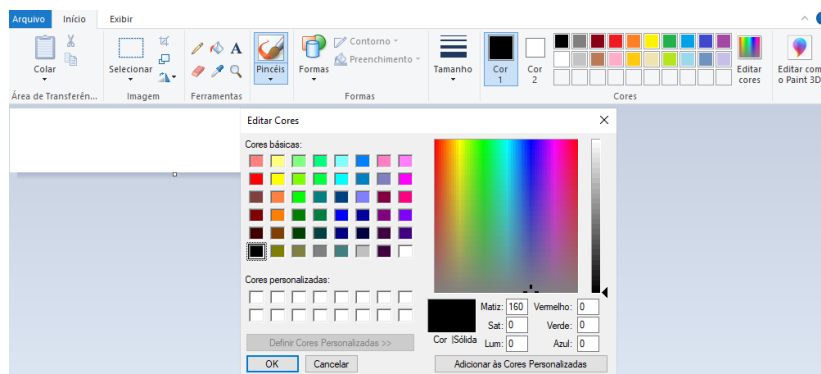


Figura 4.1: Imagem no software Paint

Na Figura 4.1 vemos a cor preta como cor sólida e os códigos do vermelho, verde e azul estão com o valor 0 (zero). Ao colocar outros valores para os códigos vamos formar novas cores sólidas. É importante observarmos que os valores numéricos para os códigos podem variar de 0 a 255.

Para esta atividade usaremos como imagem a bandeira do Brasil, ou seja, construiremos no *software Paint* a imagem desta bandeira considerando, para cada cor, os valores numéricos de acordo com a Tabela 4.1.

Tabela 4.1: Cores no Paint

Cor sólida	código do vermelho	código do verde	código do azul
Verde	30	189	78
Amarelo	240	230	13
Azul	50	47	203

Como resultado, obtemos a Figura 4.2.



Figura 4.2: Bandeira do Brasil, Colorida

Num segundo momento, os alunos irão explorar o *software Paint*, e responder as seguintes perguntas:

1. Qual a cor sólida se todos os valores numéricos dos códigos forem 0?
2. Qual a cor sólida se todos os valores numéricos dos códigos forem o máximo, ou seja, 255?
3. Quais são as cores sólidas se um dos valores numéricos dos códigos for o máximo e os outros dois o mínimo?
4. Qual a cor sólida se todos os valores numéricos dos códigos forem iguais, porém, distintos do valor máximo e do valor mínimo?

As respostas para as perguntas acima serão: Na pergunta 1, a cor sólida obtida será preto. Na pergunta 2, a cor sólida obtida será branco. Na pergunta 3, se o valor máximo (255) for colocado para o código vermelho, a cor sólida será um tom de vermelho; se o valor máximo for colocado para o código verde, a cor sólida será um tom de verde e finalmente, se o valor máximo for colocado no código azul, a cor sólida será um tom de azul. Na pergunta 4, a cor sólida obtida será um tom de cinza.

Na sequência o professor convidará os alunos para transformar a bandeira colorida em uma com cores em escala de cinza.

Para descobrir os novos valores que deverão ser colocados em cada código os alunos deverão utilizar a seguinte função $f : [0, 255] \times [0, 255] \times [0, 255] \rightarrow \mathbb{Z}$ definida por:

$$f(x, y, z) = \left\lfloor \frac{0,30 \cdot x + 0,59 \cdot y + 0,11 \cdot z}{3} \right\rfloor$$

onde x denota o valor do código a ser inserido no vermelho; y denota o valor do código a ser inserido no verde e z denota o valor do código a ser inserido no azul.

O símbolo $\lfloor \cdot \rfloor$ é utilizado na definição da “função piso” e representa que devemos tomar o maior inteiro menor que ou igual a

$$\frac{0,30 \cdot x + 0,59 \cdot y + 0,11 \cdot z}{3} \quad (4.2)$$

Os valores que aparecem na média ponderada na equação (4.2), 30% para vermelho, 59% para o verde e 11% para o azul, foram sugeridos pelo *Wikipedia*.¹

A resposta da aplicação desta função para obtermos os vários tons de cinza é apresentado na Tabela 4.2.

Tabela 4.2: Cores no Paint

cor sólida	código do vermelho	código do verde	código do azul	$f(x, y, z)$
verde	30	189	78	$\lfloor 43,03 \rfloor = 43$
amarelo	240	230	13	$\lfloor 69,71 \rfloor = 69$
azul	50	47	203	$\lfloor 21,68 \rfloor = 21$

Assim, para substituir a cor sólida verde por um tom de cinza devemos colocar o código 129 tanto para o vermelho, quanto para o verde e azul. Para substituir a cor sólida amarela devemos colocar o código 209 tanto para o vermelho, quanto para o verde e azul. Finalmente, Para substituir a cor sólida azul devemos colocar o código 65 tanto para o vermelho, quanto para o verde e azul. O resultado é apresentado na Figura 4.3.

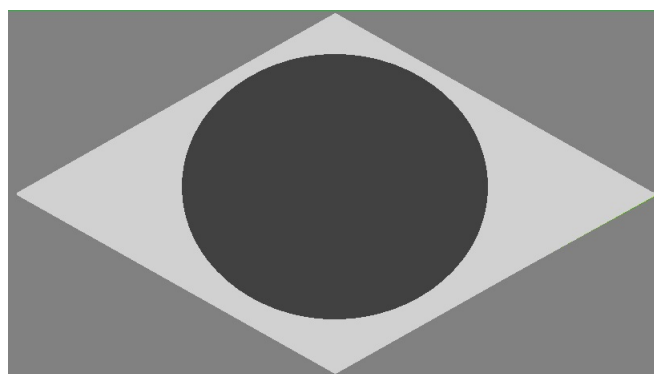


Figura 4.3: Bandeira do Brasil, escala cinza pela média ponderada

¹https://pt.wikipedia.org/wiki/N%C3%ADvel_de_cinza

4.3 Atividade 3: Maximização da área de um quadrilátero

O objetivo desta atividade é trabalhar o conceito de maximização de uma função de várias variáveis.

Problema: A quantidade de aves que podem ser criadas depende da área construída da granja. Se um avicultor dispõe de 800 metros de muro para a construção da granja, a qual deve ser retangular, qual a medida dos lados desta granja de forma a maximizar a sua área?

Vamos denotar por x e y o comprimento dos lados desta granja e por $A(x, y)$ a área desta granja, conforme a Figura 4.4.

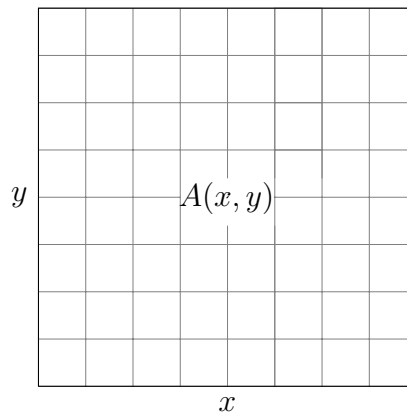


Figura 4.4: Retângulo com área máxima

Sabemos que o perímetro da granja retangular é

$$2x + 2y = 800$$

que equivale a

$$x + y = 400$$

$$y = 400 - x \tag{4.3}$$

A área $A(x, y)$ da granja retangular é dada por:

$$A(x, y) = xy \tag{4.4}$$

Substituindo a equação (4.3) na equação (4.4), temos:

$$A(x, 400 - x) = x(400 - x)$$

$$A(x, 400 - x) = 400x - x^2$$

É possível observar que a função que representa a área da granja tornou-se uma função de uma variável cujo gráfico é uma parábola côncava; assim, o vértice desta parábola representa o ponto máximo desta curva.

Dada a simetria da parábola, a abscissa do vértice da parábola pode ser obtida pela média aritmética das raízes x_1 e x_2 da equação do segundo grau: $400x - x^2 = 0$. Como

as raízes desta equação são $x_1 = 0$ e $x_2 = 400$ a abscissa do vértice da parábola será $x = 200$. Logo, da equação $y = 400 - 200$ temos $y = 200$.

Portanto, esta granja, para ter área máxima, deverá ser construída com lados de comprimento 200 metros, ou seja, esta granja será construída no formato de um quadrado. Além disso, podemos concluir que a área desta granja será $A(200, 200) = 200 \cdot 200 = 40.000$ metros quadrados.

4.4 Atividade 4: Maximização da área de um quadrilátero justaposto

Problema: Com a demanda crescente, o avicultor pretende construir uma nova granja para a criação de outras aves. Para isto ele dispõe, novamente, de 800 metros de muro para construção do segundo barracão, o qual deve ser construído de forma retangular e justaposto no primeiro, afim de aproveitar uma parte do muro. Qual deverá ser a medida da largura e comprimento desta nova granja?

Vamos denotar por x e y o comprimento dos lados desta granja e por $A(x, y)$ a área desta granja, conforme a Figura 4.5.

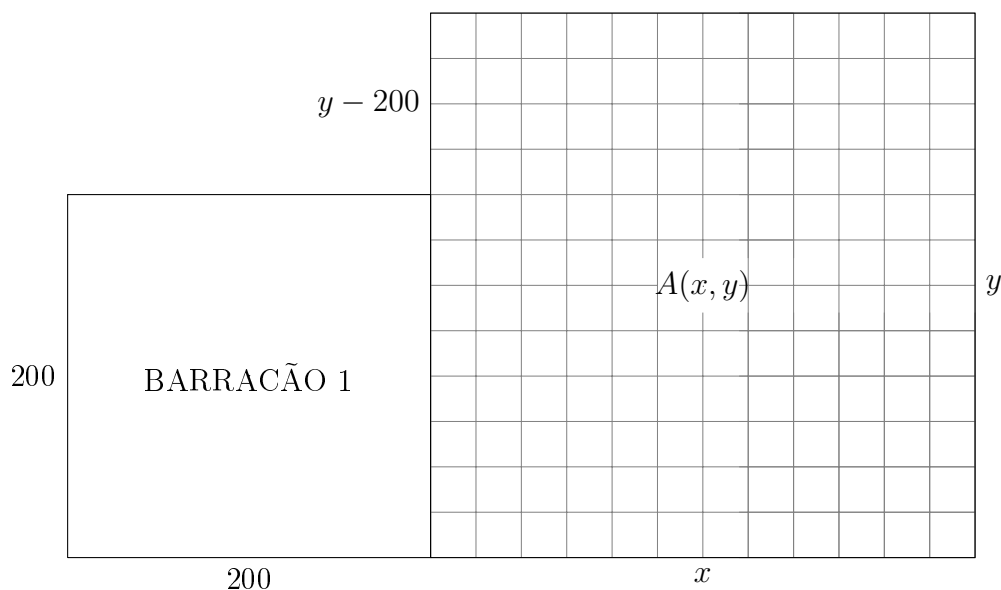


Figura 4.5: Retângulo justaposto com área máxima

Agora o perímetro do barracão retangular será:

$$2x - 200 + 2y = 800$$

que equivale a

$$\begin{aligned} x + y &= 500 \\ y &= 500 - x \end{aligned} \tag{4.5}$$

A área $A(x, y)$ da granja retangular é dada por:

$$A(x, y) = xy \tag{4.6}$$

Substituindo a equação (4.5) na equação (4.6) temos:

$$A(x, 500 - x) = x(500 - x) \quad (4.7)$$

É possível observar que a função que representa a área da granja tornou-se uma função de uma variável cujo gráfico é uma parábola côncava; assim, o vértice desta parábola representa o ponto máximo desta curva.

Dada a simetria da parábola, a abscissa do vértice da parábola pode ser obtida pela média aritmética das raízes x_1 e x_2 da equação do segundo grau: $500x - x^2 = 0$. Como as raízes desta equação são $x_1 = 0$ e $x_2 = 500$ a abscissa do vértice da parábola será $x = 250$. Logo, da equação $y = 500 - 250$ temos $y = 250$.

Portanto, esta granja, para ter área máxima, deverá ser construída com lados de comprimento 250 metros, ou seja, esta granja será construída no formato de um quadrado. Além disso, podemos concluir que a área desta granja será $A(250, 250) = 250 \cdot 250 = 62.500$ metros quadrados.

5 Considerações Finais

No desenvolver deste trabalho foi possível fazer uma breve introdução histórica sobre o desenvolvimento do Teorema da Função Implícita, apresentar como diferentes personagens contribuíram para o desenvolvimento da matemática, em especial dos teoremas centrais deste trabalho.

Definimos conceitos importantes do espaço euclidiano, \mathbb{R}^n , aplicações contínuas, conjuntos abertos, conjuntos fechados, conexidade, derivadas e diferenciabilidades no espaço n -dimensional que podem ser aplicados a um modelo matemático. De posse destes resultados foi possível enunciar, demonstrar e apresentar aplicações do Teorema da Função Inversa e do Teorema da Função implícita.

Posto isso, percebemos que é importante o professor dominar as ferramentas de trabalho e a construção adequada dos conceitos de função, pois as habilidades deste conteúdo são firmadas na Base Nacional Comum Curricular (BNCC) que preconiza competências, as quais os alunos devem dominar ao longo de sua jornada no ensino básico. Assim, propomos atividades que estimulem o estudo de funções de várias variáveis para o ensino médio.

As propostas tentaram introduzir reflexões a respeito de temas relevantes, a saber: modelagem da vida financeira, introdução à computação gráfica e otimização de função.

Devido aos poucos trabalhos de pesquisa neste tema no repositório de dissertações do PROFMAT, podemos indagar sobre o interesse dos discentes, uma vez que foi encontrado apenas um trabalho para a pesquisa “função implícita” de Husein (2014) e outro para “função inversa” de Tezoto (2014).

Conseqüente, com a apresentação deste trabalho, esperamos colaborar com o desenvolvimento de pesquisas no programa nesta temática.

Referências

- [1] APOSTOL, Tom M., Calculus, vol.II. 2 edição. Indiana: John Willey & Sons, 1969.
- [2] BALESTRI, Rodrigo. Matemática: Interação e Tecnologia. 2. ed. São Paulo: Leya, 2016.
- [3] BOYCE, W. E.; DIPRIMA, R. C. 8. ed. Rio de Janeiro: LTC, 2006.
- [4] BRASIL. Ministério da Educação. Base Nacional Comum Curricular. Brasília, 2018.
- [5] BRASIL. Ministério da saúde. Coronavírus. Disponível em: < covid.saude.gov.br >. Acesso em: 10 de julho de 2020.
- [6] DELGADO, Jorge; FRENSEL. Katia; CRISSAFF. Layla. Geometria analítica. 2 ed. Rio de Janeiro: SBM, 2017.
- [7] DE LIMA, Ronaldo Freire. Topologia e Análise no Espaço \mathbb{R}^n . 1 ed. Rio de Janeiro: SBM, 2015.
- [8] DE OLIVEIRA, Oswaldo. The implicit and the inverse function theorems: easy proofs. Real Analysis Exchange, v. 39, n. 1, p. 207-218, 2014.
- [9] FERREIRA, Mateus et al. Aspectos comportamentais no hábito de poupar em adultos jovens: uma comparação entre alta e baixa renda. XVI Seminários de Administração, São Paulo, Universidade de São Paulo. Anais do XVI Seminários de Administração, USP, São Paulo, 2013.
- [10] GUIDORIZZI. Hamilton Luiz. Um curso de Cálculo - Vol. I. 5. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 2001.
- [11] HUSEIN. M. S. M. H. Diferenciação em espaço de ordem n e o Teorema da Função Implícita com aplicações. Dissertação (mestrado profissional) - Universidade Federal de Santa Catarina. Florianópolis-SC, p.90, 2014.
- [12] IEZZI. et al. Matemática: Ciência e aplicações. 7. ed. São Paulo: Saraiva, 2016.
- [13] KRANTZ, S. G.; PARKS. H. R., The implicit Function Theorem: History, Theory, and Applications. 2 edição. Boston: Birkhäuser, 1951.
- [14] LEONARDO. Fábio Martins de (ed.), Conexões com a Matemática. 3. ed. São Paulo: Moderna, 2016.

- [15] LIMA, Elon Lages. *Análise Real: Funções de uma variável*. vol.I. 11 ed. Rio de Janeiro: IMPA, 2018.
- [16] LIMA, Elon Lages. *Análise Real*, vol.II. 2 ed. Rio de Janeiro: IMPA, 2004.
- [17] LIMA, Elon Lages. *Curso de Análise*, vol.II. 11 ed. Rio de Janeiro: IMPA, 2018.
- [18] LIMA, Elon Lages. *Números e funções Reais*. 1 ed. Rio de Janeiro: SBM, 2017.
- [19] MURAKAMI, Carlos et al. *Fundamentos de Matemática Elementar: Conjuntos e funções*. 7. ed. So Paulo: Atual.
- [20] ROYDEN, Halsey Lawrence; FITZPATRICK, Patrick. *Real analysis*. 4 ed. New York: Macmillan, 1988.
- [21] SCARPELLO, Giovanni Mingari; RITELLI, Daniele. A historical outline of the theorem of implicit functions. *Divulgaciones Matemáticas*, v. 10, n. 2, p. 171-180, 2002.
- [22] STEWART, James. *Cálculo*. 5. ed. São Paulo: Thomson, 2006.
- [23] TEZOTO, Leandro. *Sobre o teorema da função Inversa*. Dissertação (mestrado profissional) - Universidade Estadual Paulista. Rio Claro-SP, p.35, 2014.
- [24] WIKIPÉDIA. *Wikipédia: a enciclopédia livre*. Disponível em <<https://pt.wikipedia.org>>. Acesso em 03 mar. 2021.
- [25] ZUFFI, Edna Maura; PACCA, Jesuína LA. Sobre funções e a linguagem matemática de professores do Ensino Médio, p. 7-28. *Zetetike*, v. 8, n. 1-2, 2000.