

Universidade Estadual do Sudoeste da Bahia
Departamento de Ciências Exatas e Tecnológicas

Licenciatura em Matemática

Allan Amorim Oliveira

O TEOREMA DO MULTIPLICADOR DE
LAGRANGE NO ESPAÇO \mathbb{R}^n

AD PLENAM VITAM

Vitória da Conquista - BA

2024

Allan Amorim Oliveira

O TEOREMA DO MULTIPLICADOR DE LAGRANGE NO ESPAÇO \mathbb{R}^n

Monografia apresentada ao Departamento de Ciências Exatas e Tecnológicas da Universidade Estadual do Sudoeste da Bahia - Campus Vitória da Conquista-BA, para obtenção do título de Licenciado em Matemática, sob orientação do Prof. Dr. Júlio César dos Reis.

Vitória da Conquista - BA
2024

Allan Amorim Oliveira

O TEOREMA DO MULTIPLICADOR DE LAGRANGE NO ESPAÇO \mathbb{R}^n

Monografia apresentada ao Colegiado do Curso de Matemática como requisito parcial para aprovação na disciplina Seminário de Pesquisa II do Curso de Licenciatura em Matemática. Trabalho aprovado em 12 de dezembro de 2024.

BANCA EXAMINADORA

Prof. Dr. Júlio César dos Reis - UESB

Orientador

Prof^a. Dra. Clênia Andrade Oliveira de Melo - UESB

Convidado

Prof. Dr. Fernando dos Santos Silva - UESB

Convidado

Vitória da Conquista - BA

2024

“Dream on! Dream until your dreams come true.”
Aerosmith, *“Dream On”*.

AGRADECIMENTOS

À priori, agradeço a Deus por todas as bênçãos ao longo da minha trajetória acadêmica, por me dar força, saúde e sabedoria para enfrentar os desafios. Senhor, só tenho a agradecer por tudo que me proporcionou nestes anos.

À minha mãe, Valdirene, por todo suporte, confiança e apoio que me permitiram seguir em busca dos meus sonhos. Mãe, sem você, nada disso seria possível. Tu és a pessoa mais importante da minha vida, e não tenho palavras para descrever o quanto te amo! Muito obrigado por tudo, mãe.

Aos meus irmãos Welton, Thiago e Rodrigo, pelo incentivo e conselhos.

À UESB e a todos os professores do Curso de Licenciatura em Matemática, pelo ambiente acadêmico propício ao desenvolvimento deste trabalho e pela qualidade do ensino oferecido. Menciono, em especial, os professores: Fernando e Clênia (por aceitarem o convite para a banca e contribuírem com este trabalho), Altemar (*in memoriam*), Claudinei e Irani (que me apresentaram à pesquisa científica), Bárbara, Galvina, Gerson (meu chefe), Jonson (meu orientador no PIBID), Lucas e Ricardo (pela parceria no GEMPA) e Marcio (meu tutor no PETIMAT).

Ao meu orientador, Professor Júlio César dos Reis, pelo suporte, sabedoria e orientação de extrema qualidade que foram fundamentais para a realização deste trabalho. Júlio, o senhor é um professor e um ser humano fantástico. Eu o admiro bastante e me inspiro muito no senhor como professor e educador. Muito obrigado por tudo! Foi um prazer ser seu orientando.

À Poliane, por seu amor, cuidado e apoio incondicional ao longo desse período, que foram fundamentais para mim. Agradeço ao incentivo e contribuição para que eu me tornasse um estudante e, principalmente, um ser humano melhor. Você foi e é extremamente importante para mim. Muito obrigado, Polly!

Aos meus colegas de curso, especialmente da turma 2020.1, a minha turma, pelo apoio, parceria e momentos compartilhados durante esses 4 anos. Obrigado, turma 2020.1!

Aos meus amigos e familiares, em especial aos meus tios Vanildo, Nóbia, Ni, Vassonio e Nena, por todo o apoio e, à minha vó, Estelina, por tudo que fez por mim e minha mãe.

À AEESP e seus funcionários, por proporcionarem o transporte que permitiu a realização dos meus estudos em Vitória da Conquista, viajando dia após dia nessa luta.

Ao Sport Club Corinthians Paulista, o maior do mundo, que desde a sua fundação em 1910, com origem humilde entre operários, se transformou em um verdadeiro símbolo de superação e luta. Ser corinthiano é sinônimo de viver com paixão, loucura e alegria, e eu não poderia concluir este trabalho sem dizer isso: VAI, CORINTHIANS!

Por fim, a todos que, de alguma forma, contribuíram para a conclusão deste trabalho e dessa graduação, meu muito obrigado.

RESUMO

Este trabalho tem como objetivo principal apresentar e demonstrar o Teorema do Multiplicador de Lagrange, no contexto do espaço euclidiano n -dimensional, o \mathbb{R}^n . A metodologia utilizada nessa pesquisa envolveu a revisão e estudos de fontes bibliográficas na área do Cálculo Diferencial e da Análise no \mathbb{R}^n , sendo estas, livros destas áreas. A partir dessas referências, implementou-se uma abordagem analítica e exploratória, na qual foram apresentados os conceitos e definições fundamentais para a formulação e a demonstração do teorema.

Palavras-chave: Lagrange; Máximo condicionado; Análise no \mathbb{R}^n ; Topologia.

ABSTRACT

This work aims to present and demonstrate the Lagrange Multiplier Theorem in the context of the n -dimensional Euclidean space, \mathbb{R}^n . The methodology employed in this research involved a review and study of bibliographic sources in the fields of Differential Calculus and Analysis in \mathbb{R}^n , primarily drawing from textbooks in these areas. Based on these references, an analytical and exploratory approach was implemented, presenting the fundamental concepts and definitions necessary for the formulation and demonstration of the theorem.

Keywords: Lagrange; Constrained Maximum; Analysis in \mathbb{R}^n ; Topology.

Lista de Figuras

1.1	Gráfico de uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida em um intervalo fechado $[a, b]$	9
1.2	Gráfico das curvas de nível de $f(x, y)$ com restrição $\varphi(x, y)$	11
1.3	Maximização da função $f(x, y)$ sujeita à restrição $\varphi(x, y) = 0$	12
1.4	Caixa retangular de dimensões x, y e z	13
1.5	Gráfico de $f(x, y) = x^2 + 2y^2$ com a restrição $x^2 + y^2 = 1$	14
1.6	Gráfico da esfera dada por $x^2 + y^2 + z^2 = 4$	15
2.1	Representação de uma bola aberta $B(a; r)$ e de uma bola menor $B(x; \delta) \subset B(a; r)$	19
3.1	Vetor velocidade, reta tangente e vetor secante no caminho f	38
3.2	Círculo unitário S^1 com $f(t)$ e o vetor velocidade $f'(t)$ do caminho f	39
3.3	Trajectoria da curva $h(\mathbb{R}) = g(\mathbb{R})$ com ponto angular em $(0, 0)$	40
3.4	Representação geométrica da derivada parcial $\partial f / \partial x_i$ no plano $y = b$	44
3.5	Derivada direcional de f no ponto a ao longo de v	47
3.6	Gráfico de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = xy / (x^2 + y^2)$	56
3.7	Gráfico de $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g(x, y) = x^2 y / (x^2 + y^2)$	56
3.8	Gráfico de $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(x, y) = x^3 y / (x^6 + y^2)$	57
3.9	Gráfico de $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\varphi(x, y) = x^3 y / (x^4 + y^2)$	57
3.10	Gradiente $\nabla f(a)$ indicando a direção de maior crescimento de f	66
3.11	Curvas de nível de $f(x, y) = ax + by$ com retas paralelas perpendiculares a (a, b)	68
3.12	Curvas de nível de $g(x, y) = x^2 + y^2$, círculos com centro na origem e raio \sqrt{c}	68
3.13	Curvas de nível de $x^2 - y^2 = c$ e vetor gradiente $(2x, -2y)$	69
3.14	Superfícies de nível de $x^2 + y^2 - z^2 = c$: cone ($c = 0$), hiperbolóides ($c \neq 0$).	70
3.15	Gráfico de $f^{-1}(c)$, ilustrando a função implícita $y = \xi(x)$ definida por $f(x, y) = c$	71
3.16	Curva $f^{-1}(c)$ em $I \times J$, com $y = \xi(x)$ definida implicitamente por $f(x, y) = c$	72
3.17	Curva C de classe \mathcal{C}^k com representação local como gráfico de funções em V e W	74
4.1	Retas $ax + by = c$ e pontos críticos no círculo unitário S^1	84

Sumário

Introdução	7
1 Máximos e Mínimos Condicionados	9
1.1 Definições	10
1.2 Exemplos de máximos condicionados	13
2 Topologia do Espaço \mathbb{R}^n	18
2.1 Conjuntos abertos	18
2.2 Conjuntos fechados	22
2.3 Sequências no espaço euclidiano	27
2.4 Conjuntos compactos	31
3 Caminhos Diferenciáveis e Funções Reais de n variáveis	37
3.1 Caminhos diferenciáveis	37
3.2 Derivadas parciais	42
3.3 Derivadas direcionais	46
3.4 Funções diferenciáveis	50
3.5 A diferencial de uma função	61
3.6 O gradiente de uma função diferenciável	64
3.7 O Teorema da Função Implícita	70
4 Multiplicador de Lagrange	81
4.1 Teorema do Multiplicador de Lagrange	82
4.2 Demonstração do Teorema do Multiplicador de Lagrange.	82
4.3 Observações e Exemplos	83
4.4 Exemplo de uma aplicação: Ensemble Microcanônico	88
5 Considerações Finais	89
Referências bibliográficas	90

Introdução

Introdução

O Teorema do Multiplicador de Lagrange leva o nome de Joseph-Louis Lagrange (1736–1813), um destacado matemático do século XVIII. Este teorema trata de relações entre derivadas em problemas de análise que envolvem restrições, estabelecendo condições necessárias para determinar pontos críticos de uma função $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ quando esta está restrita a uma hipersuperfície $M \subset \mathbb{R}^{n+1}$. Baseando-se no conceito de gradiente, o teorema identifica situações em que a variação de f é compatível com a restrição imposta por M , introduzindo o parâmetro λ , conhecido como multiplicador de Lagrange, para expressar essa relação de maneira precisa.

O objetivo principal desta monografia é apresentar o Teorema do Multiplicador de Lagrange no contexto do espaço \mathbb{R}^n , fornecendo uma formulação rigorosa e uma demonstração detalhada do teorema. Para alcançar esse objetivo, é necessário construir uma base teórica que abrange conceitos de Topologia e Análise de funções reais em várias variáveis. Essa construção progressiva nos permite contextualizar e compreender os elementos necessários para que o teorema seja corretamente enunciado e provado.

Este trabalho tem como referência principal o livro [4], e está organizado em cinco capítulos.

No Capítulo 1, inicia-se a discussão com definições e exemplos de máximos e mínimos condicionados, no contexto de funções em dimensões menores, como vemos normalmente no estudo do Cálculo Diferencial em várias variáveis. Para este capítulo, utilizamos os livros [5] e [1] como referências. Esses conceitos iniciais são apresentados de forma a garantir uma análise teórica mais aprofundada que será desenvolvida nos capítulos seguintes, onde iremos abordar conceitos de Análise no \mathbb{R}^n .

O Capítulo 2 aborda noções de topologia no espaço euclidiano \mathbb{R}^n , discutindo conjuntos abertos, conjuntos fechados, sequências e conjuntos compactos. Entender esses conceitos é de extrema importância para compreendermos o enunciado e ambientação do teorema.

No Capítulo 3, o foco está em funções reais de várias (n) variáveis e caminhos diferenciáveis. Esse capítulo apresenta as noções de derivadas parciais e direcionais, além do conceito de funções diferenciáveis e suas propriedades. Também é introduzido o gradiente de uma função,

que desempenha um papel central no entendimento do Teorema do Multiplicador de Lagrange. Por fim, esse capítulo apresenta o Teorema da Função Implícita.

O Capítulo 4, este o ponto principal desta monografia, é inteiramente dedicado ao Teorema do Multiplicador de Lagrange. Nele, o teorema é enunciado demonstrado detalhadamente, seguindo uma abordagem estruturada e lógica, utilizando na íntegra os conceitos abordados na . Este capítulo constitui o núcleo da monografia, consolidando os conceitos desenvolvidos nos capítulos anteriores e conectando-os diretamente ao resultado principal. Além disso, são incluídas observações importantes e exemplos que ilustram as ideias centrais do teorema e reforçam a compreensão de sua aplicação teórica no contexto abordado.

Os capítulos 2, 3 e 4 tiveram o livro [4] como referência.

Por fim, no Capítulo 5, são apresentadas as considerações finais. Esse capítulo oferece uma reflexão sobre os resultados alcançados e também aponta possíveis direções para estudos futuros, que podem incluir o aprofundamento em generalizações do teorema e/ou em aplicações do mesmo.

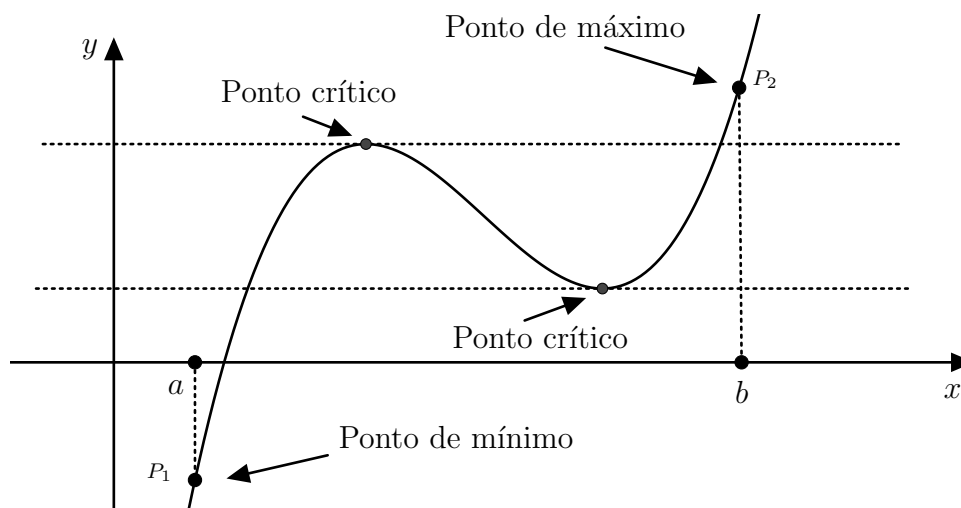
Capítulo 1

Máximos e Mínimos Condicionados

Quando se busca resolver problemas práticos que envolvem a maximização ou minimização de funções, torna-se essencial compreender os conceitos de máximos e mínimos condicionados.

No caso de funções de uma variável, ao analisar máximos e mínimos, é possível que surjam situações em que a variável em questão assume valores em um intervalo fechado. Nesses casos, os pontos de máximo ou mínimo podem ocorrer não apenas em pontos críticos da função, mas também nas extremidades desse intervalo. Observe o gráfico abaixo, que ilustra uma função definida em um intervalo fechado $[a, b]$:

Figura 1.1: Gráfico de uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida em um intervalo fechado $[a, b]$.



Fonte: Criado pelos autores.

Note que os valores de máximo e mínimo absolutos estão nas extremidades do intervalo $[a, b]$.

Além disso, observe que a função é contínua. Uma função contínua definida em um intervalo fechado sempre admite máximos e mínimos. Se a função tiver apenas um ponto crítico, os valores extremos podem estar nas extremidades. Da mesma forma, caso a função não possua pontos críticos, os máximos e mínimos também podem ocorrer nas extremidades.

Agora, consideremos como esse comportamento se estende para funções de várias variáveis.

1.1 Definições

Definição 1. *Uma função $f : Dom(f) \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ (respectivamente $f : Dom(f) \subset \mathbb{R}^3 \rightarrow \mathbb{R}$) é uma regra que atribui a cada par ordenado $(x, y) \in Dom(f) \subset \mathbb{R}^2$ (respectivamente tripla ordenada $(x, y, z) \in Dom(f) \subset \mathbb{R}^3$) um único número real, denotado por $f(x, y)$ (respectivamente $f(x, y, z)$). O conjunto $Dom(f)$ é chamado de domínio da função f .*

O aumento no grau de liberdade trazido pelas dimensões adicionais enriquece as possibilidades, tornando as situações mais complexas e, por vezes, mais sutis em termos de Análise. Esse desafio, porém, é compensado pela riqueza das novas possibilidades que surgem.

Assim, é importante analisar a questão dos máximos e mínimos condicionados. Esse conceito envolve encontrar os valores máximos e mínimos de uma função de várias variáveis $f(x, y, z, \dots)$ sob a condição $\varphi(x, y, z, \dots) = c$. Essa é a essência da ideia de “condicionado”. Nosso objetivo é determinar os valores extremos da função f , mas não em todo o seu domínio, e sim apenas nos pontos que respeitam uma condição imposta. Portanto, os pontos analisados são condicionados e não variam livremente por todo o domínio.

Antes de explorarmos um exemplo prático, vamos considerar o seguinte conceito: suponha uma função cujo domínio está em \mathbb{R}^2 , ou seja, uma função f de duas variáveis, $f(x, y)$, sujeita à restrição $\varphi(x, y) = c$. Sabemos que o gráfico de uma função de duas variáveis é uma superfície em \mathbb{R}^3 . No entanto, vamos focar apenas em seu domínio, que neste caso é no plano \mathbb{R}^2 .

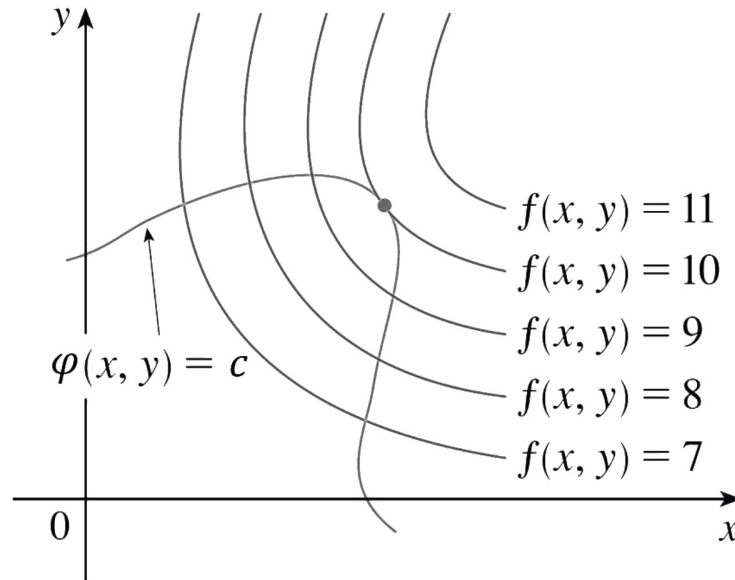
É importante destacar que nosso interesse não é encontrar os extremos em todo o domínio de f . Se estivéssemos interessados nisso, buscaríamos os pontos críticos, analisando as derivadas parciais, igualando-as a zero e prosseguindo com a análise a partir dos resultados. No entanto, aqui estamos interessados em encontrar máximos e mínimos de $f(x, y)$ apenas para os pontos que satisfazem $\varphi(x, y) = c$. neste caso, significa trabalhar com curvas de nível.

Definição 2. *As curvas de nível de uma função f de duas variáveis são conjuntos de pontos $(x, y) \in Dom(f)$ tal que $f(x, y) = c$, onde c é uma constante.*

Ou seja, a condicionante constante c é uma curva de nível da função φ . Se estivéssemos falando de uma função com 3 variáveis, a condicionante seria chamada de superfície de nível de φ .

Observemos o gráfico abaixo que representa as curvas de nível de uma função genérica $f(x, y)$, com a função de restrição $\varphi(x, y)$.

Figura 1.2: Gráfico das curvas de nível de $f(x, y)$ com restrição $\varphi(x, y)$.



Fonte: STEWART, James. *Cálculo, Volume 2*. 6. ed. São Paulo: Cengage Learning, 2013.

A questão que devemos levantar ao observar o gráfico é a seguinte: se estivermos “caminhando” ao longo da curva de nível da função φ e a função f está aumentando ou diminuindo, onde ocorre o ponto de máximo?

Note que, se estivermos andando sobre a curva de nível de φ e a função f estiver aumentando, isso implica que o valor de f está subindo, ou seja, ainda não alcançamos o ponto de máximo. O raciocínio é semelhante no caso de um ponto de mínimo. Mas, afinal, onde encontramos o ponto de máximo? A resposta é: quando a curva de nível de φ for tangente à curva de nível de f . No gráfico acima, isso ocorre na curva de nível dada por $f(x, y) = 10$. Como se pode observar, a curva de nível de φ intersecta as curvas de nível de f até atingir o ponto marcado acima, onde a tangência acontece.

Assim, os pontos de máximo ou mínimo ocorrem nos locais onde as curvas de nível de f e φ (que já é uma curva de nível) são tangentes.

Agora, lembremos que os vetores gradientes são sempre perpendiculares às curvas de nível. Quando duas curvas de nível são tangentes, as direções perpendiculares a elas coincidem. Portanto, os vetores gradientes dessas funções devem estar na mesma direção nesses pontos. Em termos de Álgebra Linear, podemos afirmar que esses vetores são múltiplos escalares entre si.

Logo, quando a função $f(x, y)$ atinge um ponto de máximo ou mínimo sujeito à restrição $\varphi(x, y)$, as condições que devem ser satisfeitas são:

- 1) o vetor gradiente da função f deve ser um múltiplo escalar do vetor gradiente da função φ , ou seja, satisfazer a seguinte igualdade:

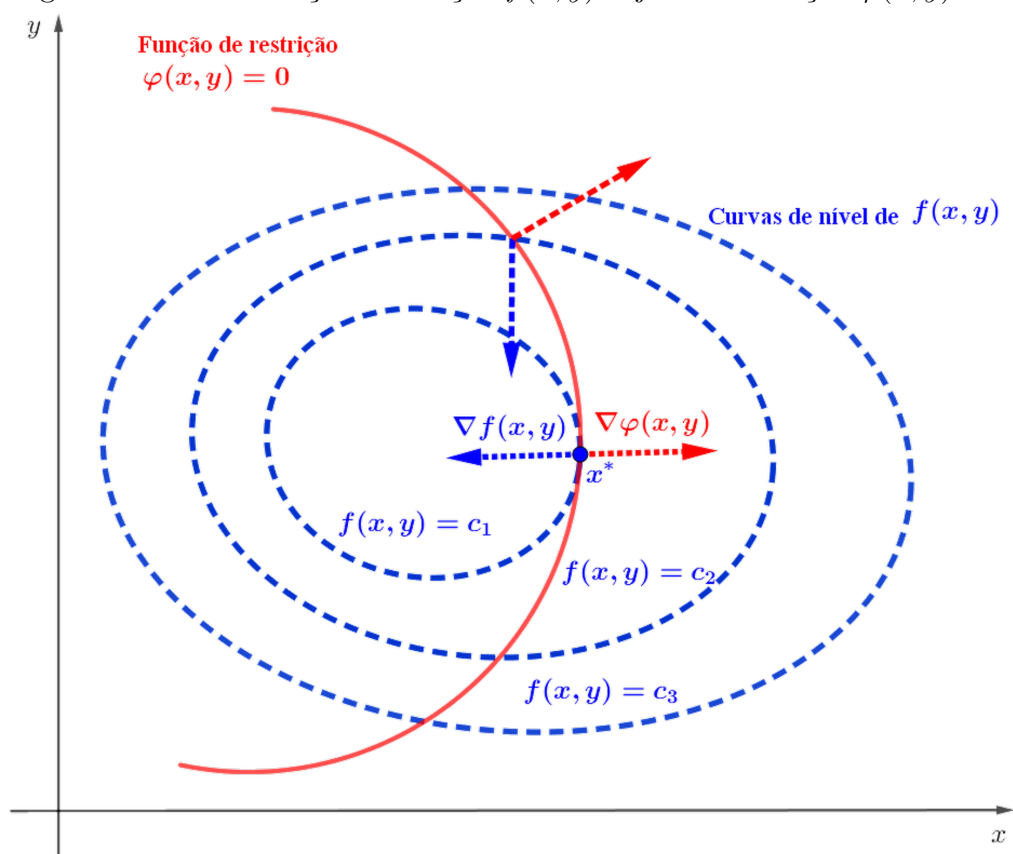
$$\nabla f(x, y) = \lambda \cdot \nabla \varphi(x, y), \quad \lambda \in \mathbb{R},$$

onde λ é chamado de multiplicador de Lagrange;

- 2) $\varphi(x, y) = c$.

Na figura abaixo, no ponto de máximo local com restrição, os gradientes de f e φ , ou seja, $\nabla f(x, y)$ e $\nabla \varphi(x, y)$, são paralelos.

Figura 1.3: Maximização da função $f(x, y)$ sujeita à restrição $\varphi(x, y) = 0$.



Fonte: Baseado em Vadlamani, S. K., Xiao, T. P., & Yablonoitch, E. (2020). *Physics successfully implements Lagrange multiplier optimization.*

Dessa forma, chegamos às famosas condições dos multiplicadores de Lagrange. Dessas condições, resultará um sistema que nos indicará alguns candidatos a serem máximos e mínimos.

É válido ressaltar, novamente, que o raciocínio é análogo para 3 ou mais variáveis, porém, as curvas de nível passariam a ser superfícies de nível.

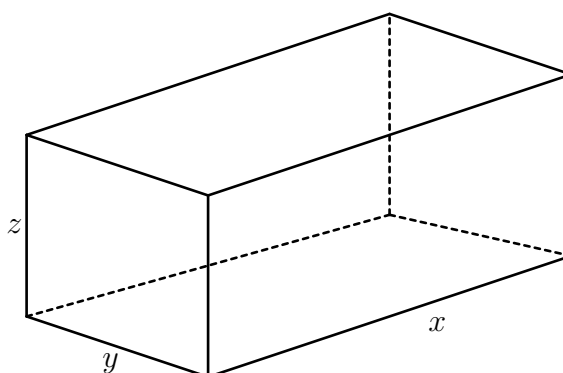
Na próxima seção, esse raciocínio será aplicado através de alguns exemplos.

1.2 Exemplos de máximos condicionados

Nesta seção, serão vistos 3 exemplos para aplicar esse raciocínio. Tais exemplos foram retirados de [5].

Exemplo 3. *Uma caixa retangular sem tampa deve ser feita com 12 m^2 de papelão. Vamos determinar o volume máximo dessa caixa.*

Figura 1.4: Caixa retangular de dimensões x , y e z .



Fonte: Criado pelos autores.

Acima, observemos a representação da figura da caixa retangular que queremos maximizar. Neste caso, sejam x , y e z o comprimento, a largura e a altura, respectivamente, da caixa em metros. Queremos maximizar $V = xyz$ sujeita à restrição $\varphi(x, y, z) = 2xz + 2yz + xy = 12$. Note que, pelo fato de querermos maximizar o volume da caixa sem a tampa, a área xy não é dobrada. Por isso, a restrição φ é essa acima. Sabemos que o volume tem a forma $V = x \cdot y \cdot z$. Vamos utilizar o método dos multiplicadores de Lagrange. Aqui, tomaremos $V = f(x, y, z)$. Primeiramente, olhamos para os valores de x , y , z e λ , tais que $\nabla f = \lambda \nabla \varphi$ e $\varphi(x, y, z) = 12$. Isso gera as equações:

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial \varphi}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial \varphi}{\partial y}, \quad \frac{\partial f}{\partial z} = \lambda \frac{\partial \varphi}{\partial z}, \quad 2xz + 2yz + xy = 12,$$

ou seja:

$$yz = \lambda(2z + y); \quad xz = \lambda(2z + x); \quad xy = \lambda(2x + 2y); \quad 2xz + 2yz + xy = 12.$$

Observe que, se multiplicarmos o primeiro resultado por x , o segundo por y , e o terceiro por z , os lados esquerdos dessas equações ficam idênticos. Fazendo isso, temos

$$xyz = \lambda(2xz + xy); \quad xyz = \lambda(2yz + xy); \quad xyz = \lambda(2xz + 2yz).$$

Note que $\lambda \neq 0$, pois nos levaria numa contradição em $\varphi(x, y, z)$.

Logo, temos

$$2xz + xy = 2yz + xy.$$

Note que $z \neq 0$. Com isso, chegamos a $x = y$.

Além disso, temos $2yz + xy = 2xz + 2yz$, o que nos leva a $2xz = xy$. Dessa igualdade, temos

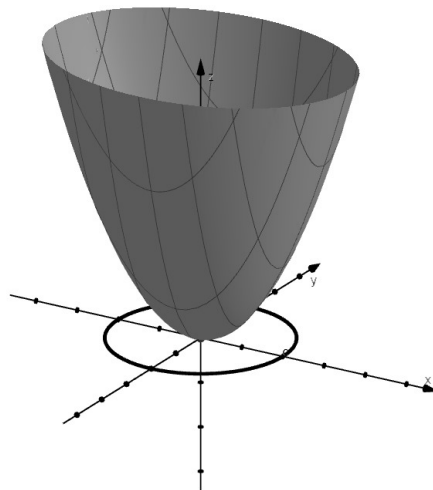
$$4z^2 + 4z^2 + 4z^2 = 12.$$

Como x, y e z todos são positivos, teremos $z = 1$ e, portanto, $x = 2$ e $y = 2$.

Multiplicando os três valores encontrados, chegamos em 4 m^3 de volume máximo dessa caixa.

Exemplo 4. Vamos determinar os valores extremos da função $f(x, y) = x^2 + 2y^2$ no círculo $x^2 + y^2 = 1$.

Figura 1.5: Gráfico de $f(x, y) = x^2 + 2y^2$ com a restrição $x^2 + y^2 = 1$.



Fonte: Criado pelos autores com o auxílio do *software* GeoGebra.

Analogamente ao Exemplo (3), dessa vez para 2 variáveis, calcularemos

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial \varphi}{\partial x}; \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial \varphi}{\partial y}.$$

Temos, portanto:

$$2x = 2x\lambda; \quad 4y = 2y\lambda; \quad x^2 + y^2 = 1.$$

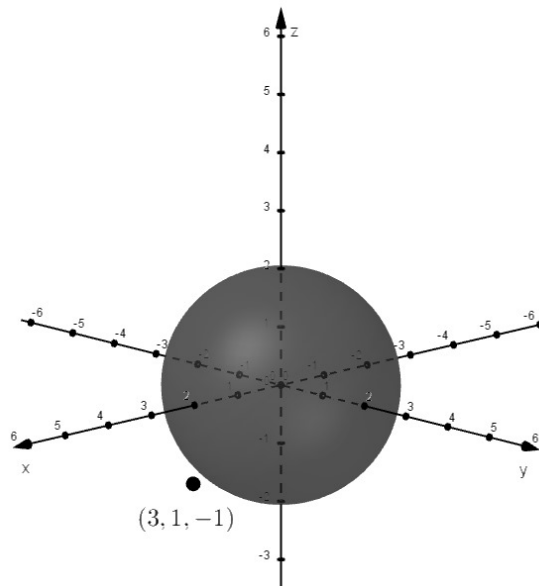
Observe que $x = 0$ ou $\lambda = 1$. Se $x = 0$, então $y = \pm 1$. Se $\lambda = 1$, então $y = 0$, e assim temos $x = \pm 1$. Dessa forma, os valores extremos possíveis de f são os pontos $(0, 1)$, $(0, -1)$, $(1, 0)$ e $(-1, 0)$. Calculando f nesses quatro pontos, achamos

$$f(0, 1) = 2; \quad f(0, -1) = 2; \quad f(1, 0) = 1; \quad f(-1, 0) = 1.$$

Portanto, o valor máximo de f no círculo $x^2 + y^2 = 1$ é $f(0, \pm 1) = 2$, e o valor mínimo é $f(\pm 1, 0) = 1$.

Exemplo 5. Vamos determinar os pontos da esfera $x^2 + y^2 + z^2 = 4$ que estão mais próximos e mais distantes do ponto $(3, 1, -1)$.

Figura 1.6: Gráfico da esfera dada por $x^2 + y^2 + z^2 = 4$.



Fonte: Criado pelos autores com o auxílio do *software* GeoGebra.

A distância de um ponto (x, y, z) ao ponto $(3, 1, -1)$ é

$$d = \sqrt{(x-3)^2 + (y-1)^2 + (z+1)^2}.$$

Elevando os dois lados ao quadrado:

$$d^2 = f(x, y, z) = (x-3)^2 + (y-1)^2 + (z+1)^2.$$

A restrição é que o ponto (x, y, z) pertença à esfera, ou seja, $\varphi(x, y, z) = x^2 + y^2 + z^2 = 4$. Precisamos resolver $\nabla f = \lambda \nabla \varphi$, $\varphi = 4$. Isso dá

$$2(x-3) = 2x\lambda; \quad 2(y-1) = 2y\lambda; \quad 2(z+1) = 2z\lambda; \quad x^2 + y^2 + z^2 = 4$$

O modo mais simples de resolver essas equações é determinar x , y e z em termos de λ e depois substituir os valores.

$$x-3 = x\lambda \quad \text{ou} \quad x(1-\lambda) = 3 \quad \text{ou} \quad x = \frac{3}{1-\lambda}$$

[Observe que $1-\lambda \neq 0$ porque $\lambda = 1$ é impossível.] Da mesma forma, temos

$$y = \frac{1}{1-\lambda} \quad \text{e} \quad z = \frac{-1}{1-\lambda}$$

Portanto, temos

$$\frac{3^2}{(1-\lambda)^2} + \frac{1^2}{(1-\lambda)^2} + \frac{(-1)^2}{(1-\lambda)^2} = 4 \quad \text{que nos dá} \quad (1-\lambda)^2 = \frac{11}{4}, \quad 1-\lambda = \pm \frac{\sqrt{11}}{2}, \quad \text{logo}$$

$$\lambda = 1 \pm \frac{\sqrt{11}}{2}.$$

Esses valores de λ então fornecem os pontos correspondentes (x, y, z) :

$$\left(\frac{6}{\sqrt{11}}, \frac{2}{\sqrt{11}}, \frac{-2}{\sqrt{11}} \right) \quad \text{e} \quad \left(\frac{-6}{\sqrt{11}}, \frac{-2}{\sqrt{11}}, \frac{2}{\sqrt{11}} \right)$$

É fácil ver que f tem valor menor no primeiro desses pontos; dessa forma, o ponto mais próximo é $\left(\frac{6}{\sqrt{11}}, \frac{2}{\sqrt{11}}, \frac{-2}{\sqrt{11}} \right)$ e o mais distante é $\left(\frac{-6}{\sqrt{11}}, \frac{-2}{\sqrt{11}}, \frac{2}{\sqrt{11}} \right)$.

Até aqui, apresentamos uma breve introdução aos conceitos fundamentais dos multiplicadores de Lagrange, exemplificando algumas aplicações em duas e três variáveis/dimensões.

Contudo, para desenvolvermos uma formulação mais ampla e alcançar uma demonstração formal para o Teorema do Multiplicador de Lagrange (Teorema 148), é necessário que estudemos e criemos uma base em conceitos de Análise no \mathbb{R}^n . No próximo capítulo, abordaremos conceitos da Topologia do espaço euclidiano, como conjuntos abertos e fechados, sequências e conjuntos compactos.

Capítulo 2

Topologia do Espaço \mathbb{R}^n

Neste capítulo, serão introduzidos conceitos fundamentais no estudo do espaço euclidiano \mathbb{R}^n . Começaremos com a definição de conjuntos abertos, fechados, sequências e compactos, que formam a base da topologia nesse espaço.

2.1 Conjuntos abertos

Nesta seção, apresentamos a definição de conjuntos abertos no espaço euclidiano \mathbb{R}^n e exploramos suas propriedades fundamentais. A noção de ponto interior é central, pois permite caracterizar um conjunto como aberto se todos os seus pontos forem interiores. Mostraremos que bolas abertas em \mathbb{R}^n são exemplos de conjuntos abertos, e discutimos como o conceito se estende para subconjuntos $X \subset \mathbb{R}^n$, onde um conjunto é dito “aberto em X ” ao satisfazer uma condição de inclusão de pontos de X em torno de cada um dos seus elementos.

Com base nessas definições, demonstraremos as principais propriedades dos conjuntos abertos: a estabilidade da abertura em interseções finitas e uniões arbitrárias, e que o vazio e \mathbb{R}^n inteiro são abertos. Além disso, também apresentamos o Teorema da Continuidade, que relaciona a continuidade de funções a imagens inversas de conjuntos abertos, além de exemplos e proposições sobre a abertura de produtos cartesianos e projeções.

Definição 6. *Seja X um subconjunto do espaço euclidiano \mathbb{R}^n . Um ponto $a \in X$ chama-se um ponto interior a X quando é centro de alguma bola aberta contida em X , ou seja, quando existe $\delta > 0$ tal que $|x - a| < \delta \Rightarrow x \in X$.*

Definição 7. *O interior de X é o conjunto $\text{int } X$, formado pelos pontos interiores a X .*

Definição 8. *Quando $X \in \text{int } V$, dizemos que o conjunto V é uma vizinhança do ponto x .*

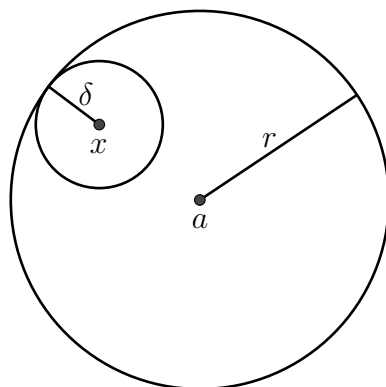
Dizer que um ponto $a \in X$ não é interior a X equivale a afirmar que toda bola aberta de centro a contém pontos do complementar de X , ou seja, que, para todo $\delta > 0$ existe $y \in \mathbb{R}^n - X$ com $|y - a| < \delta$.

Definição 9. Um conjunto $X \in \mathbb{R}^n$ chama-se aberto quando todos os seus pontos são interiores, isto é, quando para cada $x \in X$ existe $\delta > 0$ tal que $B(x; \delta) \subset X$. Assim, X é aberto $\iff \text{int } X = X$.

Proposição 10. Uma bola aberta é um exemplo de conjunto aberto.

Demonstração: Dado qualquer $x \in B(a; r)$, temos $|x - a| < r$, logo o número $\delta = r - |x - a|$ é positivo. Afirmamos que $B(x; \delta) \subset B(a; r)$. De fato, $y \in B(x; \delta) \Rightarrow |y - a| \leq |y - x| + |x - a| < \delta + |x - a| = r \Rightarrow y \in B(a; r)$. ■

Figura 2.1: Representação de uma bola aberta $B(a; r)$ e de uma bola menor $B(x; \delta) \subset B(a; r)$.



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Também é aberto em \mathbb{R}^n o conjunto $X = \mathbb{R}^n - B[a; r]$, complementar da bola fechada $B[a; r]$. Temos $X = \{x \in \mathbb{R}^n; |x - a| > r\}$. Dado arbitrariamente $x \in X$, seja $\delta = |x - a| - r$. Afirmamos que $B(x; \delta) \subset X$. Com efeito, $y \in B(x; \delta) \Rightarrow |x - a| \leq |x - y| + |y - a| < \delta + |y - a| = |x - a| - r + |y - a| \Rightarrow |y - a| > r \Rightarrow y \in X$.

Proposição 11. Para todo conjunto $X \in \mathbb{R}^n$, $\text{int } X$ é um conjunto aberto.

Demonstração: De fato, se $a \in \text{int } X$ então existe $r > 0$ tal que $B(a; r) \subset X$. Se $x \in B(a; r)$ então pondo $\delta = r - |x - a|$, vemos que $B(x; \delta) \subset B(a; r)$, donde $B(x; \delta) \subset X$ e portanto $x \in \text{int } X$. Assim, todo ponto $a \in \text{int } X$ é centro de bola $B(a; r)$ contida em $\text{int } X$, o que prova que $\text{int } X$ é aberto. ■

Observação 12. Uma bola fechada $B[a; r]$ em \mathbb{R}^n não é um conjunto aberto pois, se tomarmos arbitrariamente um vetor unitário $u \in \mathbb{R}^n$, o ponto $x = a + r \cdot u$ é tal que $|x - a| = r$, logo $x \in B[a; r]$. Mas nenhuma bola aberta $B(x; \delta)$ está contida em $B[a; r]$. Com efeito, tomando $y = a + (r + \delta/2)u$, temos $|y - x| = \delta/2 < \delta$ mas $|y - a| = r + \delta/2 > r$. Assim, $y \in B(x, \delta)$ mas $y \notin B[a; r]$. Este argumento mostra, de fato, que os pontos da esfera $S[a; r]$ não são interiores à bola fechada de centro a e raio r . Portanto, $\text{int } B[a; r] = B(a; r)$.

Dados um conjunto X e um ponto $a \in \mathbb{R}^n$, há três possibilidades que se excluem mutuamente: ou $a \in \text{int } X$, ou $a \in \text{int } \mathbb{R}^n - X$ ou então toda bola aberta de centro a contém pontos de X e pontos do complementar de X . Os pontos com esta última propriedade constituem ∂X , que chamaremos a *fronteira* de X . Os pontos $y \in \partial X$ são chamados *pontos fronteira* de X .

Assim, por exemplo, se X é a bola fechada $B[a; r]$, temos $\partial X = S[a; r]$ = esfera de centro a e raio r . Observe que se chamarmos de Y a bola aberta $B(a; r)$ teremos $\partial X = \partial Y$.

Um conjunto $A \subset \mathbb{R}^n$ é aberto se, e somente se, nenhum dos seus pontos é ponto fronteira de A , ou seja, se, e somente se, $A \cap \partial A = \emptyset$.

Teorema 13. Os conjuntos abertos do espaço euclidiano \mathbb{R}^n gozam das seguintes propriedades:

- 1) O conjunto vazio e o espaço \mathbb{R}^n inteiro são abertos;
- 2) A interseção $A = A_1 \cap \dots \cap A_k$ de um número finito de conjuntos abertos A_1, \dots, A_k é um conjunto aberto;
- 3) A reunião $A = \bigcup_{\lambda \in L} A_\lambda$ de uma família qualquer $(A_\lambda)_{\lambda \in L}$ de conjuntos abertos A_λ é um conjunto aberto.

Demonstração: Um conjunto só pode deixar de ser aberto se contiver algum ponto que não seja interior. Como \emptyset não contém ponto algum, é aberto. \mathbb{R}^n é obviamente aberto. Para provar 2), seja $a \in A$. Então, para cada $i = 1, \dots, k$, temos $a \in A_i$. Como A_i é aberto, existe $\delta_i > 0$ tal que $B(a; \delta_i) \subset A_i$. Seja $\delta = \min\{\delta_1, \dots, \delta_k\}$. Então, $B(a; \delta) \subset A_i$ para cada i , donde $B(a; \delta) \subset A$. Finalmente, provemos 3). Dado $a \in A$, existe $\lambda \in L$ tal que $a \in A_\lambda$. Sendo A_λ aberto, existe $\delta > 0$ com $B(a; \delta) \subset A_\lambda \subset A$. Logo, A é aberto. ■

Fixemos um conjunto $X \subset \mathbb{R}^n$.

Definição 14. Um subconjunto $A \subset X$ diz-se aberto em X quando, para cada $a \in A$ existe $\delta > 0$ tal que $B(a; \delta) \cap X \subset A$. Noutras palavras, para cada $a \in A$ existe $\delta > 0$ tal que os pontos x , pertencentes a X , que cumprem a condição $|x - a| < \delta$ estão em A .

Exemplo 15. $A = (0, 1]$ é aberto em $X = [0, 1]$.

Chega-se à noção de conjunto aberto em X quando se faz abstração dos demais pontos do espaço \mathbb{R}^n , considerando-se apenas os pontos de X , e se procura então imitar a definição de conjunto aberto.

Como se vê facilmente, quando $X \subset \mathbb{R}^n$ é aberto, um subconjunto $A \subset X$ é aberto em X se, e somente se, é aberto no sentido usual de \mathbb{R}^n .

Proposição 16. *Mais geralmente, um conjunto $A \subset X$ é aberto em X se, e somente se, existe um aberto $B \subset \mathbb{R}^n$ tal que $A = X \cap B$.*

Demonstração: Se for A aberto em X , tome B igual à reunião das bolas $B(a; \delta)$ com centro nos pontos $a \in A$, tais que $B(a; \delta) \cap X \subset A$. Reciprocamente, se for $A = X \cap B$, com B aberto em \mathbb{R}^n , para cada $a \in A$ existe uma bola $B(a; \delta) \subset B$, logo, $B(a; \delta) \cap X \subset B \cap X = A$. Portanto, A é aberto em X . ■

Vale para os abertos em X um resultado análogo ao do Teorema (13): \emptyset e X são abertos em X ; uma interseção finita e uma reunião qualquer de abertos em X é ainda um conjunto aberto em X .

O interesse desta nova noção reside no teorema abaixo.

Teorema 17. *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 $f^{-1}(A)$ 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 significa 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, $a \in A$. Logo, existe $\delta > 0$ tal que $B(a; \delta) \cap X \subset A$. Isto significa, porém, 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. ■

Observação 18. *O mesmo resultado continuaria válido se substituíssemos, no enunciado acima, 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, embora aparentemente mais geral, reduz-se ao que foi demonstrado no Teorema (17).*

Como aplicação do teorema acima, vemos que se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função contínua, então, para todo número real a , o conjunto $A = \{x \in \mathbb{R}^n; f(x) < a\}$ é aberto, pois A é a

imagem inversa do intervalo aberto $(-\infty, a)$ pela função f . Usando também o fato de que a interseção finita de abertos ainda é um aberto, concluímos que, dadas k funções contínuas $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$, o conjunto $A = \{x \in \mathbb{R}^n; f_1(x) < a_1, \dots, f_k(x) < a_k\}$ é aberto. A mesma conclusão vale se as funções f_1, \dots, f_k forem definidas num subconjunto aberto $X \subset \mathbb{R}^n$.

Reobtemos assim o fato de que a bola $B(a; r)$ é um subconjunto aberto de \mathbb{R}^n . Com efeito, a função $x \mapsto |x - a|$ é contínua e $B(a; r)$ é o conjunto dos pontos de \mathbb{R}^n onde esta função assume valores $< r$.

Proposição 19. *Se $A \subset \mathbb{R}^m$ e $B \subset \mathbb{R}^n$ são abertos, então o produto cartesiano $A \times B \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ é aberto.*

Demonstração: Considerando as projeções $\pi_1 : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $\pi_2 : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, que são aplicações contínuas, temos $A \times B = \pi_1^{-1}(A) \cap \pi_2^{-1}(B)$. Evidentemente, se tivermos $A_1 \subset \mathbb{R}^{m_1}, \dots, A_k \subset \mathbb{R}^{m_k}$ abertos, o produto cartesiano $A_1 \times \dots \times A_k \subset \mathbb{R}^{m_1 + \dots + m_k}$ será aberto por uma razão semelhante. ■

Definição 20. *Dados $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, uma aplicação $f : X \rightarrow Y$ diz-se aberta quando, para cada $A \subset X$ aberto em X , sua imagem $f(A)$ é um subconjunto aberto em Y .*

Proposição 21. *Cada uma das m projeções $\pi_i : \mathbb{R}^m \rightarrow \mathbb{R}$ é uma aplicação aberta.*

Demonstração: Isto fica mais fácil de ver se usarmos em \mathbb{R}^m a norma do máximo, em relação a qual uma bola aberta de centro $a = (a_1, \dots, a_m)$ é um produto cartesiano $B(a; \delta) = \prod_{j=1}^m (a_j - \delta, a_j + \delta)$.

Seja então $A \subset \mathbb{R}^m$ aberto. Para provar que $\pi_i(A)$ é aberto em \mathbb{R} , tomemos $a_i = \pi_i(a) \in \pi_i(A)$, $a \in A$. Como A é aberto, existe $\delta > 0$ tal que $B = \prod_{j=1}^m (a_j - \delta, a_j + \delta) \subset A$. Então $(a_i - \delta, a_i + \delta) = \pi_i(B) \subset \pi_i(A)$, logo $\pi_i(A)$ é aberto. ■

Na próxima seção, exploraremos os conjuntos fechados, que complementam os conjuntos abertos e trazem novos aspectos fundamentais à análise topológica.

2.2 Conjuntos fechados

Nesta seção, é apresentada a definição de conjuntos fechados no espaço euclidiano \mathbb{R}^n e são discutidas as suas propriedades fundamentais. Além disso, traremos outras definições importantes, tais como: fecho, densidade, etc.

Definição 22. Um ponto $a \in \mathbb{R}^n$ diz-se aderente a um conjunto $X \in \mathbb{R}^n$ quando é limite de uma sequência de pontos desse conjunto.

Por exemplo, todo ponto $a \in X$ é aderente a X pois podemos escrever $a = \lim x_k$, com $x_k = a$ para todo $k \in \mathbb{N}$. Mas a pode ser aderente a X sem pertencer a X ; neste caso, a é necessariamente um ponto de acumulação do conjunto X . Por exemplo, se $X = B(0; 1) \subset \mathbb{R}^n$ é a bola aberta de centro na origem e raio 1 em \mathbb{R}^n , o ponto $e_1 = (1, 0, \dots, 0)$ não pertence a X . Mas, pondo $x_k = (1 - 1/k, 0, \dots, 0)$, vemos que $x_k \in X$ para todo $k \in \mathbb{N}$ e $\lim x_k = e_1$, logo, e_1 é aderente a X .

Afim de que o ponto a seja aderente ao conjunto X , é necessário e suficiente que toda bola aberta de centro a contenha algum ponto de X . Com efeito, a condição é necessária em virtude da definição de limite de uma sequência e é suficiente porque, se ela se verifica, em cada bola $B(a; 1/k)$ podemos escolher um ponto $x_k \in X$ e assim obtemos uma sequência com $\lim x_k = a$.

Definição 23. O conjunto dos pontos aderentes a X chama-se o fecho de X e é indicado com a notação \overline{X} .

Pelo que vemos acima, afim de que um ponto $b \in \mathbb{R}^n$ não pertença ao fecho de X , é necessário e suficiente que exista uma bola aberta de centro b que não contém pontos de X . Noutros termos, $b \in \mathbb{C}\overline{X} \iff \exists r > 0; B(b; r) \cap X = \emptyset$.

Como toda bola aberta é um conjunto aberto e todo aberto que contém um ponto contém também uma bola aberta com centro nesse ponto, as condições acima podem ser reformuladas com abertos, em vez de bolas:

1. Tem-se $a \in \overline{X}$ se, e somente se, todo aberto que contém a intersecta o conjunto X . (Isto é, A é aberto, $a \in A \Rightarrow A \cap X \neq \emptyset$.)
2. Tem-se $b \notin \overline{X}$ se, e somente se, existe um aberto contendo b e disjunto de X . (Isto é, existe A aberto com $b \in A$ e $A \cap X = \emptyset$.)

O fecho de uma bola aberta $B(a; r)$ é a bola fechada $B[a; r]$. Se $X = \mathbb{Q}^n$ é o conjunto dos pontos de \mathbb{R}^n cujas coordenadas são números racionais, então $\overline{X} = \mathbb{R}^n$.

Definição 24. Um conjunto $X \subset \mathbb{R}^n$ chama-se fechado quando contém todos os seus pontos aderentes, isto é, quando $X = \overline{X}$.

Dizer que $X \subset \mathbb{R}^n$ é fechado significa, portanto, o seguinte: se $\lim x_k = a$ e $x_k \in X$ para todo $k \in \mathbb{N}$, então $a \in X$.

Por exemplo, uma bola fechada $B[a; r]$ é um subconjunto fechado do espaço \mathbb{R}^n pois se $|x_k| \leq r$ para todo k e $\lim x_k = b$ então $|b| = \lim |x_k| \leq r$.

Daí resulta que o fecho de todo conjunto limitado $X \subset \mathbb{R}^n$ é limitado. Com efeito, temos $X \subset B$, onde B é uma bola fechada. Logo, $\overline{X} \subset \overline{B} = B$, donde \overline{X} é limitado.

Para todo $X \subset \mathbb{R}^n$, o complementar do fecho de X é um aberto. Com efeito, seja $A = \mathbb{C}\overline{X}$. Para todo $b \in A$ existe $r > 0$ tal que $B(b, r) \cap X = \emptyset$. 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 X , logo $y \in \mathbb{C}\overline{X} = A$.

Em particular, se $X \subset \mathbb{R}^n$ é um conjunto fechado, seu complementar \overline{X} é aberto em \mathbb{R}^n (pois $\mathbb{C}X = \mathbb{C}\overline{X}$).

Reciprocamente, se $X \subset \mathbb{R}^n$ é tal que $A = \mathbb{C}X$ 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.

Acabamos portanto de demonstrar o

Teorema 25. *Um conjunto é fechado se, e somente se, seu complementar é aberto.*

Corolário 26. *O fecho de todo conjunto é um conjunto fechado.*

O Corolário acima diz que $\overline{\overline{X}} = \overline{X}$ para todo $X \subset \mathbb{R}^n$.

O Teorema (25) faz com que o resultado seguinte seja uma consequência imediata do Teorema (13).

Teorema 27. *Os conjuntos fechados do espaço euclídeano \mathbb{R}^n gozam das seguintes propriedades:*

- 1) *O conjunto \emptyset e o espaço inteiro \mathbb{R}^n são fechados;*
- 2) *A reunião $F = F_1 \cup \dots \cup F_k$ de um número finito de conjuntos fechados F_1, \dots, F_k é um conjunto fechado;*
- 3) *A interseção $F = \bigcap_{\lambda \in L} F_\lambda$ de uma família qualquer $(F_\lambda)_{\lambda \in L}$ de conjuntos fechados F_λ é um conjunto fechado.*

Demonstração: A afirmação 1) é evidente. Quanto a 2), se F_1, \dots, F_k são fechados, então $A_1 = \mathbb{C}F_1, \dots, A_k = \mathbb{C}F_k$ são abertos, portanto, $A_1 \cap \dots \cap A_k$ é aberto. Logo, $F_1 \cup \dots \cup F_k = \mathbb{C}(A_1 \cap \dots \cap A_k) = \mathbb{C}A$ é fechado. Finalmente, se cada $F_\lambda, \lambda \in L$, é fechado então cada $A_\lambda = \mathbb{C}F_\lambda$ é aberto, logo $A = \bigcup_{\lambda \in L} A_\lambda$ também é aberto. Sendo assim, o conjunto

$$F = \bigcap_{\lambda \in L} F_\lambda = \bigcap_{\lambda \in L} \mathbb{C}A_\lambda = \mathbb{C} \left(\bigcup_{\lambda \in L} A_\lambda \right) = \mathbb{C}A \text{ é fechado.}$$

Note que uma reunião finita de conjuntos fechados pode ser fechado ou não. De fato, para cada ponto $x \in \mathbb{R}^n$, o conjunto $\{x\}$ é fechado. Ora, todo conjunto $X \subset \mathbb{R}^n$ é reunião dos seus pontos: $X = \bigcup_{x \in X} \{x\}$. Como há conjuntos em \mathbb{R}^n que não são fechados, há reuniões (infinitas) de conjuntos fechados que não são fechados.

Segue-se da definição de fronteira que um ponto a pertence à fronteira do conjunto X se, e somente se, a é aderente a X e a $\mathbb{R}^n - X$. Ou seja, $\partial X = \overline{X} \cap \overline{(\mathbb{R}^n - X)}$. Em particular, a fronteira de todo conjunto $X \subset \mathbb{R}^n$ é um conjunto fechado.

Fixemos um conjunto $X \subset \mathbb{R}^n$.

Definição 28. Um subconjunto $F \subset X$ diz-se fechado em X quando se tem $F = X \cap G$, onde G é um conjunto fechado em \mathbb{R}^n . Afim de que o subconjunto $F \subset X$ seja fechado em X é necessário e suficiente que F contenha todos os seus pontos aderentes que pertençam a X .

Se $X \subset \mathbb{R}^n$ é fechado, então um subconjunto $F \subset X$ é fechado em X se, e somente se, é fechado em \mathbb{R}^n .

Os conjuntos fechados em X gozam de propriedades análogas às que foram demonstradas no Teorema 27 para os fechados em \mathbb{R}^n . A saber: \emptyset e X são fechados em X ; uma reunião finita ou interseção arbitrária de fechados em X é ainda um conjunto fechado em X .

Evidentemente, se $F \subset \mathbb{R}^n$ é fechado e $F \subset X$ então F é fechado em X . Seja $X = \{x \in \mathbb{R}^n; x > 0\}$ a semirreta positiva aberta. O intervalo semiaberto $(0, 1]$ é fechado em X .

Proposição 29. Seja $F \subset X$. Afim de que F seja fechado em X é necessário e suficiente que o conjunto $A = X - F$ (complementar de F relativamente a X) seja aberto em X .

Demonstração: Dados $F' \subset \mathbb{R}^n$ e $A' = \complement F'$, o complementar de F' em \mathbb{R}^n , temos $F = X \cap F' \iff X - F = X \cap A'$. Ora, F' é fechado em \mathbb{R}^n se, e somente se, A' é aberto. Assim, F é fechado em X se, e somente se, $X - F$ é aberto em X . ■

A Proposição acima foi dada para que tivéssemos o

Teorema 30. Seja $f : X \rightarrow \mathbb{R}^n$ uma aplicação definida no subconjunto $X \subset \mathbb{R}^m$. Afim de que f seja contínua, é necessário e suficiente que a imagem inversa $f^{-1}(F)$ de todo conjunto fechado $F \subset \mathbb{R}^n$ seja um conjunto fechado em X .

Demonstração: Isto decorre do teorema análogo para abertos, que é o Teorema (17), juntamente com o fato de que os conjuntos abertos (em X) são exatamente aqueles cujos complementares são fechados (em X). Com efeito, seja f contínua. Então, para cada $F \subset \mathbb{R}^n$ fechado, temos $F = \complement A$, onde $A \subset \mathbb{R}^n$ é aberto. Logo, $f^{-1}(A)$ é aberto em X . Mas $f^{-1}(F) = f^{-1}(\complement A) = X - f^{-1}(A)$, portanto, $f^{-1}(F)$ é fechado em X . Reciprocamente,

se a imagem inversa por f de todo fechado em \mathbb{R}^n é um fechado em X , então a relação $f^{-1}(A) = X - f^{-1}(F)$, com $A = \mathbb{C}F$, mostra que a imagem inversa por f de todo aberto $A \subset \mathbb{R}^n$ é um aberto em X . Portanto, f é contínua. ■

Assim, por exemplo, se $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ são funções reais contínuas, então, dados os números a_1, \dots, a_k , o conjunto X dos pontos $x \in \mathbb{R}^n$ para os quais se tem simultaneamente $f_1(x) \leq a_1, \dots, f_k(x) \leq a_k$, é um conjunto fechado de \mathbb{R}^n .

Em particular, fixado um ponto $a \in \mathbb{R}^n$, a função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, dada por $f(x) = |x - a|$, é contínua. Logo, para todo $r > 0$, a bola fechada $B[a; r] = \{x \in \mathbb{R}^n; f(x) \leq r\}$ é um conjunto fechado.

Nas mesmas condições acima, é fechado o conjunto Y dos pontos $y \in \mathbb{R}^n$ para os quais se tem simultaneamente $f_1(y) = a_1, \dots, f_k(y) = a_k$. Com efeito, vale $Y = f^{-1}(a)$, onde $f = (f_1, \dots, f_k) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ e $a = (a_1, \dots, a_k)$.

Em particular, a esfera $S[a; r] = \{y \in \mathbb{R}^n; |y - a| = r\}$, é um subconjunto fechado de \mathbb{R}^n .

Observação 31. Se $f : X \rightarrow \mathbb{R}^n$ é contínua e $F \subset f(X)$ é um conjunto apenas fechado em $f(X)$, sua imagem inversa $f^{-1}(F)$ é ainda fechada em X . Com efeito, temos $F = F' \cap f(X)$, onde F' é fechado em \mathbb{R}^n . Como $f^{-1}(F) = f^{-1}(F') \cap f^{-1}(f(X))$, segue-se do Teorema (30) que $f^{-1}(F)$ é fechada em X .

Proposição 32. Se $F \subset \mathbb{R}^m$ e $G \subset \mathbb{R}^n$ são subconjuntos fechados, então o produto cartesiano $F \times G \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ é fechado.

Demonstração: As projeções $\pi_1 : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $\pi_2 : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, definidas por $\pi_1(x, y) = x$ e $\pi_2(x, y) = y$, são contínuas e $F \times G = \pi_1^{-1}(F) \cap \pi_2^{-1}(G)$. ■

Exemplo 33. Um exemplo de conjunto fechado é o gráfico de uma aplicação contínua $f : X \rightarrow \mathbb{R}^n$, definida num conjunto $X \subset \mathbb{R}^m$. Como sabemos, o gráfico de f é o subconjunto $G \subset X \times \mathbb{R}^n \subset \mathbb{R}^m \times \mathbb{R}^n$, formado pelos pontos $(x, f(x)) \in \mathbb{R}^m \times \mathbb{R}^n$, onde x varia em X . Afirmamos que G é um subconjunto fechado em $X \times \mathbb{R}^n$. (Quando $X \subset \mathbb{R}^m$ for fechado, então G será fechado no espaço euclidiano \mathbb{R}^{m+n}). Com isso, temos $G = \{(x, y) \in X \times \mathbb{R}^n; y - f(x) = 0\}$. Como a aplicação $\varphi : X \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, dada por $\varphi(x, y) = y - f(x)$, é contínua, resulta que $G = \varphi^{-1}(0)$ é fechado em $X \times \mathbb{R}^n$.

Ao contrário do que ocorre para conjuntos abertos, a projeção $\pi_1 : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ não transforma necessariamente um conjunto fechado $F \subset \mathbb{R}^m \times \mathbb{R}^n$ num conjunto fechado $\pi_1(F) \subset \mathbb{R}^m$. Por exemplo, a hipérbole $H = \{(x, y) \in \mathbb{R}^2; xy = 1\}$ é um subconjunto fechado do plano, imagem inversa do fechado $\{1\} \subset \mathbb{R}$ pela aplicação contínua $(x, y) \mapsto xy$. Mas sua projeção no eixo das abscissas, $\pi_1(H) = \mathbb{R} - \{0\}$, não é fechada.

Dados $Y \subset X \subset \mathbb{R}^n$, podemos também definir o *fecho de Y relativamente a X* como sendo o conjunto $\bar{Y} \cap X$, dos pontos aderentes a Y que pertencem ao conjunto X . Então, Y é fechado em X se, e somente se, coincide com seu fecho relativamente a X .

Um caso particular importante se dá quando o fecho de Y relativamente a X é todo o conjunto X . Para descrever esta situação, damos a definição abaixo.

Definição 34. *Sejam $Y \subset X \subset \mathbb{R}^n$. Dizemos que Y é denso em X quando $\bar{Y} \cap X = X$, ou seja, $X \subset \bar{Y}$.*

Isto significa que todo ponto de X é limite de uma sequência cujos termos pertencem a Y . Ou, ainda de outra maneira: dado $Y \subset X$, tem-se Y denso em X se, e somente se, toda bola aberta com centro em algum ponto de X contém pontos de Y .

Os fatos mais importantes a respeito de conjuntos densos em \mathbb{R}^n são as seguintes observações.

Observação 35. *Sejam $f, g : X \rightarrow \mathbb{R}^m$ aplicações contínuas num conjunto $X \subset \mathbb{R}^n$ e $Y \subset X$ um subconjunto denso. Se $f(y) = g(y)$ para todo $y \in Y$, então $f = g$ (isto é, $f(x) = g(x)$ para todo $x \in X$).*

Com efeito, para cada $x \in X$ existe uma sequência de pontos $y_k \in Y$ com $\lim y_k = x$. Então, $f(x) = f(\lim y_k) = \lim f(y_k) = \lim g(y_k) = g(\lim y_k) = g(x)$.

Observação 36. *Todo conjunto $X \subset \mathbb{R}^n$ contém um subconjunto enumerável E , denso em X .*

Com efeito, a coleção \mathcal{B} das bolas abertas $B(q; r)$, com centro num ponto $q \in \mathbb{Q}^n$ e raio racional, é enumerável: $\mathcal{B} = \{B_1, \dots, B_i, \dots\}$. Para cada $i \in \mathbb{N}$, escolhemos um ponto $x_i \in B_i \cap X$, se $B_i \cap X$ não for vazio. Caso $B_i \cap X$ seja \emptyset , x_i não existirá. O conjunto E dos pontos x_i assim obtidos é um subconjunto enumerável de X . Para mostrar que E é denso em X , tomemos arbitrariamente $x \in X$ e $\epsilon > 0$. Existe $r > 0$ racional com $2r < \epsilon$. Como \mathbb{Q}^n é denso em \mathbb{R}^n , encontramos $q \in \mathbb{Q}^n$ tal que $|q - x| < r$. Logo, $x \in B(q; r) = B_i$. Assim, $B_i \cap X \neq \emptyset$. Existe, portanto, $x_i \in E$. Como x e x_i pertencem à bola B_i , de raio r , temos $|x - x_i| < 2r < \epsilon$. Ficou provado que toda bola aberta $B(x; \epsilon)$ com centro em algum ponto de X contém um ponto $x_i \in E$. Então, E é denso em X .

Os conjuntos fechados complementam os abertos e desempenham um papel importante na compreensão de continuidade e limites. Essa base topológica abre caminho para o estudo de sequências e suas propriedades de convergência, que veremos na próxima seção.

2.3 Sequências no espaço euclidiano

O estudo de sequências e de sua convergência é fundamental para a análise de limites e continuidade. Nesta seção, são exploradas como as propriedades das sequências em \mathbb{R}^n permitem descrever o comportamento dos pontos de uma função ao longo do tempo. O Teorema

de Bolzano-Weierstrass, que garante a existência de subsequências convergentes em sequências limitadas, é central para entender como esses pontos podem acumular-se, formando a base para diversas noções de compactação e continuidade.

Definição 37. *Uma sequência em \mathbb{R}^n é uma aplicação $x : \mathbb{N} \rightarrow \mathbb{R}^n$, definida no conjunto \mathbb{N} dos números naturais. O valor que essa aplicação assume no valor k é indicado com x_k e é chamado o k -ésimo termo da sequência.*

Usaremos as notações (x_k) , $(x_k)_{k \in \mathbb{N}}$ ou $(x_1, x_2, \dots, x_k, \dots)$ para indicar a sequência cujo k -ésimo termo é $x_k \in \mathbb{R}^n$.

Definição 38. *Uma subsequência de (x_k) é a restrição da sequência a um subconjunto infinito $\mathbb{N}' = \{k_1 < k_2 < \dots < k_i < \dots\} \subset \mathbb{N}$. A subsequência é indicada pelas notações $(x_k)_{k \in \mathbb{N}'}$, $(x_{k_i})_{i \in \mathbb{N}}$ ou $(x_{k_1}, x_{k_2}, \dots, x_{k_i}, \dots)$.*

Definição 39. *Diz-se que a sequência (x_k) é limitada quando o conjunto dos seus termos é limitado em \mathbb{R}^n , ou seja, quando existe um número real $c > 0$ tal que $|x_k| \leq c$ para todo $k \in \mathbb{N}$.*

Uma sequência $(x_k) \in \mathbb{R}^n$ equivale a n sequências de números reais. Com efeito, para cada $k \in \mathbb{N}$ temos $x_k = (x_{k_1}, x_{k_2}, \dots, x_{k_n})$, onde $x_{k_i} = \pi_i(x_k) = i$ -ésima coordenada de $x_k(1, 2, \dots, n)$. As n sequências $(x_{k_i})_{k \in \mathbb{N}}$ ($i = 1, \dots, n$) são chamadas as *sequências das coordenadas* de (x_k) . Assim, por exemplo, no plano \mathbb{R}^2 , uma sequência de pontos $z_k = (x_k, y_k)$ é o mesmo que um par de sequências (x_k) , (y_k) de números reais.

Sabe-se que uma sequência (x_k) em \mathbb{R}^n é limitada se, e somente se, cada uma de suas sequências de coordenadas $(x_{k_i})_{k \in \mathbb{N}}$ ($1 \leq i \leq n$) é limitada em \mathbb{R} .

Definição 40. *Diz-se que o ponto $a \in \mathbb{R}^n$ é o limite da sequência de pontos $x_k \in \mathbb{R}^n$ quando, para todo $\varepsilon > 0$ dado, é possível obter $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |x_k - a| < \varepsilon$.*

Neste caso, diz-se também que (x_k) converge para a ou tende para a e escreve-se $\lim x_k = a$, $\lim_{k \rightarrow \infty} x_k = a$, $\lim_{k \in \mathbb{N}} x_k = a$, ou simplesmente $x_k \rightarrow a$.

Definição 41. *Quando existe o limite $a = \lim x_k$, diz-se que a sequência (x_k) é convergente. Caso contrário, diz-se que (x_k) é divergente.*

Exemplo 42. *Uma sequência constante (a, a, \dots, a, \dots) é obviamente convergente e seu limite é a . Por outro lado, se $a \neq b$, então (a, b, a, b, \dots) é uma sequência divergente.*

Tem-se $\lim x_k = a \iff \lim |x_k - a| = 0$. Isto reduz a convergência em \mathbb{R}^n à convergência de números reais ≥ 0 .

Observação 43. Em termos de bolas, tem-se $\lim x_k = a$ se, e somente se, qualquer bola aberta de centro a contém todos os termos x_k salvo, possivelmente para um número finito de índices k . Com efeito, se $\varepsilon > 0$ é o raio da bola e k_0 é o número natural que corresponde a ε na definição de limite, fora da bola $B(a; \varepsilon)$ só poderão estar, no máximo, alguns dos termos x_1, \dots, x_{k_0} .

Observação 44. Resulta da observação acima que toda sequência convergente é limitada. De fato, se $\lim x_k = a$ então fora da bola aberta de centro a e raio 1 existem no máximo os termos x_1, \dots, x_{k_0} da sequência. Se r é o maior dos números $1, |x_1 - a|, \dots, |x_{k_0} - a|$, vemos que todos os termos da sequência estão contidos na bola $B[a; r]$. A recíproca é falsa: se $a \neq b$, a sequência (a, b, a, b, \dots) é divergente e limitada.

Observação 45. Segue-se também da caracterização do limite por meio de bolas que se $\lim x_k = a$, então toda subsequência de (x_k) tem ainda limite igual a a . Ou seja: toda subsequência de uma sequência convergente é ainda convergente e tem o mesmo limite.

Observação 46. Outro fato elementar, porém essencial, é que o limite de uma sequência convergente é único. Ou seja, se $\lim x_k = a$ e $\lim x_k = b$, então $a = b$. Com efeito, para todo $k \in \mathbb{N}$, temos:

$$0 \leq |a - b| \leq |x_k - a| + |x_k - b|.$$

Logo, $\lim |x_k - a| = \lim |x_k - b| = 0 \Rightarrow a = b$.

Em particular, se $\lim x_k = a$ e uma subsequência de (x_k) converge para o ponto $b \in \mathbb{R}^n$, então $a = b$.

A definição de limite de uma sequência em \mathbb{R}^n faz uso de uma norma. As desigualdades que relacionam as 3 normas usuais do espaço euclidiano nos dão, entretanto:

$$|x_k - a|_M \leq |x_k - a| \leq |x_k - a|_S \leq n|x_k - a|_M.$$

Segue-se daí que $\lim |x_k - a|_M = 0 \iff \lim |x_k - a| = 0 \iff \lim |x_k - a|_S = 0$. Portanto, a afirmação $\lim x_k = a$ independe de qual das três normas usuais estamos considerando. (No final deste parágrafo, mostraremos que noção de limite de uma sequência em \mathbb{R}^n permanece a mesma seja qual for a norma que considerarmos.)

Na demonstração do teorema abaixo, usamos em \mathbb{R}^n a norma do máximo, por conveniência. Resulta do que acabamos de dizer que o fato nele enunciado é válido seja qual for a norma, dentre as 3 usuais, que seja tomada na definição de limite.

Teorema 47. Uma sequência (x_k) em \mathbb{R}^n converge para o ponto $a = (a_1, \dots, a_n)$ se, e somente se, para cada $i = 1, 2, \dots, n$, tem-se $\lim_{k \rightarrow \infty} x_{k_i} = a_i$, ou seja, cada coordenada de x_k converge para a coordenada correspondente de a .

Demonstração: Como $|x_{k_i} - a_i| \leq |x_k - a|$, vemos que $\lim_{k \rightarrow \infty} x_k = a \Rightarrow \lim_{k \rightarrow \infty} x_{k_i} = a_i$ para todo $i = 1, 2, \dots, n$, então, dado $\varepsilon > 0$, existem números naturais k_1, \dots, k_n tais que $k > k_i \Rightarrow |x_k - a| = \max_i |x_{k_i} - a| < \varepsilon$. Logo, $\lim x_k = a$. ■

Corolário 48. *Dadas as seqüências convergentes de pontos $x_k, y_k \in \mathbb{R}^n$ e $\alpha_k \in \mathbb{R}$, sejam $\lim x_k = a$, $\lim y_k = b$ e $\lim \alpha_k = \alpha$. Então:*

- 1) $\lim(x_k + y_k) = a + b$;
- 2) $\lim \alpha_k \cdot x_k = \alpha \cdot a$;
- 3) $\lim \langle x_k, y_k \rangle = \langle a, b \rangle$;
- 4) $\lim |x_k| = |a|$.

Demonstração: Utilizando os fatos conhecidos sobre limites de somas v de produtos de números reais, vemos que, para cada $i = 1, \dots, n$, valem $\lim_{k \rightarrow \infty} (x_{k_i} + y_{k_i}) = a_i + b_i$ e $\lim_{k \rightarrow \infty} \alpha_k x_{k_i} = \alpha \cdot a_i$. Em vista do Teorema (47), as igualdades 1) e 2) ficam demonstradas. Analogamente, $\lim_{k \rightarrow \infty} \langle x_k, y_k \rangle = \lim_{k \rightarrow \infty} (x_{k_1} y_{k_1} + \dots + x_{k_n} y_{k_n}) = a_1 b_1 + \dots + a_n b_n = \langle a, b \rangle$. Isto prova 3). Finalmente, $\lim |x_k| = \lim \sqrt{\langle x_k, x_k \rangle} = \sqrt{\lim \langle x_k, x_k \rangle} = \sqrt{\langle a, a \rangle} = |a|$, o que prova 4). Também poderíamos provar 4) observando que $||x_k| - |a|| \leq |x_k - a|$. Esta maneira tem a vantagem de valer para qualquer norma, euclidiana ou não. ■

Teorema 49. (Bolzano-Weierstrass.) *Toda seqüência limitada em \mathbb{R}^n possui uma subseqüência convergente.*

Demonstração: Sabemos que o Teorema de Bolzano-Weierstrass é válido na reta: toda seqüência limitada de números reais possui uma subseqüência convergente. Consulte [3]. Dada a seqüência limitada (x_k) em \mathbb{R}^n , as primeiras coordenadas dos seus termos formam uma seqüência limitada $(x_{k_1})_{k \in \mathbb{N}}$ de números reais, a qual possui uma subseqüência convergente. Isto é, existem num subconjunto infinito $\mathbb{N}_1 \subset \mathbb{N}$ e um número real a_1 tais que $\lim_{k \in \mathbb{N}_1} x_{k_1} = a_1$. Por sua vez, a seqüência limitada $(x_{k_2})_{k \in \mathbb{N}_1}$ de números reais, possui uma subseqüência convergente; podemos obter um subconjunto infinito $\mathbb{N}_2 \subset \mathbb{N}_1$ e $a_2 \in \mathbb{R}$ tais que $\lim_{k \in \mathbb{N}_2} x_{k_2} = a_2$. E assim por diante, até encontrarmos conjuntos infinitos $\mathbb{N} \supset \mathbb{N}_1 \supset \mathbb{N}_2 \supset \dots \supset \mathbb{N}_n$ e números reais a_1, \dots, a_n tais que $\lim_{k \in \mathbb{N}_i} x_{k_i} = a_i$ para $i = 1, 2, \dots, n$. Então pomos $a = (a_1, \dots, a_n)$ e vemos, pelo Teorema (47), que $\lim_{k \in \mathbb{N}_n} x_k = a$, o que conclui a demonstração. ■

A análise de convergência em seqüências, especialmente com o suporte do Teorema de Bolzano-Weierstrass, fornece uma base sólida para o estudo de limites e continuidade. Esses conceitos são essenciais para avançarmos na compreensão dos conjuntos compactos, tema da próxima seção.

2.4 Conjuntos compactos

Nesta seção, são discutidos os conjuntos compactos. Uma característica importante desses conjuntos é que, por meio do Teorema de Bolzano-Weierstrass, toda sequência limitada dentro de um conjunto compacto possui uma subsequência convergente. Essa propriedade permite garantir a existência de pontos de acumulação e torna os conjuntos compactos ferramentas essenciais para a continuidade e para diversos teoremas de existência em análise.

Definição 50. *Diremos que um conjunto $K \in \mathbb{R}^n$ é compacto quando ele for limitado e fechado.*

Exemplo 51. *São compactas todas as esferas e bolas fechadas do espaço euclidiano, mas o espaço \mathbb{R}^n inteiro não é compacto (salvo se $n = 0$!).*

Em virtude do Teorema de Bolzano-Weierstrass (Teorema 49), 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 seguintes propriedades decorrem imediatamente da definição:

1. K_1, \dots, K_p compactos em $\mathbb{R}^n \Rightarrow K_1 \cup \dots \cap K_p$ 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 $K \times L \subset \mathbb{R}^{m+n}$ é compacto.

Menos óbvio é o seguinte fato, conhecido como a *propriedade de Cantor*:

4. Dada uma sequência decrescente $K_1 \supset \dots \supset K_k$ de compactos não vazios, a interseção $K = \bigcap_{k=1}^{\infty} K_k$ (é compacta e) não é vazia.

Demonstração: Escolhamos, para cada $k \in \mathbb{N}$, um ponto $x_k \in K_k$, obtendo assim uma sequência, a qual possui uma subsequência $(x_{k_1}, \dots, x_{k_i}, \dots)$ convergindo para um ponto $x \in \mathbb{R}^n$. Dado arbitrariamente $k \in \mathbb{N}$, temos $x_{k_i} \in K_k$ para todo $k_i > k$, logo $x = \lim x_{k_i} \in K_k$. Assim, o ponto x pertence a K_k para todo $k \in \mathbb{N}$, ou seja, $x \in K = \bigcap_k K_k$, o que mostra que K não é vazio. ■

Teorema 52. *Seja $f : X \rightarrow \mathbb{R}^n$ contínua no conjunto $X \subset \mathbb{R}^n$. 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 pois $y \in \mathbb{R}^n$ aderente a $f(K)$. Então $y = \lim f(x_k)$, $x_k \in K$ para todo $k \in \mathbb{N}$. Pela compacidade de K , uma subsequência (x_{k_i}) converge para um certo ponto $x \in K$. Segue-se 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, para cada $k \in \mathbb{N}$, obter um ponto $x_k \in K$ tal 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 dá então $f(x) = f(\lim x_{k_i}) = \lim f(x_{k_i})$, uma contradição. ■

Note que uma aplicação contínua pode transformar um conjunto limitado num conjunto ilimitado ($f(x) = 1/x$ leva o intervalo aberto $(0, 1)$ no intervalo infinito $(1, +\infty)$) ou um conjunto fechado num conjunto não fechado ($f(x) = 1/(1+x^2)$ leva \mathbb{R} no intervalo $(0, 1]$). Mas se $X \subset \mathbb{R}^m$ for limitado e fechado, sua imagem por qualquer aplicação contínua é limitada e fechada.

Corolário 53. (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 $x \in K$.

Demonstração: Temos que $f(K) \subset \mathbb{R}$ 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 $x \in K$. ■

A função $f : \mathbb{R} \rightarrow \mathbb{R}$, definida por $f(x) = x/(1 + |x|)$, é contínua, cumpre $-1 < f(x) < 1$ para todo $x \in \mathbb{R}$, mas nenhum valor $f(x)$ é menor nem maior do que todos os demais valores de f .

Em particular, toda aplicação contínua $f : K \rightarrow \mathbb{R}^n$, definida num compacto $K \subset \mathbb{R}^m$, é limitada, isto é, existe $c > 0$ tal que $|f(x)| \leq c$ para todo $x \in K$.

Note ainda que se $f : K \rightarrow \mathbb{R}$ é contínua no compacto $K \subset \mathbb{R}^n$ e $f(x) > 0$ para todo $x \in K$, então existe $\varepsilon > 0$ tal que $f(x) \geq \varepsilon$ para todo $x \in K$. Isto seria falso se K não fosse compacto: bastaria tomar $f : (0, +\infty) \rightarrow \mathbb{R}$, dada por $f(x) = 1/x$.

Corolário 54. Seja $K \subset \mathbb{R}^m$ compacto. Toda aplicação contínua $f : K \rightarrow \mathbb{R}^n$ é fechada, isto é, $F \subset K$ fechado $\Rightarrow f(F) \subset \mathbb{R}^n$ fechado.

Demonstração: $F \subset K$ fechado $\Rightarrow F$ compacto $\Rightarrow f(F)$ compacto $\Rightarrow f(F)$ compacto em \mathbb{R}^n .

Corolário 55. A inversa de uma bijeção contínua definida num compacto é contínua. (Ou seja: toda bijeção contínua definida num compacto é um homeomorfismo sobre sua imagem.)

Demonstração: Seja $f : K \rightarrow L$ uma bijeção contínua do compacto $K \subset \mathbb{R}^m$ no compacto $L \subset \mathbb{R}^n$. Pondo $g = f^{-1} : L \rightarrow K$ vemos que, para todo $F \subset K$ fechado, que imagem inversa $g^{-1}(F) = f(F)$ é fechada em \mathbb{R}^n , pelo Corolário 54. Segue-se do Teorema 30 que g é contínua.

Vemos agora que só foi possível definir a bijeção contínua $f : [0, 2\pi) \rightarrow S^1$ com inversa descontínua porque seu domínio é um intervalo não-compacto.

Uma consequência do Corolário 54 é a seguinte:

Proposição 56. *Seja $\varphi : K \rightarrow L$ uma aplicação contínua do compacto $K \subset \mathbb{R}^m$ sobre o conjunto (necessariamente compacto) $L = \varphi(K) \subset \mathbb{R}^n$. Dado $F \subset L$, se sua imagem inversa $\varphi^{-1}(F)$ é fechada, então F é fechado.*

Demonstração: Temos que φ é sobrejetiva, logo, $\varphi[\varphi^{-1}(F)] = F$. Pelo Corolário 54, F é fechado. ■

Este fato tem como corolário:

Corolário 57. *Seja $\varphi : K \rightarrow L$ uma aplicação contínua do compacto $K \subset \mathbb{R}^m$ sobre o compacto $L \subset \mathbb{R}^n$. Então uma aplicação $f : L \rightarrow \mathbb{R}^p$ é contínua se, e somente se, $f \circ \varphi : K \rightarrow \mathbb{R}^p$ é contínua.*

Demonstração: Se f é contínua, evidentemente $f \circ \varphi$ também é. Reciprocamente, supondo $f \circ \varphi$ contínua, para cada $F \subset \mathbb{R}^p$ fechado temos $\varphi^{-1}[f^{-1}(F)] = (f \circ \varphi)^{-1}(F)$ fechado em K . Pelo que vimos acima, isto implica $f^{-1}(F)$ fechado em L . Logo, f é contínua (Teorema 30).

Como aplicação, seja $g : [0, 2\pi] \rightarrow \mathbb{R}^n$ contínua tal que $g(0) = g(2\pi)$. Através de g , podemos definir uma aplicação contínua $f : S^1 \rightarrow \mathbb{R}^n$, pondo $f(e^{it}) = f(\cos t, \sin t) = g(t)$, $0 \leq t \leq 2\pi$. Como $g(0) = g(2\pi)$, f está bem definida. Seja $\varphi : [0, 2\pi] \rightarrow S^1$ a sobrejeção contínua dada por $\varphi(t) = e^{it} = (\cos t, \sin t)$. Então, $f \circ \varphi = g$ é contínua. Pelo resultado acima, $f : S^1 \rightarrow \mathbb{R}^n$ é contínua. Isto se exprime dizendo que “definir uma aplicação contínua no círculo S^1 é o mesmo que defini-la no intervalo $[0, 2\pi]$, assumindo valores iguais nos extremos”.

Teorema 58. *Se $f : X \rightarrow \mathbb{R}^n$ é contínua e $K \subset X$ é compacto, então, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $y \in K$, $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.*

Demonstração: Suponha, por absurdo, que o teorema fosse falso. Então existiriam $\varepsilon > 0$ e duas seqüências de pontos $x_k \in X$, $y_k \in K$ tais que $|x_k - y_k| < 1/k$ e $|f(x_k) - f(y_k)| \geq \varepsilon$, para todo $k \in \mathbb{N}$. Passando a uma subsequência, se necessário, podemos supor que $\lim y_k = y \in K$, donde $\lim x_k = y$ também. Então, pela continuidade de f , viria $\varepsilon \leq \lim |f(x_k) - f(y_k)| = |f(y) - f(y)|$, uma contradição. ■

Outro tipo de uniformidade ocorre quando uma função (contínua) depende continuamente de um parâmetro α que varia num compacto. Mais precisamente:

Teorema 59. *Seja $f : X \times K \rightarrow \mathbb{R}^n$ contínua, onde K é compacto. Fixemos $x_0 \in X$. Para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $|x - x_0| < \delta \Rightarrow |f(x, \alpha) - f(x_0, \alpha)| < \varepsilon$, seja qual for $\alpha \in K$.*

Demonstração: Supondo o contrário, existiriam $\varepsilon > 0$ e seqüências de pontos $x_k \in X$, $\alpha_k \in K$ tais que $|x_k - x_0| < 1/k$ e $|f(x_k, \alpha_k) - f(x_0, \alpha_k)| \geq \varepsilon$. Passando a uma subsequência, se necessário, podemos admitir que $\lim \alpha_k = \alpha \in K$. Como, evidentemente, $\lim x_k = x_0$, a continuidade de f nos daria $\varepsilon \leq \lim |f(x_k, \alpha_k) - f(x_0, \alpha_k)| = |f(x_0, \alpha) - f(x_0, \alpha)|$, uma contradição. ■

Um exemplo de aplicação do teorema acima é:

Exemplo 60. *Seja $f : X \times [a, b] \rightarrow \mathbb{R}$ contínua. Definamos $\varphi : X \rightarrow \mathbb{R}$ pondo, para cada $x \in X$,*

$$\varphi(x) = \int_a^b f(x, t) dt.$$

Então, φ é contínua em cada ponto $x_0 \in X$. Com efeito, $|\varphi(x) - \varphi(x_0)| \leq \int_a^b |f(x, t) - f(x_0, t)| dt$. Pelo Teorema 59, dado $\varepsilon > 0$, podemos achar $\delta > 0$ tal que $x \in X$, $|x - x_0| < \delta \Rightarrow |f(x, t) - f(x_0, t)| < \varepsilon/(b - a)$, seja qual for $t \in [a, b]$, logo, tem-se $|\varphi(x) - \varphi(x_0)| < \varepsilon$.

Caracterizamos agora a compacidade por meio de coberturas.

Definição 61. *Uma cobertura de um conjunto $X \subset \mathbb{R}^n$ é uma família $(C_\lambda)_{\lambda \in L}$ de subconjuntos $C_\lambda \subset \mathbb{R}^n$ tal que $X \subset \bigcup_{\lambda \in L} C_\lambda$. Isto significa que, para cada $x \in X$, existe um $\lambda \in L$ tal que $x \in C_\lambda$.*

Definição 62. *Uma subcobertura é uma subfamília $(C_\lambda)_{\lambda \in L'}$, $L' \subset L$, tal que ainda se tem $X \subset \bigcup_{\lambda \in L'} C_\lambda$.*

Definição 63. *Diz-se que a cobertura $X \subset \bigcup C_\lambda$ é aberta quando os C_λ forem todos abertos, finita se L é um conjunto finito, enumerável se L é enumerável, etc.*

Teorema 64. (Lindelof). *Seja $X \subset \mathbb{R}^n$ um conjunto arbitrário. Toda cobertura aberta $X \subset \bigcup A_\lambda$ admite uma subcobertura enumerável $X \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_i} \cup \dots$.*

Demonstração: Seja $E = \{x_1, \dots, x_i, \dots\}$ um subconjunto enumerável, denso em X . Consideremos o conjunto \mathcal{B} de todas as bolas abertas $B(x; r)$, com centro num ponto de E , raio racional e tais que cada uma delas está contida em algum A_λ . \mathcal{B} é um conjunto enumerável de bolas. Afirmamos que as bolas $B \in \mathcal{B}$ cobrem X . Com efeito, dado $X \in X$, existe $\lambda \in L$ tal

que $x \in A_\lambda$. Como A_λ é aberto, existe $r > 0$ racional tal que $B(x; 2r) \subset A_\lambda$. Sendo E denso em X , podemos encontrar $x_i \in E$ com $|x - x_i| < r$. Então, $x \in B(x_i, r)$. Para mostrar que $B(x_i, r) \in \mathcal{B}$, resta ver que esta bola está contida em A_λ . Ora, $y \in B(x_i, r) \Rightarrow |y - x_i| < r \Rightarrow |y - x| \leq |y - x_i| + |x_i - x| < 2r \Rightarrow y \in B(x; 2r) \subset A_\lambda$. Isto conclui a verificação de que as bolas $B \in \mathcal{B}$ cobrem X . Tomando uma enumeração B_1, \dots, B_i, \dots para essas bolas e escolhendo, para cada $i \in \mathbb{N}$, um índice $\lambda_i \in L$ tal que $B_i \subset A_{\lambda_i}$, concluímos que $X \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_i} \cup \dots$, o que conclui a prova. ■

Teorema 65. (Borel-Lebesgue). *Seja $K \subset \mathbb{R}^n$ compacto (isto é, limitado e fechado). Toda cobertura aberta $K \subset \bigcup_{\lambda \in L} A_\lambda$ admite uma subcobertura finita $K \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_i}$.*

Demonstração: Pelo Teorema 64, obtemos uma subcobertura enumerável $K \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_i} \cup \dots$. Ponhamos $K_i = K \cap \mathcal{C}(A_{\lambda_1} \cup \dots \cup A_{\lambda_i})$ para cada $i \in \mathbb{N}$. Isto nos dá uma sequência decrescente $K_1 \supset K_2 \supset \dots \supset K_i \supset \dots$ de compactos. Dado qualquer $x \in K$, existe algum $i \in \mathbb{N}$ tal que $x \in A_{\lambda_i}$. Então, $x \notin K_i$. Isto mostra que nenhum ponto $x \in K$ está em todos os K_i , ou seja, que $\bigcap_{i=1}^{\infty} K_i = \emptyset$. Segue-se então da propriedade de Cantor que algum dos compactos K_i é vazio, o que significa $K \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_i}$. ■

Vale também a recíproca do Teorema de Borel-Lebesgue.

Teorema 66. *Se toda cobertura aberta do conjunto $K \subset \mathbb{R}^m$ admite uma subcobertura finita, então K é limitado e fechado (isto é, compacto).*

Demonstração: Em primeiro lugar, as bolas abertas de raio 1 e centros nos pontos de K constituem uma cobertura aberta $K \subset \bigcup_{x \in K} B(x; 1)$, a qual possui uma subcobertura finita $K \subset B(x_1; 1) \cup \dots \cup B(x_i, 1)$. Assim, K está contido numa reunião finita de conjuntos limitados, logo é limitado. Além disso, K é fechado pois, se não fosse, existiria um ponto $a \in \overline{K} - K$. Então, para cada $i \in \mathbb{N}$, tomamos $A_i =$ complementar da bola fechada $B[a; 1/i]$. Para todo $x \in K$, temos $x \neq a$, logo, $|x - a| > 1/i$ para algum i , o que nos dá $x \in A_i$. Portanto, $K \subset \bigcup_{i=1}^{\infty} A_i$, uma cobertura aberta, da qual extraímos uma subcobertura finita: $K \subset A_{i_1} \cup \dots \cup A_{i_p}$. Como $A_1 \subset A_2 \subset \dots \subset A_i \subset \dots$, toda reunião de uma coleção finita de conjuntos A_i é igual ao conjunto de maior índice na coleção. Assim, temos $K \subset A_i$ para algum i . Esta inclusão significa que a bola $B[a; 1/i]$ não tem pontos em comum com K , o que contradiz ser $a \in \overline{K}$ e prova o teorema. ■

Os Teoremas 65 e 66 mostram que poderíamos, equivalentemente, ter definido um conjunto compacto K pela condição de que toda cobertura aberta $K \subset \cup A_\lambda$ admita uma subcobertura finita $K \subset K_{\lambda_1} \cup \dots \cup K_{\lambda_k}$. Tal definição é, de fato, a mais conveniente para estudos mais gerais. A que demos no texto é mais simples, porém interessante apenas para espaços euclidianos.

Como aplicação do Teorema de Borel-Lebesgue, (65), demonstraremos o seguinte:

Proposição 67. *Se o aberto U contém a interseção $\bigcap_{i=1}^{\infty} K_i$ de uma seqüência decrescente $K_1 \supset K_2 \supset \dots \supset K_i \supset \dots$ de conjuntos compactos, então existe $i \in \mathbb{N}$ tal que $K_i \subset U$.*

Demonstração: Temos que os abertos $U_i \in \mathbb{R}^n - K_i$, juntamente com U , constituem uma cobertura de K_1 , da qual extraímos uma subcobertura finita $K_1 \subset U \cup U_{i_1} \cup \dots \cup U_{i_p}$. Seja i o maior dos índices i_1, \dots, i_p . Como $U_1 \subset U_2 \subset \dots$, temos $U_{i_1} \cup \dots \cup U_{i_p} = U_i$, logo $K_1 \subset U \cup U_i$. Com maior razão, $K_i \subset U \cup U_i$. Como nenhum ponto de K_i pode pertencer ao seu complementar U_i devemos ter $K_i \subset U$, como queríamos demonstrar. ■

Os conjuntos compactos estendem nossa compreensão de continuidade e de limites, garantindo que certas propriedades se mantenham dentro de um contexto controlado. Esse conceito estabelece uma base para a análise de funções ao longo de trajetórias específicas, que serão exploradas na primeira seção do próximo capítulo.

Capítulo 3

Caminhos Diferenciáveis e Funções Reais de n variáveis

Neste capítulo, são apresentadas as noções dos Caminhos Diferenciáveis e abordadas as propriedades de Derivadas Parciais e Direcionais, além do conceito de Funções Diferenciáveis. O capítulo é concluído com uma introdução ao Gradiente de uma função diferenciável e apresentando o Teorema da Função Implícita.

3.1 Caminhos diferenciáveis

Os Caminhos Diferenciáveis representam trajetórias suaves que permitem observar a variação de funções ao longo de direções específicas. Nesta seção, abordamos a definição e as propriedades desses caminhos e exploramos como eles oferecem uma base para entender como funções se comportam ao longo de trajetórias diferenciáveis em espaços multidimensionais.

Definição 68. *Um caminho em \mathbb{R}^n é uma aplicação $f : I \rightarrow \mathbb{R}^n$, cujo domínio é um intervalo $I \subset \mathbb{R}$. Para cada $t \in I$, temos $f(t) = (f_1(t), \dots, f_n(t))$. As n funções $f_i : I \rightarrow \mathbb{R}$ são chamadas as funções coordenadas de f . Escreve-se, então, $f = (f_1, \dots, f_n)$.*

O caminho $f : I \rightarrow \mathbb{R}^n$ é contínuo no ponto $a \in I$ se, e somente se, cada uma das suas funções coordenadas é contínua nesse ponto. Mais geralmente, sabemos que se $f = (f_1, \dots, f_n)$ é definida num conjunto $X \subset \mathbb{R}$ e a é um ponto de acumulação de X , então $\lim_{t \rightarrow a} f(t) = b = (b_1, \dots, b_n)$ se, e somente se, $\lim_{t \rightarrow a} f_i(t) = b_i$ para cada $i = 1, 2, \dots, n$. Como $X \subset \mathbb{R}$, têm sentido os limites laterais $\lim_{t \rightarrow a^-} f(t)$ e $\lim_{t \rightarrow a^+} f(t)$, o primeiro quando a é ponto de acumulação à esquerda e o segundo quando a é ponto de acumulação à direita.

Definição 69. O vetor velocidade do caminho $f : I \rightarrow \mathbb{R}^n$ no ponto $a \in I$ é, por definição, o limite

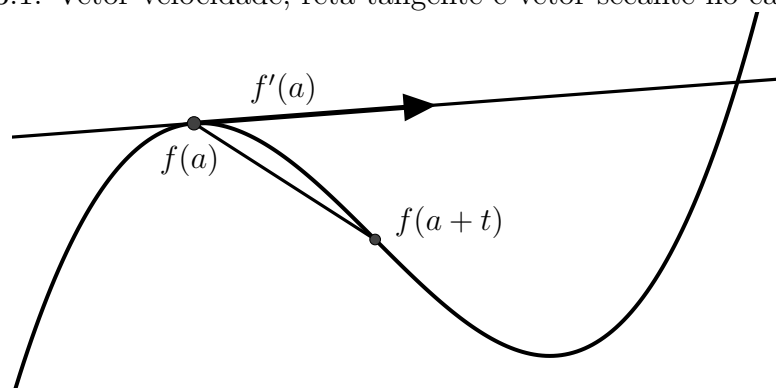
$$f'(a) = \lim_{t \rightarrow 0} \frac{f(a+t) - f(a)}{t},$$

quando tal limite existe. Às vezes escrevemos $\frac{df}{dt}(a)$ ou $Df(a)$ em vez de $f'(a)$. A norma $|f'(a)|$ chama-se a velocidade escalar de f no ponto a .

Definição 70. Quando o caminho f possui vetor velocidade num ponto, dizemos que f é diferenciável nesse ponto. Se existe $f'(a)$ para todo $a \in I$, dizemos que $f : I \rightarrow \mathbb{R}^n$ é um caminho diferenciável.

O vetor velocidade $f'(a)$, quando é diferente de zero, determina a *reta tangente* ao caminho f no ponto $f(a)$, a saber, a reta $L = \{f(a) + t \cdot f'(a); t \in \mathbb{R}\}$.

Figura 3.1: Vetor velocidade, reta tangente e vetor secante no caminho f .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Como as coordenadas do “vetor secante” $[f(a+t) - f(a)]/t$ são os números $[f_i(a+t) - f_i(a)]/t$, vemos que o caminho f possui vetor tangente no ponto a se, e somente se, cada uma das suas funções coordenadas f_i possui derivada nesse ponto. No caso afirmativo, tem-se $f'(a) = (f'_1(a), \dots, f'_n(a))$. Assim, o caminho $f : I \rightarrow \mathbb{R}^n$ é diferenciável se, e somente se, cada uma das suas funções coordenadas $f_i : I \rightarrow \mathbb{R}$ é uma função real diferenciável no intervalo I .

Em particular, a diferenciabilidade do caminho f não depende da norma que estivermos utilizando em \mathbb{R}^n .

A fim de que o caminho $f : I \rightarrow \mathbb{R}^n$ seja diferenciável no ponto $a \in I$, é necessário e suficiente que exista um vetor $v \in \mathbb{R}^n$ tal que, para $a+t \in I$ se tenha

$$f(a+t) = f(a) + t \cdot v + r(t), \quad \text{onde} \quad \lim_{t \rightarrow 0} \frac{r(t)}{t} = 0.$$

No caso afirmativo, tem-se $v = f'(a)$.

Com efeito, a primeira igualdade acima significa que, para $t \neq 0$,

$$\frac{r(t)}{t} = \frac{f(a+t) - f(a)}{t} - v.$$

A condição de diferenciabilidade de f no ponto a pode ser expressa, equivalentemente, assim:

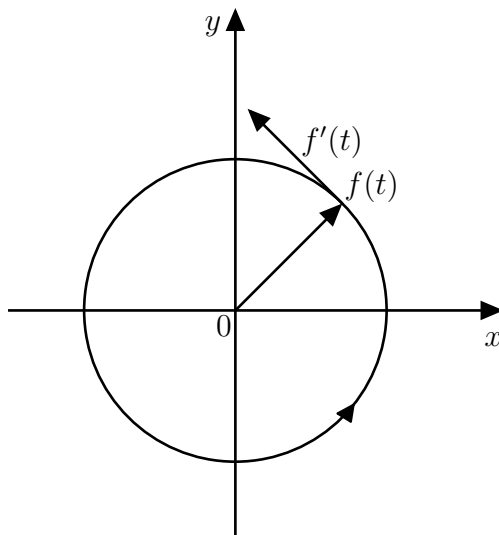
$$f(a+t) = f(a) + t[f'(a) + \rho(t)], \quad \text{onde } \lim_{t \rightarrow 0} \rho(t) = 0.$$

Basta por $\rho(t) = r(t)/t$, se $t \neq 0$, e $\rho(0) = 0$.

A noção de derivada lateral se define, para caminhos, de modo análogo ao caso de funções reais. Tem-se $f'(a+) = (f'_1(a+), \dots, f'_n(a+))$ e $f'(a-) = (f'_1(a-), \dots, f'_n(a-))$. Existe o vetor velocidade de f no ponto a se, e somente se, existem e são iguais as derivadas laterais nesse ponto.

Exemplo 71. *Seja $f : \mathbb{R} \rightarrow \mathbb{R}^2$ dado por $f(t) = (\cos t, \sin t) = e^{it}$. A imagem do caminho f é o círculo unitário S^1 . Para todo $t \in \mathbb{R}$, o vetor velocidade de f é $f'(t) = (-\sin t, \cos t) = ie^{it}$. A velocidade escalar é constante: $|f'(t)| = 1$.*

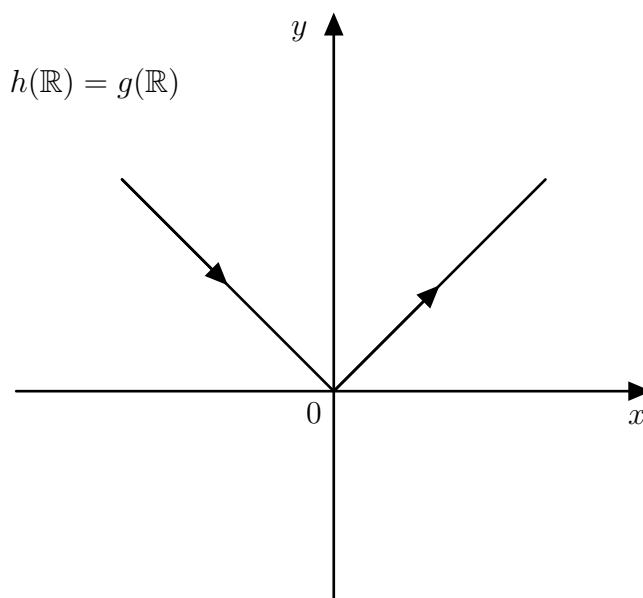
Figura 3.2: Círculo unitário S^1 com $f(t)$ e o vetor velocidade $f'(t)$ do caminho f .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Exemplo 72. O caminho $g : \mathbb{R} \rightarrow \mathbb{R}^2$, definido por $g(t) = (t, |t|)$, possui vetor velocidade $g'(t) = (1, 1)$ para todo $t > 0$ e $g'(t) = (1, -1)$ para $t < 0$. No ponto $t = 0$, as derivadas laterais $g'(0-) = (1, -1)$ e $g'(0+) = (1, 1)$ existem mas são diferentes. Logo, g não possui velocidade no ponto 0 . A imagem de g é o gráfico da função $y = |x|$, que apresenta um ponto angular na origem. Podemos descrever a mesma imagem por meio de outra “parametrização”, considerando, por exemplo, o caminho $h : \mathbb{R} \rightarrow \mathbb{R}^2$, com $h(t) = (t^3, |t|^2)$. Temos $h(\mathbb{R}) = g(\mathbb{R})$. Para $t < 0$, $h'(t) = (3t^2, -3t^2)$ e, para $t > 0$, $h'(t) = (3t^2, 3t^2)$. No ponto $t = 0$, as derivadas laterais são ambas nulas, logo, existe $h'(0) = (0, 0)$. Em outras palavras: para descrever a rota $h(\mathbb{R})$, o ponto cuja posição no tempo t é $h(t)$ precisou dar uma parada instantânea ao atingir o ponto angular $(0, 0)$ de sua trajetória.

Figura 3.3: Trajetória da curva $h(\mathbb{R}) = g(\mathbb{R})$ com ponto angular em $(0, 0)$.



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

As propriedades usuais da derivada de uma função real de uma variável real, quando aplicadas às funções coordenadas de um caminho, conduzem imediatamente às seguintes regras de derivação:

$$\begin{aligned}\frac{d}{dt}[f(t) + g(t)] &= f'(t) + g'(t), \\ \frac{d}{dt}[\alpha(t) \cdot f(t)] &= \alpha'(t)f(t) + \alpha(t)f'(t), \\ \frac{d}{dt}\langle f(t), g(t) \rangle &= \langle f'(t), g(t) \rangle + \langle f(t), g'(t) \rangle, \\ \frac{d}{dt}|f(t)| &= \frac{\langle f(t), f'(t) \rangle}{|f(t)|} \text{ se } f(t) \neq 0.\end{aligned}$$

Acima, $f, g : I \rightarrow \mathbb{R}^n$ são caminhos diferenciáveis e $\alpha : I \rightarrow \mathbb{R}$ é uma função diferenciável. A norma $|f(t)|$ é a que provém do produto interno $\langle f(t), g(t) \rangle$, isto é, $|f(t)| = \sqrt{\langle f(t), f(t) \rangle}$. Esta hipótese será feita sempre que tomarmos a derivada de $|f(t)|$.

Exemplo 73. *Se a norma não provém de um produto interno, podemos perfeitamente ter um caminho diferenciável $f : I \rightarrow \mathbb{R}^n$, com $f(t) \neq 0$ para todo $t \in I$, para o qual a função $\varphi(t) = |f(t)|$ não é diferenciável. Por exemplo, consideremos a norma do máximo em \mathbb{R}^2 . O caminho $f : \mathbb{R} \rightarrow \mathbb{R}^2$, dado por $f(t) = (1, t)$ é obviamente diferenciável. Mas $|f(t)|$ não possui derivada em cada um dos pontos $t = -1$ e $t = 1$.*

Resulta da fórmula da derivada de $|f(t)|$ que, dado um caminho diferenciável $f : I \rightarrow \mathbb{R}^n$, o vetor $f(t)$ tem comprimento constante se, e somente se, para cada valor do parâmetro t , o vetor velocidade $f'(t)$ é perpendicular ao “vetor posição” $f(t)$.

Exemplo 74. *O caminho $f : \mathbb{R} \rightarrow \mathbb{R}^2$, dado por $f(t) = (\cos t, \sin t)$, é tal que $|f(t)| = 1$ para todo t . Por isso, $f'(t) = (-\sin t, \cos t)$ é, para todo $t \in \mathbb{R}$, perpendicular a $f(t)$. Neste exemplo, temos também $|f'(t)| = 1$, mas isto é acidental. Se tomarmos $g : \mathbb{R} \rightarrow \mathbb{R}^2$, com $g(t) = (\cos t^2, \sin t^2)$, ainda vale $|g(t)| = 1$ para todo t , mas $g'(t) = (-2t \sin t^2, 2t \cos t^2)$ tem comprimento variável, igual a $2|t|$. Continua, porém, sendo $g'(t)$ perpendicular a $g(t)$.*

Se o caminho $f : I \rightarrow \mathbb{R}^n$ é diferenciável, tem sentido considerar o caminho derivado $f' : I \rightarrow \mathbb{R}^n$ e indagar se ele é contínuo, diferenciável, etc.

Observação 75. *Quando f' é contínuo, diz-se que f é um caminho de classe \mathcal{C}^1 . Pode-se então investigar a existência da derivada de f' . Quando existe, o vetor $(f')'(a) = f''(a)$ chama-se a derivada segunda de f no ponto a , ou o vetor aceleração do caminho f no ponto a . Tem-se $f''(a) = (f''_1(a), \dots, f''_n(a))$. Se existe $f''(t)$ para todo $t \in I$, diz-se que f é duas vezes diferenciável e fica definido o caminho $f'' : I \rightarrow \mathbb{R}^n$. Quando f'' é contínuo, dizemos que f é um caminho de classe \mathcal{C}^2 .*

Definição 76. Diremos que o caminho $f : I \rightarrow \mathbb{R}^n$ é $p + 1$ vezes diferenciável quando existir o caminho $f^{(p)} : I \rightarrow \mathbb{R}^n$ (derivada de ordem p de f) e for derivável. Então, poremos $f^{(p+1)} = (f^{(p)})'$. Quando $f^{(p)}$ for de classe \mathcal{C}^1 , diremos que f é de classe \mathcal{C}^{p+1} .

Definição 77. Diremos que o caminho $f : I \rightarrow \mathbb{R}^n$ é p vezes diferenciável no ponto $a \in I$ quando existir $\delta > 0$ tal que f é de classe \mathcal{C}^{p-1} no intervalo $J = \{t \in I; |t - a| < \delta\}$ e $f^{(p-1)}$ for diferenciável no ponto a .

Por extensão, diremos que um caminho contínuo é de classe \mathcal{C}^0 e que $f = f^{(0)}$ é sua própria “derivada de ordem zero”.

Definição 78. Quando existirem as derivadas de todas as ordens do caminho f , diremos que f é infinitamente derivável, ou de classe \mathcal{C}^∞ .

Para $0 \leq p \leq \infty$, escrevemos $f \in \mathcal{C}^p$ para indicar que f é um caminho de classe \mathcal{C}^p . Dado $f = (f_1, \dots, f_n)$, tem-se $f \in \mathcal{C}^p$ se, e somente se, $f_i \in \mathcal{C}^p$ para cada $i = 1, \dots, n$.

Definição 79. Seja $p > 0$. Diremos que o caminho $f : I \rightarrow \mathbb{R}^n$ é de classe \mathcal{C}^p por partes quando f for contínuo e, além disso, possuir derivadas contínuas até a ordem p , inclusive, salvo num conjunto finito de pontos de I . Nesses pontos, f deve possuir derivadas laterais contínuas até a ordem p , inclusive.

Exemplo 80. Para todo $p \geq 0$, o caminho $f : \mathbb{R} \rightarrow \mathbb{R}^2$, definido por $f(t) = (t^{p+1}, t^p|t|)$, é de classe \mathcal{C}^p , e de classe \mathcal{C}^∞ , por partes.

O estudo de caminhos diferenciáveis fornece um recurso para analisar variações direcionais de forma controlada e suave. Essa compreensão é essencial para o desenvolvimento das ideias de derivadas parciais e direcionais, temas das próximas seções, que expandem ainda mais nossa análise de funções em várias variáveis.

3.2 Derivadas parciais

As derivadas parciais capturam a taxa de variação de uma função em relação a cada uma de suas variáveis, proporcionando uma análise detalhada do comportamento local das funções. Nesta seção, exploramos o cálculo e as propriedades das derivadas parciais, fundamentais para descrever como as funções variam em um ambiente multidimensional.

Quando se estudam funções reais de n variáveis, isto é, definidas em subconjuntos do espaço \mathbb{R}^n , e se busca para essas funções uma noção de derivada que tenha propriedades análogas às da derivada de uma função definida num intervalo, a idéia que se apresenta mais naturalmente é a de “derivada parcial”, que exporemos agora.

Para efeito de derivação, onde se compara o acréscimo $f(a + h) - f(a)$ da função 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 suficiente pequeno h .

Definição 81. *Seja, pois, $f : U \rightarrow \mathbb{R}$ 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 tal limite existe.

Às vezes, usaremos também a notação $\partial f(a)$.

Observação 82. *O símbolo $\frac{\partial f}{\partial x_i}$ terá para nós o mesmo significado que $\frac{\partial f}{\partial y_i}$, $\frac{\partial f}{\partial z_i}$, etc. O que importa num símbolo destes não é o “nome” da variável, que tanto pode ser x , como y ou z , etc. O importante é o índice i : trata-se da derivada de f em relação à sua i -ésima variável, seja qual for o sinal usado para indicá-la. Estritamente falando, a melhor notação para a i -ésima derivada parcial é $\partial_i f$, mas continuaremos escrevendo $\frac{\partial f}{\partial x_i}$ por respeito à tradição, pelo apelo estético e, principalmente, porque isto torna mais naturais certas fórmulas, como por exemplo a Regra da Cadeia.*

Quando $U \subset \mathbb{R}^2$, uma função $f : U \rightarrow \mathbb{R}$ é o que se chama uma “função real de duas variáveis reais”. Escreve-se $f(x, y)$ para indicar seu valor no ponto $z = (x, y)$. Desta forma, as derivadas parciais de f num ponto $c = (a, b) \in U$ podem também ser representadas por $\frac{\partial f}{\partial x}(c)$ e $\frac{\partial f}{\partial y}(c)$, em vez de $\frac{\partial f}{\partial x_1}(c)$, $\frac{\partial f}{\partial x_2}(c)$. Temos:

$$\frac{\partial f}{\partial x}(c) = \lim_{t \rightarrow 0} \frac{f(a + t, b) - f(a, b)}{t}, \quad \frac{\partial f}{\partial y}(c) = \lim_{t \rightarrow 0} \frac{f(a, b + t) - f(a, b)}{t}.$$

Analogamente, se $U \subset \mathbb{R}^3$, uma função $f : U \rightarrow \mathbb{R}$ é uma “função real de três variáveis reais”. Seu valor num ponto $p = (x, y, z)$ se escreve $f(x, y, z)$ e suas derivadas parciais no ponto $q = (a, b, c)$ podem ser escritas como $\frac{\partial f}{\partial x}(q)$, $\frac{\partial f}{\partial y}(q)$ e $\frac{\partial f}{\partial z}(q)$.

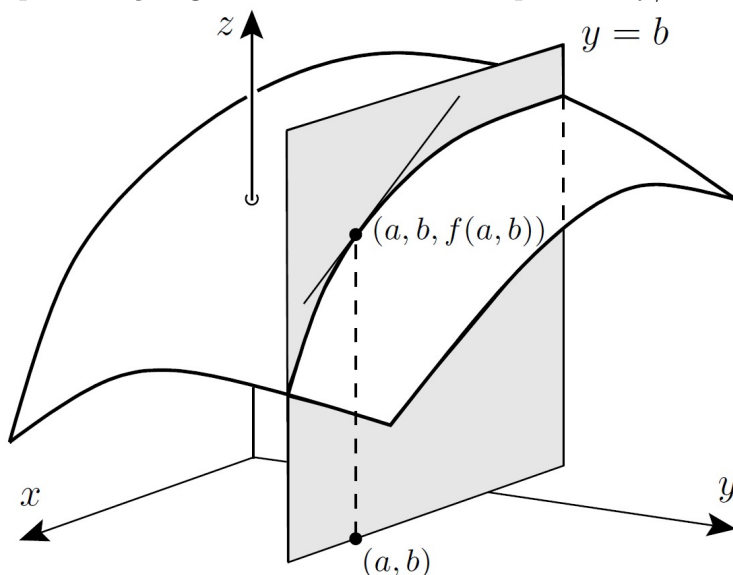
Voltando ao caso geral, seja $f : U \rightarrow \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$. Dados o ponto $a \in U$ e o inteiro $i \in [1, n]$, a imagem do caminho $\lambda : \mathbb{R} \rightarrow \mathbb{R}^n$, $\lambda(t) = a + te_i$, é o que se chama “a reta que passa por a e é paralela ao i -ésimo eixo”. (Note que $\lambda(0) = a$.) 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 $t = 0$.

Quando $n = 2$, o gráfico de f é uma superfície em \mathbb{R}^3 ; a restrição de f ao segmento de reta que passa por $c = (a, b)$ e é paralelo ao eixo das abscissas tem como gráfico a curva plana obtida nessa superfície fazendo y constante, igual a b . Logo, $\frac{\partial f}{\partial x}(c)$ é a inclinação da reta tangente a essa curva, no ponto $(a, b, f(a, b))$, relativamente ao plano horizontal.

O cálculo prático da i -ésima derivada parcial de uma função $f(x_1, \dots, x_n)$ se faz considerando todas as variáveis como se fossem constantes, exceto a i -ésima, e aplicando as regras usuais de derivação relativamente a essa variável.

Figura 3.4: Representação geométrica da derivada parcial $\partial f / \partial x_i$ no plano $y = b$.



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

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. Assim, por exemplo, se $f : U \rightarrow \mathbb{R}$ está definida em $U \subset \mathbb{R}^2$, o segmento de reta $J = \{(a, t); 0 \leq t \leq 1\}$, paralelo ao eixo dos y , está contido em U e, além disso, $\frac{\partial f}{\partial y}(z) > 0$

para todo $z \in J$, então f é crescente sobre J , isto é, $0 \leq s < t \leq 1 \Rightarrow f(a, s) > f(a, t)$.

Definição 83. 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 $b = (a_1, \dots, a_{i-1}, y, a_{i+1}, \dots, a_n)$ em U , tem-se $f(a) = f(b)$. Noutros termos, se $a, b \in U$ com $a + te_i$ então $f(a) = f(b)$.

Definição 84. Um conjunto $U \subset \mathbb{R}^n$ diz-se i -convexo quando $a, b \in U$, $b = a + te_i \Rightarrow [a, b] \subset U$.

A Proposição abaixo resulta das definições:

Proposição 85. 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.

Demonstração: 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 $f(a) = \xi(0) = \xi(t) = f(b)$. ■

Observação 86. No plano, diz-se horizontalmente, ou verticalmente, convexo, em vez de 1-convexo e 2-convexo, respectivamente.

Exemplo 87. Seja $X = \{(x, 0) \in \mathbb{R}^2; x \geq 0\}$ o semi-eixo positivo fechado das abscissas. O aberto $U = \mathbb{R}^2 \rightarrow X$ 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 $f(x, -y) = 0$.

Note-se que as derivadas parciais sozinhas não permitem conclusões sobre o comportamento “ n -dimensional” da função. Por exemplo, a existência de todas as derivadas parciais num ponto não implica a continuidade da função nesse ponto, como veremos agora.

Exemplo 88. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = xy/(x^2 + y^2)$ se $x^2 + y^2 \neq 0$ e $f(0, 0) = 0$. Se $z = (x, y)$ não é a origem, temos $\partial_1 f(z) = (y^3 - x^2y)/(x^2 + y^2)^2$ e $\partial_2 f(z) = (x^3 - xy^2)/(x^2 + y^2)^2$. Na origem, vale:

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

Assim, f possui derivadas parciais em todos os pontos do plano. Entretanto, f é descontínua na origem.

[Se $x^2 + y^2 \neq 0$, então $f(x, y) = (x/\sqrt{x^2 + y^2})(y/\sqrt{x^2 + y^2}) = \cos \theta \cdot \sin \theta$, onde θ é o ângulo formado pelo semi-eixo positivo das abscissas e a semi-reta que passa na origem e contém o ponto (x, y) . Ao longo de cada uma dessas semi-retas, $f(x, y)$ tem valor constante, dependendo da semi-reta, logo não existe $\lim f(x, y)$ na origem.]

As derivadas parciais são ferramentas cruciais para entender variações locais ao longo de cada direção independente, permitindo uma análise precisa de funções multivariáveis. Esse conceito prepara o terreno para o estudo das derivadas direcionais, que generalizam essa ideia para direções arbitrárias.

3.3 Derivadas direcionais

As derivadas direcionais estendem o conceito de derivadas parciais, permitindo a análise da taxa de variação de uma função em qualquer direção desejada. Nesta seção, examinamos como essas derivadas fornecem uma visão mais abrangente das mudanças que ocorrem em funções multivariáveis, independente dos eixos coordenados.

Vendo que as derivadas parciais, desacompanhadas de hipóteses adicionais, apenas fornecem informações sobre a função ao longo de retas paralelas aos eixos, tentamos estender a noção de derivada a outras direções além dessas. Isto nos leva ao importante conceito de derivada direcional.

Definição 89. *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 tal limite existe.

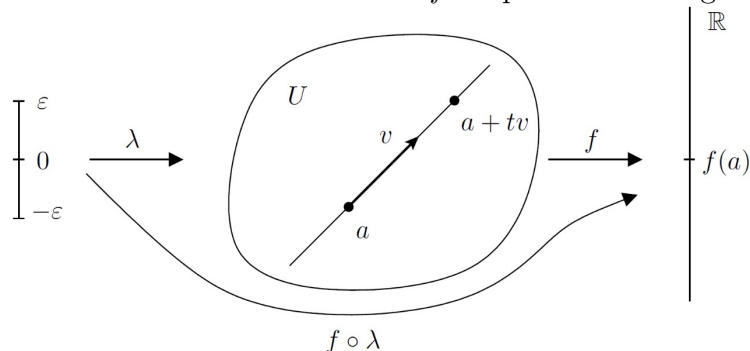
As derivadas parciais tornam-se casos particulares das derivadas direcionais:

$$\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i}(a) = \text{derivada direcional de } f \text{ segundo o vetor } e_i.$$

A derivada direcional $(\partial f / \partial v)(a)$ é a derivada, no ponto $t = 0$, da função composta $f \circ \lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$, onde $\lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ é o caminho retilíneo, $\lambda(t) = a + tv$, para o qual se tem $\lambda(0) = a$ e $\lambda'(t) = v$ para todo t .

Aqui, $\varepsilon > 0$ é escolhido tão pequeno que a imagem de λ esteja contida em U .

Figura 3.5: Derivada direcional de f no ponto a ao longo de v .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

A função do Exemplo 88 possui as derivadas direcionais $(\partial f/\partial v)(0,0)$ para $v = (\alpha, 0)$ ou $(0, \beta)$, as quais são nulas. Se, porém, tomarmos $v = (\alpha, \beta)$, com $\alpha \neq 0$ e $\beta \neq 0$, veremos que não existe $(\partial f/\partial v)(0,0)$, pois

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

e o último limite acima, evidentemente, não existe.

Admitimos que $v \in \mathbb{R}^n$ seja um vetor arbitrário porque desejamos que $\partial f/\partial v$ dependa linearmente de v . Vejamos se isto realmente ocorre. Em primeiro lugar, se $0 \neq \alpha \in \mathbb{R}$, então existe $\partial f/\partial(\alpha v)$ num ponto $a \in U$ se, e somente se, existe $\partial f/\partial v$ e, no caso afirmativo, temos

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

Por outro lado, o exemplo abaixo mostra que a derivada direcional $\partial f/\partial v$ pode existir em todos os pontos do domínio de f , segundo todos os vetores $v \in \mathbb{R}^n$, sem que se tenha necessariamente

$$\frac{\partial f}{\partial(v+w)}(a) = \frac{\partial f}{\partial v}(a) + \frac{\partial f}{\partial w}(a).$$

Exemplo 90. Seja $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $g(0,0) = 0$ e $g(x,y) = x^2y/(x^2 + y^2)$ quando $x^2 + y^2 \neq 0$. Um cálculo direto, a partir da definição, mostra que existem as derivadas direcionais $\partial g/\partial v$ em todos os pontos de \mathbb{R}^2 , segundo qualquer vetor $c = (\alpha, \beta)$. Em particular, na origem,

temos

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

Evidentemente, para $a = (0,0)$ não vale $\frac{\partial g}{\partial(v+w)}(a) = \frac{\partial g}{\partial v}(a) + \frac{\partial g}{\partial w}(a)$.

Na próxima seção, mostraremos que $\partial f/\partial v$ dependerá linearmente de v se f for “diferenciável”, uma hipótese mais restritiva do que simplesmente possuir derivadas direcionais.

A função g do Exemplo 90 é contínua em todos os pontos do plano. [Isto é claro em $\mathbb{R}^2 - \{0\}$. Na origem, basta observar que $g(x, y) = x \cdot \cos \theta \cdot \sen \theta$. (Veja o Exemplo 88). Logo, $\lim_{x,y \rightarrow 0} g(x, y) = 0$.]

Não é verdade, porém, que a existência das derivadas direcionais implique em continuidade, como veremos agora.

Exemplo 91. Seja $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $h(0,0) = 0$ e $h(x, y) = x^3y/(x^6 + y^2)$ se $(x, y) \neq (0,0)$. Em $\mathbb{R}^2 - \{0\}$, a função h é contínua. Como, porém, $h(x, x^3) = 1/2$ para todo $x \neq 0$, vemos que h é descontínua na origem. Examinemos as derivadas direcionais. Para todo $v = (\alpha, \beta)$, temos:

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

Assim, todas as derivadas direcionais $\partial h/\partial v$ na origem existem e dependem linearmente de v . O mesmo ocorre nos demais pontos $c \in \mathbb{R}^2 - \{0\}$, como se verifica mediante o cálculo elementar de $\frac{\partial h}{\partial v}(c) = \xi'(0)$, onde $\xi(t) = h(c + tv)$.

Outra propriedade desejável para um conceito adequado de derivada de uma função de n variáveis é que a composta de duas funções deriváveis seja ainda derivável. O próximo exemplo mostra uma função contínua $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, tal que $(\partial\varphi)(z)$ existe para todo z e todo v , depende linearmente de v , mas $\varphi \circ \lambda$ não é diferenciável para um certo caminho diferenciável λ .

Exemplo 92. Definamos $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ pondo $\varphi(x, y) = x^3y/(x^4 + y^2)$ se $(x, y) \neq (0,0)$ e $\varphi(0,0) = 0$. Fora da origem, temos

$$\varphi(x, y) = x \frac{x^2}{\sqrt{x^4 + y^2}} \cdot \frac{y}{\sqrt{x^4 + y^2}} = x \frac{1}{\sqrt{1 + y^2/x^4}} \cdot \frac{\pm 1}{\sqrt{x^4/y^2 + 1}},$$

logo, $\lim_{x,y \rightarrow 0} \varphi(x,y) = 0$. Assim, φ é contínua. Além disso, para todo $v = (\alpha, \beta)$, temos

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

Portanto existem na origem, e dependem linearmente de v , todas as derivadas direcionais $\partial\varphi/\partial v$. Nos demais pontos $c \in \mathbb{R}^2 - \{0\}$ chega-se à mesma conclusão calculando-se $\frac{\partial \varphi}{\partial v}(c) = \xi'(0)$, com $\xi(t) = \varphi(c + tv)$, por meio das regras elementares de derivação. Entretanto, se considerarmos o caminho $\lambda : \mathbb{R} \rightarrow \mathbb{R}^2$, definido por $\lambda(t) = (t, t^2 \text{sen } 1/t)$, $\lambda(0) = (0, 0)$, veremos que λ é diferenciável e não existe a derivada $(\varphi \circ \lambda)'(0)$. Com efeito, seria

$$(\varphi \circ \lambda)'(0) = \lim_{t \rightarrow 0} \frac{\varphi(t, t^2 \text{sen } 1/t)}{t} = \lim_{t \rightarrow 0} \frac{\text{sen } 1/t}{1 + \text{sen}^2 1/t},$$

mas o último limite não existe.

A hipótese da existência de derivadas direcionais, embora fraca, não é inteiramente inócua. Ela permite demonstrar o teorema do Valor Médio, o qual, para funções reais de n variáveis, é verdadeiro sob forma de igualdade, como no caso de uma só variável.

Teorema 93. (Teorema do Valor Médio) *Seja $f : U \rightarrow \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$. Suponhamos que o segmento de reta $[a, a + v]$ esteja contido em U , que a restrição $f|_{[a, a + v]}$ seja contínua e que exista a derivada direcional $\frac{\partial f}{\partial v}(x)$, segundo v , em todo ponto $x \in (a, a + v)$.*

Então existe $\theta \in (0, 1)$ tal que $f(a + v) - f(a) = \frac{\partial f}{\partial v}(a + \theta v)$.

Demonstração: Definamos a função $\xi : [0, 1] \rightarrow \mathbb{R}$ pondo $\xi(t) = f(a + tv)$. Pelas hipóteses feitas sobre f , ξ é contínua em $[0, 1]$ e derivável em $(0, 1)$. Pelo Teorema do Valor Médio para funções de uma variável real, existe $\theta \in (0, 1)$ tal que $\xi(1) - \xi(0) = \xi'(\theta)$. Mas $\xi(1) = f(a + v)$, $\xi(0) = f(a)$ e

$$\begin{aligned} \xi'(\theta) &= \lim_{t \rightarrow 0} \frac{\xi(\theta + t) - \xi(\theta)}{t} = \lim_{t \rightarrow 0} \frac{f(a + (\theta + t)v) - f(a + \theta v)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(a + \theta v + tv) - f(a + \theta v)}{t} = \frac{\partial f}{\partial v}(a + \theta v), \end{aligned}$$

o que demonstra o teorema. ■

Observação: A existência de $\partial f/\partial v$ em todo ponto de $(a, a + v)$ assegura apenas a continuidade de $f|_{(a, a + v)}$ mas não de $f|_{[a, a + v]}$.

Corolário 94. *Seja $U \subset \mathbb{R}^n$ aberto e conexo. Se $f : U \rightarrow \mathbb{R}$ possui derivadas direcionais em todo ponto $x \in U$ e $\frac{\partial f}{\partial v}(x) = 0$ para qualquer vetor v , então f é constante.*

Com efeito, fixemos $a \in U$. A existência de $\partial f / \partial v$ garante a continuidade da restrição $f|_{[a,b]}$ para todo segmento de reta $[a,b]$ contido em U . Resulta então do Teorema do Valor Médio que $[a,b] \subset U$ implica $f(b) = f(a)$. Ora, qualquer ponto $x \in U$ pode (em virtude da conexidade do aberto U) ser ligado ao ponto a por uma poligonal contida em U , com vértices $a_0 = a, a_1, \dots, a_k = x$. Temos sucessivamente $f(a) = f(a_1) = \dots = f(x)$. Logo, $f(x) = f(a)$ para todo $x \in U$, donde f é constante.

Com as derivadas direcionais, obtemos uma descrição completa de como funções variam em todas as direções possíveis, enriquecendo a análise e compreensão das funções diferenciáveis. Esse conceito é essencial para avançarmos na diferenciabilidade e entender a estrutura global de uma função.

3.4 Funções diferenciáveis

Funções diferenciáveis permitem aproximações lineares precisas, revelando como elas se comportam em torno de pontos específicos. Nesta seção, discutimos o conceito de diferenciabilidade e como ele facilita a compreensão das variações locais e suaves em funções de múltiplas variáveis.

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

Definição 95. *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 constantes A_1, \dots, A_n tais que, para todo vetor $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, com $a + v \in U$, se tenha*

$$f(a + v) = f(a) + A_1 \cdot \alpha_1 + \dots + A_n \cdot \alpha_n + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

Observação 96. *Quando f é diferenciável em todos os pontos de U , dizemos simplesmente que f é diferenciável.*

Se f é diferenciável no ponto a , então, tomando $v = te_i$, temos $\alpha_j = 0$ se $j \neq i$ e $\alpha_i = t$. Logo,

$$\frac{f(a + te_i) - f(a)}{t} = A_i + \frac{r(te_i)}{t} = A_i \pm \frac{r(te_i)}{|te_i|}.$$

Fazendo $t \rightarrow 0$, vemos que existe cada derivada parcial de f no ponto a , sendo $\frac{\partial f}{\partial x_i}(a) = A_i$. A definição abaixo é, portanto, equivalente à anterior.

Definição 97. Diremos que a função $f : U \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$ 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

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

Observação 98. Na igualdade acima, o “resto” $r(v)$ é definido como sendo igual a $f(a + v) - f(a) - \Sigma(\partial f / \partial x_i)(a) \cdot \alpha_i$. 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 (direta ou indiretamente) sempre que quisermos provar que uma função é diferenciável.

De $\lim_{v \rightarrow 0} (r(v)/|v|) = 0$, concluímos que $\lim_{v \rightarrow 0} r(v) = 0$ pois $r(v) = (r(v)/|v|)|v|$. Daí resulta que toda função diferenciável num ponto é contínua nesse ponto. Com efeito, para $v = (\alpha_1, \dots, \alpha_n)$, temos

$$\lim_{v \rightarrow 0} [f(a + v) - f(a)] = \lim_{v \rightarrow 0} \left[\sum \frac{\partial f}{\partial x_i}(a) \alpha_i + r(v) \right] = 0.$$

A condição $\lim_{v \rightarrow 0} (r(v)/|v|) = 0$ significa, entretanto, mais do que $r(v) \rightarrow 0$; ela quer dizer que $r(v)$ tende a zero mais rapidamente do que v , isto é, para valores de v suficientemente próximos de zero, a norma de $r(v)$ é uma fração arbitrariamente pequena da norma de v . Às vezes, isto se exprime dizendo-se que $r(v)$ é um infinitésimo de ordem superior a v . Assim, f é diferenciável no ponto a quando o acréscimo

$$f(a + v) - f(a)$$

é igual a uma função linear de v , $\sum \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i$, mais um resto infinitamente pequeno em relação a v .

Note que a validade da afirmação $\lim_{v \rightarrow 0} (r(v)/|v|) = 0$ independe da norma adotada em \mathbb{R}^n .

Em certas ocasiões, é preferível usar, em vez de $r(v)$, a função $\rho = \rho(v)$, definida para os valores de v tais que $a + v \in U$, do seguinte modo: $\rho(v) = r(v)/|v|$ se $v \neq 0$ e $\rho(0) = 0$. Então, a função 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

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

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

Para funções $f : I \rightarrow \mathbb{R}$, definidas num intervalo aberto $I \subset \mathbb{R}$, diferenciabilidade é o mesmo que derivabilidade, pois de $f(a + t) = f(a) + A \cdot t + \rho|t|$ se tira

$$\rho = \pm \left| \frac{f(a + t) - f(a)}{t} - A \right|,$$

logo $\lim_{t \rightarrow 0} \rho = 0$ se, e somente se, $A = f'(a)$.

Seja $f : U \rightarrow \mathbb{R}$ diferenciável no ponto $a \in U$. Já vimos que f é contínua e possui derivadas parciais nesse ponto. Mostraremos agora que f admite derivada direcional segundo qualquer vetor $v = (\alpha_1, \dots, \alpha_n)$, e vale a fórmula

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

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:

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

donde

$$\frac{f(a + tv) - f(a)}{t} = \sum \frac{\partial f}{\partial x_i} \cdot \alpha_i \pm \rho(tv) \cdot |v|.$$

Como $\lim_{t \rightarrow 0} \rho(tv) = 0$, segue-se a fórmula enunciada.

Resulta da expressão $\frac{\partial f}{\partial v} = \sum \left(\frac{\partial f}{\partial x_i} \alpha_i \right)$ que se f é diferenciável num ponto então a derivada

direcional $\frac{\partial f}{\partial v}$, nesse ponto, depende linearmente de v , isto é, não somente se tem $\frac{\partial f}{\partial(\alpha v)} = \alpha \left(\frac{\partial f}{\partial v} \right)$ como também $\frac{\partial f}{\partial(v+w)} = \frac{\partial f}{\partial v} + \frac{\partial f}{\partial w}$.

Uma propriedade relevante das funções diferenciáveis é dada pela

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 $v = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ tais que $a + v \in U$. Para $v \in U_0$ e $k = 1, \dots, n$, temos

$$f_k(a + v) = f_k(a) + \sum_{i=1}^m \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + \rho_k \cdot |v|, \quad (3.1)$$

onde cada $\rho_k = \rho_k(v)$ é uma função definida em U_0 , contínua no ponto 0, que se anula quando $v = 0$. [Acima, e no que se segue, as derivadas parciais $\partial f_k / \partial x_i$ e $\partial / \partial y_k$ são consideradas nos pontos a e b , respectivamente.]

Consideremos a aplicação $w = (\beta_1, \dots, \beta_n) : U_0 \rightarrow \mathbb{R}^n$, contínua no ponto 0, cujas funções-coordenada são definidas por

$$\beta_k(v) = \sum_{i=1}^m \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + \rho_k \cdot |v|. \quad (3.2)$$

Adotando, por exemplo, a norma da soma, temos $|\alpha_i|/|v| \leq 1$ se $v \neq 0$, logo cada $|\beta_k|/|v|$, e portanto a função $|w|/|v|$, é limitada numa vizinhança do ponto $v = 0$. Escrevendo gf em vez de $g \circ f$, podemos afirmar, em virtude de (3.1), (3.2) e da diferenciabilidade de g no ponto $b = f(a)$, que, para todo $v \in U_0$, vale:

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

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

$$gf(a+v) = gf(a) + \sum_{k=1}^n \frac{\partial g}{\partial y_k} \left(\sum_{i=1}^m \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + \rho_k |v| \right) + \sigma \cdot |w| = gf(a) + \sum_{i=1}^m A_i \cdot \alpha_i + R,$$

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

Daí

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

Quando v tende a zero, sabemos que cada função ρ_k tende a zero, que o quociente $|w|/|v|$ é limitado e que $\lim_{v \rightarrow 0} \sigma = 0$. Segue-se que $\lim_{v \rightarrow 0} (R/|v|) = 0$. Isto mostra que $g \circ f$ é diferenciável no ponto a e suas derivadas parciais são os números A_i .

Observação 99. *A notação clássica do Cálculo Diferencial, às vezes imprecisa porém bastante sugestiva, além de compatível com a prática (então universal) de enfatizar grandezas (“ y é uma função de x ”) em vez de aplicações (“ f leva x em y ”), seria a seguinte, para a Regra da Cadeia: os pontos de U seriam escritos como “ x ” e os de V como “ y ”; as funções f_k seriam escritas como $y_k = [y_k(x)]$. A derivada $\partial(g \circ f)/\partial x_i$ seria a “derivada de g em relação à variável x_i ”, indicada com $\partial g/\partial x_i$. A Regra da Cadeia seria então:*

$$\frac{\partial g}{\partial x_i} = \sum_k \frac{\partial g}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_i}.$$

Não se pode negar a elegância nostálgica desta fórmula. Ela, porém, se acha demasiadamente comprometida com sistemas de coordenadas, para o gosto atual.

Corolário 100. *Se $f : U \rightarrow \mathbb{R}$ é diferenciável no ponto b e se $\lambda : (a - \varepsilon, a + \varepsilon) \rightarrow U \subset \mathbb{R}^n$ é um caminho diferenciável no ponto a , com $\lambda(a) = b$ e $\lambda(t) = (\lambda_1(t), \dots, \lambda_n(t))$, então a função composta $f \circ \lambda : (a - \varepsilon, a + \varepsilon) \rightarrow \mathbb{R}$ é diferenciável no ponto a , e tem-se*

$$(f \circ \lambda)'(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(b) \cdot \lambda'_i(a).$$

Se escrevermos $\lambda(t) = (x_1(t), \dots, x_n(t))$, então

$$\lambda'(t) = \left(\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right).$$

Indicando com $\frac{df}{dt}$ a derivada da função composta

$$t \mapsto f(\lambda(t)) = f(x_1(t), \dots, x_n(t)),$$

a Regra da Cadeia assume a forma clássica

$$\frac{df}{dt} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \frac{dx_i}{dt}.$$

Corolário 101. *Sejam $U \subset \mathbb{R}^n$ um conjunto aberto, $f : U \rightarrow \mathbb{R}$ diferenciável no ponto a , com $f(U) \subset I$, e $g : I \rightarrow \mathbb{R}$ diferenciável no ponto $b = f(a)$. Então $g \circ f : U \rightarrow \mathbb{R}$ é diferenciável no ponto a e, para cada $i = 1, \dots, n$, vale*

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = g'(b) \cdot \frac{\partial f}{\partial x_i}(a).$$

Decorre da Regra da Cadeia que, se $f : U \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$, ao calcularmos a derivada direcional $\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$, não é necessário tomar $\lambda(t) = a + tv$. Em vez de nos restringirmos a um caminho retilíneo, podemos considerar *qualquer* caminho $\lambda : (-\epsilon, \epsilon) \rightarrow U$, diferenciável no ponto a , com $\lambda(0) = a$ e $\lambda'(0) = v = (\alpha_1, \dots, \alpha_n)$ e teremos ainda

$$\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0) = \lim_{t \rightarrow 0} \frac{f(\lambda(t)) - f(a)}{t}.$$

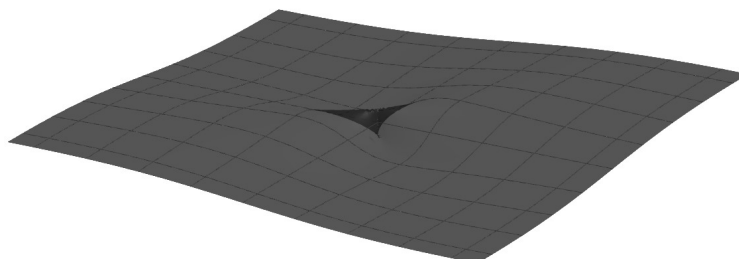
Com efeito, pela Regra da Cadeia,

$$(f \circ \lambda)'(0) = \sum \frac{\partial f}{\partial x_i}(a) \cdot \lambda'_i(0) = \sum \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i = \frac{\partial f}{\partial v}(a).$$

Revejamos agora os Exemplos 88, 90, 91 e 92 à luz da definição de diferenciabilidade. As funções que examinamos foram as seguintes: (citar os exemplos nos ambientes)

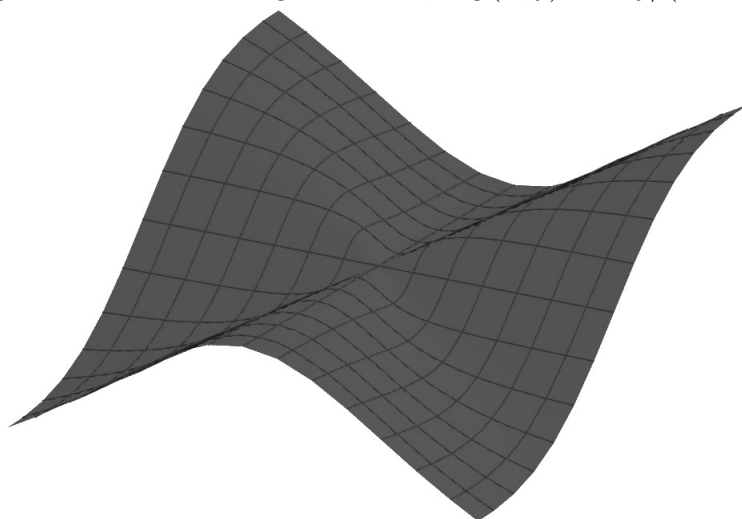
$$\begin{aligned}
f : \mathbb{R}^2 &\rightarrow \mathbb{R}, & f(x, y) &= xy/(x^2 + y^2), & f(0, 0) &= 0; \\
g : \mathbb{R}^2 &\rightarrow \mathbb{R}, & g(x, y) &= x^2y/(x^2 + y^2), & g(0, 0) &= 0; \\
h : \mathbb{R}^2 &\rightarrow \mathbb{R}, & h(x, y) &= x^3y/(x^6 + y^2), & h(0, 0) &= 0; \\
\varphi : \mathbb{R}^2 &\rightarrow \mathbb{R}, & \varphi(x, y) &= x^3y/(x^4 + y^2), & \varphi(0, 0) &= 0.
\end{aligned}$$

Figura 3.6: Gráfico de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = xy/(x^2 + y^2)$.



Fonte: Criado pelos autores com o auxílio do *software* GeoGebra.

Figura 3.7: Gráfico de $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g(x, y) = x^2y/(x^2 + y^2)$.



Fonte: Criado pelos autores com o auxílio do *software* GeoGebra.

Nenhuma dessas funções é diferenciável na origem de \mathbb{R}^2 : f porque não é contínua nem possui derivada direcional segundo todo vetor; g porque, embora sendo contínua e existindo

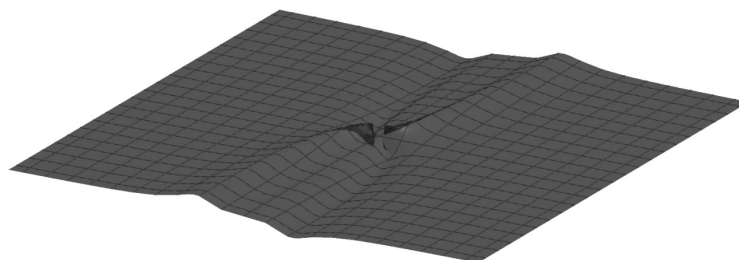
$\partial g/\partial v(0,0)$ segundo todo vetor $v \in \mathbb{R}^2$, esta derivada não depende linearmente de v ; h possui derivadas direcionais $\partial h/\partial v$, que dependem linearmente de v , mas não é contínua na origem; finalmente, φ é contínua em todo o plano, admite em todos os pontos do plano derivadas direcionais $\partial \varphi/\partial v$ que dependem linearmente de v , mas contraria a Regra da Cadeia porque, considerando o caminho $\lambda(t) = (t, t^2 \text{sen } 1/t)$, a composta $\varphi \circ \lambda : \mathbb{R} \rightarrow \mathbb{R}$ não é derivável no ponto $t = 0$.

Estas são razões indiretas pelas quais as quatro funções acima não são diferenciáveis. A razão real é que, embora cada uma delas possua derivadas parciais na origem, elas não cumprem a condição

$$\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = \lim_{\alpha, \beta \rightarrow 0} \frac{1}{\sqrt{\alpha^2 + \beta^2}} \left(f(\alpha, \beta) - \frac{\partial f}{\partial x} \cdot \alpha - \frac{\partial f}{\partial y} \cdot \beta \right) = 0,$$

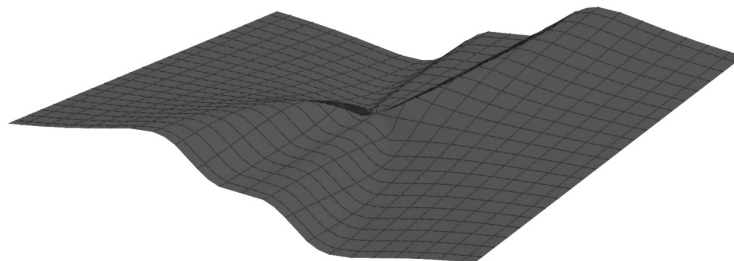
onde as derivadas parciais são tomadas no ponto $(0,0)$. Acima, $\sqrt{\alpha^2 + \beta^2} = |v|$ é a norma euclidiana do vetor $v = (\alpha, \beta)$.

Figura 3.8: Gráfico de $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(x, y) = x^3y/(x^6 + y^2)$.



Fonte: Criado pelos autores com o auxílio do *software* GeoGebra.

Figura 3.9: Gráfico de $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\varphi(x, y) = x^3y/(x^4 + y^2)$.



Fonte: Criado pelos autores com o auxílio do *software* GeoGebra.

Definição 102. Uma função real $f : U \rightarrow \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, diz-se de classe \mathcal{C}^1 quando existem, em cada ponto $x \in U$, as derivadas parciais $\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)$, e as n funções $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$, assim definidas, são contínuas. Mais geralmente, diremos que uma função $f : U \rightarrow \mathbb{R}$ é de classe \mathcal{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 \mathcal{C}^{k-1} . Aqui, k é um inteiro > 0 .

Para completar a definição indutiva, diremos que uma função $f : U \rightarrow \mathbb{R}$ é de classe \mathcal{C}^0 quando ela for contínua. Usaremos a notação $f \in \mathcal{C}^k$.

Escreveremos também $f \in \mathcal{C}^\infty$, e diremos que f é de classe \mathcal{C}^∞ quando $f \in \mathcal{C}^k$ para todo $k \geq 0$.

Evidentemente, $\mathcal{C}^0 \supset \mathcal{C}^1 \supset \dots \supset \mathcal{C}^k \supset \dots \supset \mathcal{C}^\infty$, sendo todas estas inclusões estritas.

Teorema 103. Se uma função $f : U \rightarrow \mathbb{R}$ possui derivadas parciais em todos os pontos do aberto $U \subset \mathbb{R}^n$ e cada uma delas é contínua no ponto c , então f é diferenciável no ponto c .

Demonstração: Por simplicidade, consideraremos o caso $n = 2$. A situação geral se trata de modo análogo, apenas com notação mais complicada. Fixemos $c = (a, b) \in U$ e tomemos $v = (h, k)$ tal que $c + v \in U$. Seja

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

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

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

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

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

Logo,

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

Ora, $\frac{h}{\sqrt{h^2 + k^2}}$ e $\frac{k}{\sqrt{h^2 + k^2}}$ estão, em valor absoluto, compreendidos entre 0 e 1. A conti-

nidade das derivadas parciais nos dá, então,

$$\lim_{v \rightarrow 0} \left[\frac{r(v)}{|v|} \right] = 0,$$

logo, f é diferenciável. ■

Corolário 104. *Toda função de classe C^1 é diferenciável.*

Observação 105. *Escrevendo*

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

vemos que existe $\theta \in (0, 1)$ tal que

$$\frac{r(v)}{|v|} = \left[\frac{\partial f}{\partial x}(a+\theta h, b+k) - \frac{\partial f}{\partial x}(a, b) \right] \frac{h}{|v|} + \left[\frac{f(a, b+k) - f(a, b)}{k} - \frac{\partial f}{\partial y}(a, b) \right] \frac{k}{|v|}.$$

Como $\frac{h}{|v|}$ e $\frac{k}{|v|}$ têm valor absoluto menor do que um, a primeira parcela da soma acima tem limite zero quando $|v| \rightarrow 0$, desde que $\frac{\partial f}{\partial x}$ seja contínua no ponto (a, b) . A segunda parcela também tem limite zero, em virtude da definição de derivada, sem que seja preciso supor $\frac{\partial f}{\partial y}$ contínua. Assim, para que uma função f , de duas variáveis, seja diferenciável num ponto, é suficiente que $\frac{\partial f}{\partial x}$ exista numa vizinhança do ponto e seja contínua no ponto, e que $\frac{\partial f}{\partial y}$ apenas exista no ponto em questão. Para funções de n variáveis, a diferenciabilidade num ponto é assegurada quando $n-1$ das suas derivadas parciais são contínuas no ponto e a derivada parcial restante apenas existe ali.

Exemplo 106. *Um polinômio em duas variáveis é uma função $p: \mathbb{R}^2 \rightarrow \mathbb{R}$, do tipo $p(x, y) = \sum a_{ij}x^i y^j$. Todo polinômio é evidentemente uma função contínua e possui derivadas parciais $\frac{\partial p}{\partial x} = \sum i a_{ij}x^{i-1}y^j$, $\frac{\partial p}{\partial y} = \sum j a_{ij}x^i y^{j-1}$. Tais derivadas são ainda polinômios e, portanto, são funções contínuas em \mathbb{R}^2 . Logo, todo polinômio $p: \mathbb{R}^2 \rightarrow \mathbb{R}$ é uma função de classe C^1 . Então, as derivadas de p , sendo polinômios, são de classe C^1 , portanto $p \in C^2$. Repetindo o mesmo argumento, concluímos que $p \in C^k$ para todo k , logo todo polinômio é, na realidade, uma função de classe C^∞ . Afirmativas semelhantes podem ser feitas sobre um polinômio a n variáveis, que é uma função $p: \mathbb{R}^n \rightarrow \mathbb{R}$, do tipo $p(x) = \sum a_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n}$.*

A soma $f + g$, o produto $f \cdot g$ e o quociente f/g (se $g(x) \neq 0$ para todo x no domínio de g) de funções $f, g : U \rightarrow \mathbb{R}$, de classe \mathcal{C}^k , são ainda funções de classe \mathcal{C}^k .

Exemplo 107. Como exemplo de função diferenciável que não é de classe \mathcal{C}^1 , tomemos $f : \mathbb{R} \rightarrow \mathbb{R}$, com $f(x) = x^2 \cdot \text{sen}(1/x)$ se $x \neq 0$, e $f(0) = 0$, lembrando que, para funções de uma variável, diferenciabilidade é o mesmo que existência da derivada.

Exemplo 108. O produto interno $f : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$, $f(x, y) = \sum x_i y_i$, sendo um polinômio em $2m$ variáveis, é uma função de classe \mathcal{C}^∞ . Também é de classe \mathcal{C}^∞ a função $g : \mathbb{R}^n \rightarrow \mathbb{R}$, $g(x) = |x|^2 = \sum x_i^2$, por ser ainda um polinômio, em n variáveis. Pela Regra da Cadeia, a norma euclidiana $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $h(x) = |x| = \sqrt{\sum x_i^2}$, é de classe \mathcal{C}^∞ quando restrita a $\mathbb{R}^n \setminus \{0\}$. Na origem, a norma euclidiana não é diferenciável; nem sequer existem as derivadas parciais:

$$\frac{\partial h}{\partial x_i}(0^+) = 1 \quad e \quad \frac{\partial h}{\partial x_i}(0^-) = -1.$$

(Derivadas parciais laterais diferentes.) Quanto às normas que não provêm de um produto interno, elas podem não ser diferenciáveis mesmo em pontos $x \neq 0$. Por exemplo, seja $\xi : \mathbb{R}^2 \rightarrow \mathbb{R}$ a norma da soma, $\xi(x, y) = |x| + |y|$. Nos pontos $(x, 0)$ não existe $\frac{\partial \xi}{\partial y}$ e nos pontos $(0, y)$ não existe $\frac{\partial \xi}{\partial x}$.

Para concluir, registremos um importante corolário do Teorema 103 (e da Regra da Cadeia), segundo o qual $g \circ f \in \mathcal{C}^k$ desde que $g \in \mathcal{C}^k$ e cada função coordenada de f também seja de classe \mathcal{C}^k .

Corolário 109. 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 $f_j : U \rightarrow \mathbb{R}$ é de classe \mathcal{C}^k . Seja ainda $g : V \rightarrow \mathbb{R}$ uma função de classe \mathcal{C}^k . Então a função composta $g \circ f : U \rightarrow \mathbb{R}$ é de classe \mathcal{C}^k .

Com efeito, pelo Corolário 104, g e cada f_j são diferenciáveis. (Estamos supondo $k \geq 1$, pois o Corolário 109 é trivial se $k = 0$. Podemos então aplicar a Regra da Cadeia, segundo a qual, para todo $i = 1, \dots, m$, e todo $x \in U$:

$$\frac{\partial(g \circ f)}{\partial x_i}(x) = \sum_{j=1}^n \frac{\partial g}{\partial y_j}(f(x)) \cdot \frac{\partial f_j}{\partial x_i}(x),$$

ou seja, vale a igualdade de funções

$$\frac{\partial(g \circ f)}{\partial x_i}(x) = \sum_{j=1}^n \left(\frac{\partial g}{\partial y_j} \circ f \right) \cdot \frac{\partial f_j}{\partial x_i}.$$

Suponhamos, por indução, que o Corolário 109 foi provado para classe \mathcal{C}^{k-1} . Então, para cada $j = 1, \dots, n$, a função composta $\frac{\partial g}{\partial y_j} \circ f$ é de classe \mathcal{C}^{k-1} , o mesmo ocorrendo com $\frac{\partial f_j}{\partial x_i}$ para todo i , já que $f \in \mathcal{C}^k$. Como o produto de funções de classe \mathcal{C}^{k-1} é ainda desta classe, cada parcela da soma acima é de classe \mathcal{C}^{k-1} , donde a soma também é. Assim, todas as derivadas parciais de $g \circ f$ são de classe \mathcal{C}^{k-1} , portanto $g \circ f \in \mathcal{C}^k$.

A diferenciabilidade fornece uma maneira estruturada de entender as variações locais das funções, por meio de aproximações lineares. Esse conceito é um pilar para a definição da diferencial e do gradiente, que serão explorados nas próximas seções.

3.5 A diferencial de uma função

A diferencial de uma função representa a melhor aproximação linear da variação da função em torno de um ponto. Nesta seção, analisamos o conceito de diferencial e suas propriedades, que desempenham um papel crucial na descrição da variação local de funções diferenciáveis.

A derivada de um caminho $f : \mathbb{R} \rightarrow \mathbb{R}^m$ é um vetor. Na situação dual, o papel de derivada de uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é desempenhado por um funcional linear, conforme mostraremos agora.

Definição 110. *Seja $f : U \rightarrow \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$, diferenciável no ponto $a \in U$. A diferencial de f no ponto a é o funcional linear $df(a) : \mathbb{R}^n \rightarrow \mathbb{R}$, cujo valor no vetor $v = (\alpha_1, \dots, \alpha_n)$ é dado por*

$$df(a) \cdot v = \frac{\partial f}{\partial v}(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i.$$

Observação 111. *Como toda transformação linear $\mathbb{R}^n \rightarrow \mathbb{R}$, o funcional linear $df(a)$ possui uma matriz $1 \times n$ em relação à base canônica de \mathbb{R}^n . Se identificarmos o funcional com sua matriz, teremos*

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

Quando $f : U \rightarrow \mathbb{R}$ é diferenciável em todo ponto de U , obtemos uma aplicação $df : U \rightarrow (\mathbb{R}^n)^* = \mathcal{L}(\mathbb{R}^n; \mathbb{R})$, que associa a cada ponto $x \in U$ o funcional $df(x)$, cuja matriz é

$$\left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

A aplicação df é contínua se, e somente se, cada uma das suas funções coordenadas $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ é contínua, isto é, se, e somente se, f é de classe \mathcal{C}^1 .

É comum indicar-se, em Análise, a base canônica de $(\mathbb{R}^n)^*$, com (dx_1, \dots, dx_n) , logo $dx_i \cdot v = \alpha_i$ se $v = (\alpha_1, \dots, \alpha_n)$. O motivo desta notação é o seguinte: como a i -ésima projeção $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ assume, em cada ponto $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ o valor $\pi_i(x) = x_i$, escreve-se x_i em vez de π_i . Calculando, de modo óbvio, a diferencial da i -ésima projeção $x_i : \mathbb{R}^n \rightarrow \mathbb{R}$, obtemos $dx_i(a) \cdot v = \alpha_i$ em todo ponto $a \in \mathbb{R}^n$. Escrevendo $dx_i \cdot v$ em vez de α_i , a definição da diferencial fica

$$df(a) \cdot v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot dx_i \cdot v.$$

Como esta igualdade vale para cada $v \in \mathbb{R}^n$, temos

$$df(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i.$$

Isto significa que o funcional linear $df(a)$ se exprime como combinação linear dos funcionais dx_i , sendo $(\partial f / \partial x_i)(a)$ os coeficientes da combinação. Finalmente, a igualdade acima valendo para todo ponto $a \in U$, podemos escrever

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

A expressão formal da regra da cadeia (no caso $\mathbb{R} \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$) diz que, se cada coordenada x_i é função de um parâmetro real t , então podemos “dividir ambos os membros da igualdade acima por dt ” e obter

$$\frac{df}{dt} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \frac{dx_i}{dt}.$$

Todo funcional linear $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável e, para todo $x \in \mathbb{R}^n$, $d\varphi(x) = \varphi$ (isto é, $d\varphi(x) \cdot v = \varphi \cdot v$). Com efeito, temos $\varphi(x) = c_1 x_1 + \dots + c_n x_n$, logo $\frac{\partial \varphi}{\partial x_i} = c_i$. Portanto,

$$d\varphi(x) \cdot v = \sum \frac{\partial \varphi}{\partial x_i} \cdot \alpha_i = \sum c_i \alpha_i = \varphi \cdot v.$$

Teorema 112. *Sejam $f, g : U \rightarrow \mathbb{R}$ diferenciáveis no ponto $a \in U$. Então:*

1. $f + g : U \rightarrow \mathbb{R}$ é diferenciável e $d(f + g) = df + dg$;

2. $f \cdot g : U \rightarrow \mathbb{R}$ é diferenciável e $d(f \cdot g) = f \cdot dg + g \cdot df$;

3. Se $g(x) \neq 0$ para todo $x \in U$, então $f/g : U \rightarrow \mathbb{R}$ é diferenciável e $d(f/g) = (g \cdot df - f \cdot dg)/g^2$.

Demonstração: As funções $s, m : \mathbb{R}^2 \rightarrow \mathbb{R}$, $q : \mathbb{R} \times (\mathbb{R} - \{0\}) \rightarrow \mathbb{R}$, definidas por $s(x, y) = x + y$, $m(x, y) = x \cdot y$ e $q(x, y) = x/y$, são de classe C^∞ , logo diferenciáveis. A aplicação $F : U \rightarrow \mathbb{R}^2$, definida por $F(x) = (f(x), g(x))$, tem coordenadas diferenciáveis. Como $f + g = s \circ F$, $f \cdot g = m \circ F$, e $f/g = q \circ F$, a Regra da Cadeia assegura a diferenciabilidade de $f + g$, $f \cdot g$ e f/g . Além disso, temos

$$\frac{\partial}{\partial x_i}(f + g) = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i},$$

$$\frac{\partial}{\partial x_i}(f \cdot g) = f \cdot \frac{\partial g}{\partial x_i} + g \cdot \frac{\partial f}{\partial x_i},$$

e

$$\frac{\partial}{\partial x_i} \left(\frac{f}{g} \right) = \frac{g \cdot \partial f / \partial x_i - f \cdot \partial g / \partial x_i}{g^2}.$$

Daí resultam as fórmulas enunciadas para as diferenciais. ■

O Teorema do Valor Médio (93), já demonstrado para funções que admitem derivadas direcionais ao longo de um segmento, assume, para funções diferenciáveis, a forma abaixo, que decorre da anterior (bem como o Corolário 114):

Teorema 113. (Teorema do Valor Médio) *Seja $f : U \rightarrow \mathbb{R}$ diferenciável em todos os pontos do segmento de reta aberto $(a, a + v)$ e seja contínua sua restrição ao segmento fechado $[a, a + v] \subset U \subset \mathbb{R}^n$. Existe $\theta \in (0, 1)$ tal que*

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

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

Corolário 114. *Seja $U \subset \mathbb{R}^n$ aberto e conexo. Se $f : U \rightarrow \mathbb{R}$ é diferenciável e $df(x) = 0$ (isto é, $\frac{\partial f}{\partial x_1}(x) = \dots = \frac{\partial f}{\partial x_n}(x) = 0$) para todo $x \in U$, então f é constante.*

Corolário 115. *Sejam $U \subset \mathbb{R}^n$ um aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função diferenciável. Se $|df(x)| \leq M$ para todo $x \in U$, então, para quaisquer $x, y \in U$, temos $|f(x) - f(y)| \leq M \cdot |x - y|$.*

Ou seja, num aberto convexo, toda função que possui diferencial limitada é Lipschitziana. No Corolário 115, $|df(x)|$ é a norma do funcional $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, isto é, o maior dos números

$\left| \frac{\partial f}{\partial v}(x) \right|$, para todo $v \in \mathbb{R}^n$, $|v| = 1$. Se tomarmos em \mathbb{R}^n a norma euclidiana, ou a norma da soma, ou a norma do máximo, então $|df(x)|$ assume, respectivamente, os valores

$$\sqrt{\sum_i \left(\frac{\partial f}{\partial x_i}(x) \right)^2}, \quad \max_i \left| \frac{\partial f}{\partial x_i}(x) \right| \quad \text{ou} \quad \sum_i \left| \frac{\partial f}{\partial x_i}(x) \right|.$$

Exemplo 116. Quando V não é convexo, uma função $g : V \rightarrow \mathbb{R}$ pode ter diferencial limitada em V e não ser Lipschitziana. Por exemplo, sejam f e U como no Exemplo 87 e tomemos $V = \{z \in U; |z| < 2\}$, $g = f|_V$. Então $|dg(z)| \leq 4$ para todo $z \in V$, mas g não é Lipschitziana pois, para todo $\varepsilon > 0$, temos $|g(1, \varepsilon) - g(1, -\varepsilon)| = 1$ enquanto $|(1, \varepsilon) - (1, -\varepsilon)| = 2\varepsilon$.

Uma consequência do Corolário 115 acima é que se $f : U \rightarrow \mathbb{R}$ é diferenciável e suas derivadas parciais são limitadas no aberto convexo $U \subset \mathbb{R}^n$, então f é uniformemente contínua em U . Em particular, f é a restrição de uma função (uniformemente) contínua $g : \bar{U} \rightarrow \mathbb{R}$.

A diferencial sintetiza a variação local de uma função de maneira linear, facilitando o estudo detalhado de suas mudanças próximas a cada ponto. Esse conceito se conecta diretamente ao gradiente, que oferece uma interpretação vetorial dessas variações.

3.6 O gradiente de uma função diferenciável

O gradiente de uma função diferenciável aponta na direção de maior crescimento da função e representa a intensidade dessa variação. Nesta seção, exploramos as propriedades e a interpretação do gradiente, que é essencial para a análise e otimização de funções.

O produto interno natural induz um isomorfismo entre \mathbb{R}^n e seu dual $(\mathbb{R}^n)^*$. Tal isomorfismo faz corresponder a cada vetor $v \in \mathbb{R}^n$ o funcional $v^* \in (\mathbb{R}^n)^*$ com $v^*(x) = \langle v, x \rangle$ para todo $x \in \mathbb{R}^n$. Se $v = (c_1, \dots, c_n)$, então $v^*(e_1) = \alpha_1, \dots, v^*(e_n) = \alpha_n$, logo a matriz de v^* em relação à base canônica de \mathbb{R}^n é $(\alpha_1, \dots, \alpha_n)$. A existência deste isomorfismo é responsável pelo fato de que no Cálculo Vetorial clássico (e, portanto, na Geometria e na Física tradicionais) não ocorrem funcionais lineares: em vez de um funcional, toma-se o produto interno $\langle v, x \rangle$ de um vetor fixo v por um vetor variável x . A própria expressão $\varphi(x) = c_1x_1 + \dots + c_nx_n$, que dá o valor do funcional φ no vetor $x = (x_1, \dots, x_n)$, já indica isso: $\varphi(x)$ é o produto interno de x pelo vetor $v = (c_1, \dots, c_n)$, ou seja, $\varphi = v^*$.

Definição 117. Dada a função diferenciável $f : U \rightarrow \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, definiremos o gradiente de f no ponto $a \in U$ como o vetor $\nabla f(a)$, que corresponde ao funcional $df(a)$

segundo o isomorfismo acima descrito. Isto significa, por definição, que:

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

para todo $v = (c_1, \dots, c_n)$.

Em particular, $\langle \nabla f(a), e_i \rangle = \frac{\partial f}{\partial x_i}(a)$, logo

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

Observação 118. Se usarmos apenas bases ortonormais em \mathbb{R}^n , as coordenadas do vetor $\nabla f(a)$ em relação à base (e_1, \dots, e_n) são as mesmas que as do funcional $df(a)$ com respeito à base dual (dx_1, \dots, dx_n) .

Nessas condições, o gradiente se torna praticamente indistinguível da diferencial. Mesmo no espaço euclidiano, o gradiente, por ser um vetor, apresenta aspectos geométricos muito convenientes para dar informações a respeito do comportamento da função, como veremos a seguir. Destacaremos as três propriedades mais importantes do gradiente de uma função diferenciável f . Nesta discussão, fixaremos um ponto a e suporemos que $\nabla f(a) \neq 0$. Então:

- 1^a) O gradiente aponta para uma direção segundo a qual a função f é crescente;
- 2^a) Dentre todas as direções ao longo das quais a função f cresce, a direção do gradiente é a de crescimento mais rápido;
- 3^a) O gradiente de f no ponto a é perpendicular à superfície de nível de f que passa por esse ponto.

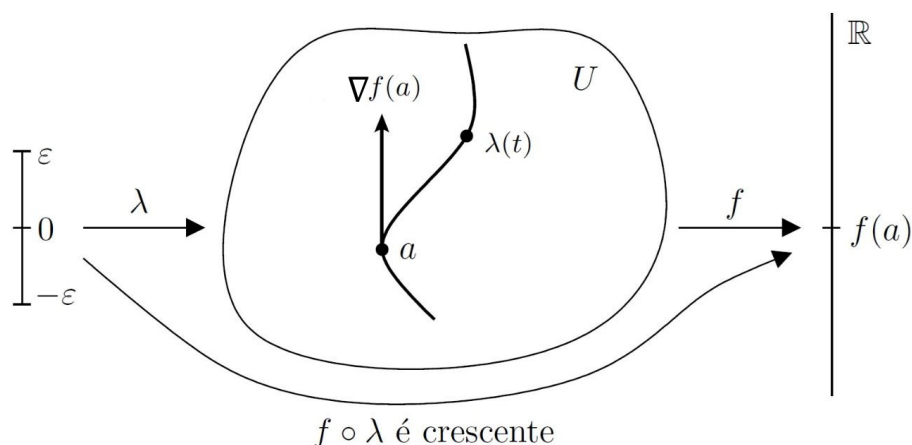
Em primeiro lugar, se $w = \nabla f(a)$ então

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

Isto significa que se $\lambda : (-\epsilon, \epsilon) \rightarrow U$ é um caminho diferenciável, com valores no domínio U da função f , tal que $\lambda(0) = a$ e $\lambda'(0) = \nabla f(a)$, então a função real $t \mapsto f(\lambda(t))$ possui derivada positiva no ponto $t = 0$. Se supusermos f e λ de classe \mathcal{C}^1 , então a derivada de $f \circ \lambda$ será ainda positiva em todos os pontos de um intervalo aberto de centro 0, isto é, se tomarmos $\epsilon > 0$

suficientemente pequeno, então $f \circ \lambda : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ será uma função crescente. Isto é o que significa afirmar que “ f cresce na direção do gradiente”. Evidentemente, não se tem $\frac{\partial f}{\partial v}(a) > 0$ apenas quando $v = \nabla f(a)$. Como $\frac{\partial f}{\partial v}(a) = \langle \nabla f(a), v \rangle$, os vetores v que apontam para direções ao longo das quais a função cresce são aqueles que formam um ângulo agudo com $\nabla f(a)$, isto é, tais que o produto interno $\langle \nabla f(a), v \rangle$ é positivo. O que distingue o gradiente é o fato de que em sua direção o crescimento de f é mais rápido do que nas outras.

Figura 3.10: Gradiente $\nabla f(a)$ indicando a direção de maior crescimento de f .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Isto quer dizer o seguinte: se v for um vetor tal que $|v| = |\nabla f(a)|$ então

$$\frac{\partial f}{\partial v}(a) \leq \frac{\partial f}{\partial(\nabla f(a))}(a).$$

Com efeito, pela desigualdade de Schwarz:

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

Finalmente, esclareçamos a terceira das afirmações acima.

Definição 119. Dada $f : U \rightarrow \mathbb{R}$, diferenciável no aberto, $U \subset \mathbb{R}^n$, e dado um número real c , diz-se que o ponto $x \in U$ está no nível c , ou tem nível c , relativamente a f , quando $f(x) = c$. Fixado c , o conjunto dos pontos de U que estão no nível c é a imagem inversa $f^{-1}(c)$, a qual é chamada a superfície de nível c da função f . Quando $n = 2$, $f^{-1}(c)$ chama-se a curva de nível c de f .

Convém, de início, chamar a atenção para o fato de que a imagem inversa $f^{-1}(c)$ às vezes não tem aspecto de curva ou superfície. (Por exemplo, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ pode ser constante, igual a c , num conjunto que contenha uma bola.) Melhor seria chamar $f^{-1}(c)$ de “conjunto de nível”. Mas a terminologia está consagrada e se justifica devido a $f^{-1}(c)$ ser mesmo uma superfície (ou uma curva) sempre que $\nabla f(x) \neq 0$ para todo x com $f(x) = c$, conforme provaremos adiante, com ajuda do teorema da função implícita.

Dizer que um vetor w é perpendicular à superfície (ou curva) de nível $f^{-1}(c)$ no ponto a significa que w é perpendicular ao vetor velocidade, no ponto a , de qualquer caminho diferenciável no ponto $t = 0$, com $\lambda(0) = a$ e $\lambda(t) \in f^{-1}(c)$, isto é, $f(\lambda(t)) = c$, para todo $t \in (-\epsilon, \epsilon)$. Com efeito, desta última igualdade segue-se que

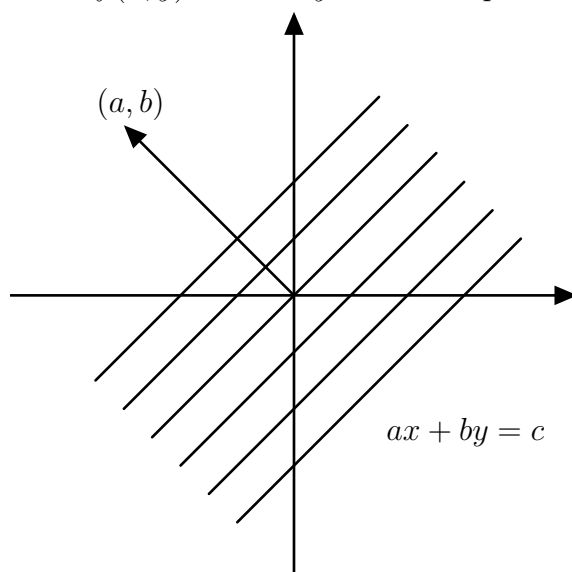
$$0 = (f \circ \lambda)'(0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot \lambda'_i(0) = \langle \nabla f(a), \lambda'(0) \rangle,$$

logo $\nabla f(a)$ é perpendicular a $\lambda'(0)$, vetor velocidade no ponto $a = \lambda(0)$ de qualquer caminho diferenciável λ , contido na superfície de nível de f que contém a .

Isto conclui a verificação das três propriedades do gradiente acima enunciadas. Vejamos agora alguns exemplos simples.

Exemplo 120. Sejam $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$ definidas por $f(x, y) = ax + by$ (onde $a^2 + b^2 \neq 0$), $g(x, y) = x^2 + y^2$ e $h(x, y) = x^2 - y^2$. As curvas de nível de f são as retas definidas pelas equações $ax + by = c$, para qualquer c real. O vetor gradiente de f é constante: $\nabla f = (a, b)$ em qualquer ponto $(x, y) \in \mathbb{R}^2$. Assim, as curvas de nível de f são todas as retas perpendiculares ao vetor (a, b) ; tais retas são, evidentemente, paralelas umas às outras.

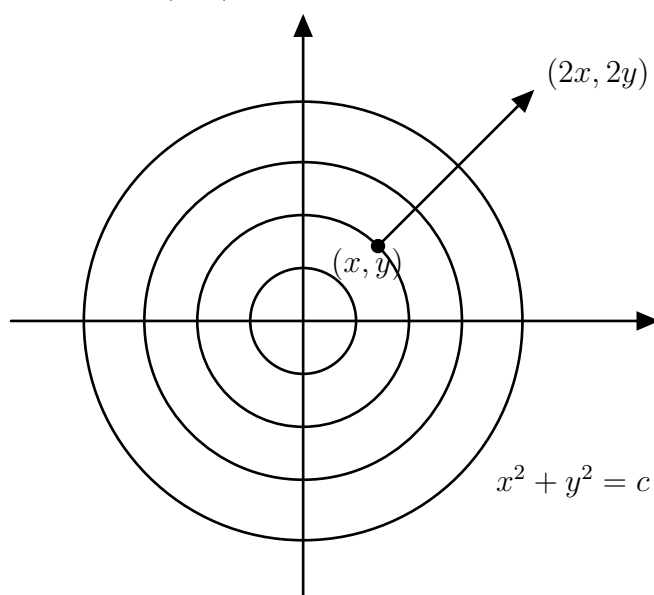
Figura 3.11: Curvas de nível de $f(x, y) = ax + by$ com retas paralelas perpendiculares a (a, b) .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

As curvas de nível c da função $g(x, y) = x^2 + y^2$ são as soluções de uma equação do tipo $g(x, y) = c$. Elas são vazias se $c < 0$. A curva de nível 0 reduz-se a um único ponto, a origem. Para $c > 0$, a curva de nível c é o círculo de centro na origem e raio \sqrt{c} . O vetor gradiente de g no ponto (x, y) é $\nabla g(x, y) = (2x, 2y)$, um vetor paralelo ao raio, o que é de se esperar, pois o raio é perpendicular a todo vetor tangente ao círculo naquele ponto.

Figura 3.12: Curvas de nível de $g(x, y) = x^2 + y^2$, círculos com centro na origem e raio \sqrt{c} .

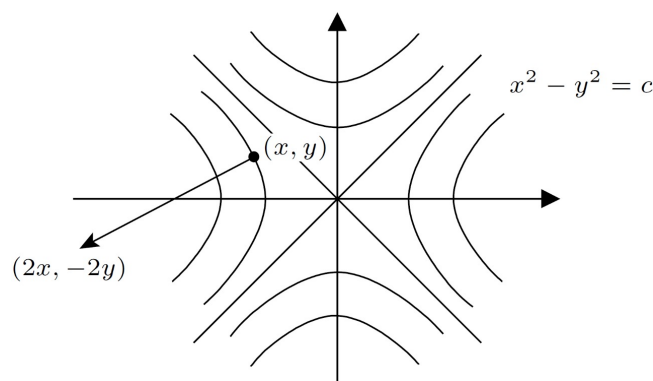


Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

A curva de nível c da função $h(x, y) = x^2 - y^2$ é formada por dois ramos de hipérbole quando $c \neq 0$.

No caso de $c > 0$, a hipérbole $x^2 - y^2 = c$ tem como eixo o eixo das abscissas; quando $c < 0$ o eixo da hipérbole é o das ordenadas. Para $c = 0$, a curva de nível $x^2 - y^2 = 0$ consiste em duas retas perpendiculares que se cortam na origem: as diagonais do primeiro e terceiro e do segundo e quarto quadrantes, dadas por $y = x$ e $y = -x$, respectivamente. O gradiente da função h é o vetor $\nabla h(x, y) = (2x, -2y)$. Atribuindo valores particulares a x e y , podemos observar que esse vetor é perpendicular à curva de nível que passa pelo ponto (x, y) , e indica a direção de crescimento de h .

Figura 3.13: Curvas de nível de $x^2 - y^2 = c$ e vetor gradiente $(2x, -2y)$.

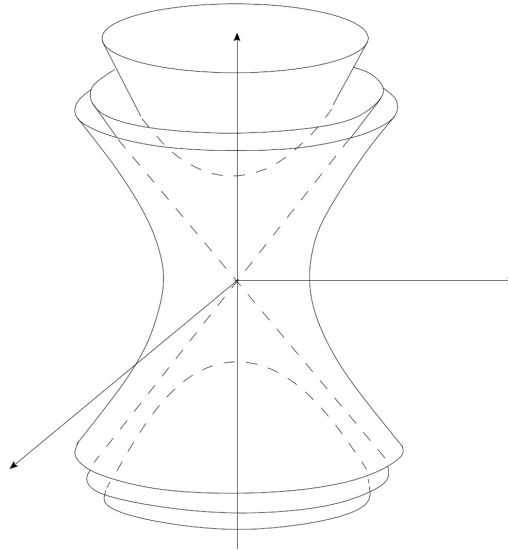


Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Note-se que, nos pontos onde o gradiente se anula (a origem na segunda e na terceira figura), ocorre uma quebra de regularidade na disposição das curvas de nível. Chamam-se *pontos singulares* ou *pontos críticos* da função os pontos onde seu gradiente é o vetor zero.

Exemplo 121. *Análogas tridimensionais das funções do exemplo anterior são $f, g, h: \mathbb{R}^3 \rightarrow \mathbb{R}$, definidas por $f(x, y, z) = ax + by + cz$ (com $a^2 + b^2 + c^2 \neq 0$), $g(x, y, z) = x^2 + y^2 + z^2$ e $h(x, y, z) = x^2 + y^2 - z^2$. As superfícies de nível de f são planos paralelos, todos perpendiculares ao vetor (a, b, c) , que é o gradiente de f . A superfície de nível c da função g é vazia se $c < 0$, reduz-se à origem se $c = 0$ e é uma esfera de centro na origem e raio \sqrt{c} quando $c > 0$. A superfície de nível 0 da função h é um cone duplo, com vértice na origem e eixo no eixo dos z . Se $c > 0$, a superfície $x^2 + y^2 - z^2 = c$ é um hiperbóloide de revolução, com o mesmo eixo do cone. Se $c < 0$, $x^2 + y^2 - z^2 = c$ define um hiperbóloide de duas folhas.*

Figura 3.14: Superfícies de nível de $x^2 + y^2 - z^2 = c$: cone ($c = 0$), hiperbolóides ($c \neq 0$).



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

O gradiente fornece uma visão central sobre a estrutura das funções diferenciáveis, ao indicar as direções de maior variação. Esse conceito conecta-se ao Teorema da Função Implícita, discutido na próxima seção.

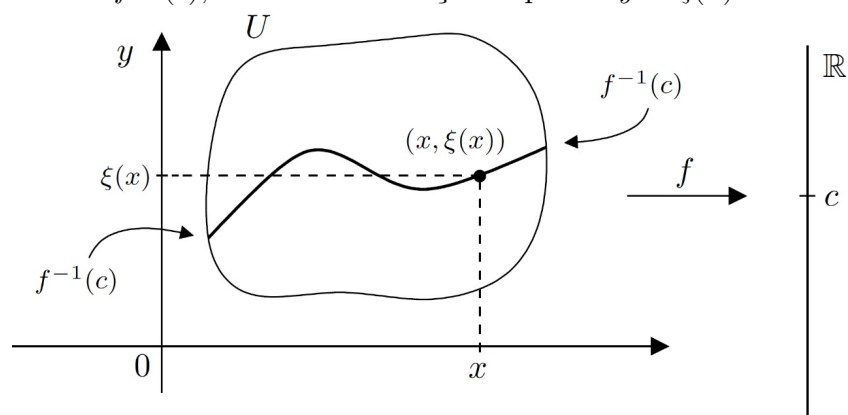
3.7 O Teorema da Função Implícita

O Teorema da Função Implícita é um dos resultados mais importantes na análise de várias variáveis, pois estabelece condições sob as quais uma função definida implicitamente por um sistema de equações pode ser expressa em termos de outras variáveis. Esse teorema permite “desmembrar” sistemas de equações e entender como variáveis interdependentes se comportam, garantindo a existência de funções que resolvem o sistema localmente. Nesta seção, investigamos o enunciado do teorema, suas hipóteses e implicações, que são essenciais para explorar a estrutura das funções em múltiplas variáveis.

Por simplicidade, consideramos inicialmente funções de duas variáveis.

Definição 122. 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 $f(x, y) = c \iff y = \xi(x)$. Isto quer dizer que $f^{-1}(c)$ é o gráfico da função ξ .

Figura 3.15: Gráfico de $f^{-1}(c)$, ilustrando a função implícita $y = \xi(x)$ definida por $f(x, y) = c$.



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Exemplo 123. *É mais comum acontecer que uma equação do tipo $f(x, y) = c$ (quando define alguma coisa) defina y como função de x , ou x como 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 2 valores de y tais que $x^2 + y^2 = 1$.] Mas, se tomarmos $U_1 = \{(x, y) \in \mathbb{R}^2; y > 0\}$, $U_2 = \{(x, y) \in \mathbb{R}^2; y < 0\}$, $U_3 = \{(x, y) \in \mathbb{R}^2; x > 0\}$ 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 $y = -\sqrt{1 - x^2}$ para $(x, y) \in U_2$, a $x = \sqrt{1 - y^2}$ se $(x, y) \in U_3$ e a $x = -\sqrt{1 - y^2}$ para $(x, y) \in U_4$. 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 4 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 $y = \xi(x)$].*

Notemos que é bem possível a uma equação do tipo $f(x, y) = c$ não definir função alguma: basta que c não pertença à imagem de f . 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 a equação $f(x, y) = c$ possua soluções, elas podem não definir funções, tal 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 elucidativo é o seguinte: a origem é solução da equação $x^2 - y^2 = 0$, mas, para nenhum aberto V contendo a origem, a interseção $f^{-1}(0) \cap V$ é o gráfico de uma função $y = \xi(x)$ ou $x = \zeta(y)$, pois tal interseção contém sempre 2 segmentos de reta de inclinação ± 1 que se cortam na origem.

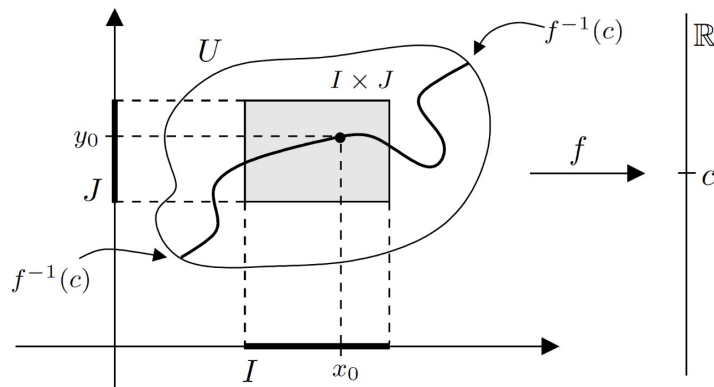
Feitas essas considerações, enunciemos o

Teorema 124. (Teorema da Função Implícita.) *Sejam $f : U \rightarrow \mathbb{R}$ uma função de classe \mathcal{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 \mathcal{C}^k . Tem-se $\xi'(x) = \frac{\partial f / \partial x}{\partial f / \partial y}$, estas derivadas sendo calculadas no ponto $(x, \xi(x))$.*

Como $(x_0, y_0) \in I \times J$, o intervalo aberto I contém x_0 , enquanto J contém y_0 .

A afirmação de que $f^{-1}(c) \cap (I \times J)$ 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 $f(x, y) = c$.

Figura 3.16: Curva $f^{-1}(c)$ em $I \times J$, com $y = \xi(x)$ definida implicitamente por $f(x, y) = c$.



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Demonstração: Para fixar as ideias, suponhamos $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $\frac{\partial f}{\partial y}$ é contínua, existem $\delta > 0$ e $\varepsilon > 0$ tais que, pondo $I = (x_0 - \delta, x_0 + \delta)$ e $J = (y_0 - \varepsilon, y_0 + \varepsilon)$, temos $I \times \bar{J} \subset U$ e $\frac{\partial f}{\partial y}(x, y) > 0$ para todo ponto $(x, y) \in I \times \bar{J}$. 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 $f(x_0, y_0 - \varepsilon) < c$ e $f(x_0, y_0 + \varepsilon) > c$. Pela continuidade de f , podemos supor δ tão pequeno que, para todo $x \in I$, tenhamos $f(x, y_0 - \varepsilon) < c$ e $f(x, y_0 + \varepsilon) > c$. Em virtude do Teorema do Valor Intermediário, 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 \mathbf{C}^k , ou seja, que existe $\xi'(x)$ para todo $x \in I$ e que $\xi' : I \rightarrow \mathbb{R}$ é de classe \mathbf{C}^{k-1} .

Ora, pondo $k = \xi(x+h) - \xi(x)$, temos $\xi(x+h) = \xi(x) + k$, logo $f(x+h, \xi(x) + k) = f(x, \xi(x)) = c$. Pelo Teorema do Valor Médio (93), existe θ , com $0 < \theta < 1$, tal que

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

Daí

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

Segundo o lema que provaremos logo a seguir, 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 $f \in \mathbf{C}^1$, sendo $\partial f/\partial x$, $\partial f/\partial y$ e ξ contínuas, esta fórmula mostra que ξ' é contínua, logo $\xi \in \mathbf{C}^1$. Se $f \in \mathbf{C}^2$, então $\partial f/\partial x$, $\partial f/\partial y$ e (como acabamos de mostrar) ξ são de classe \mathbf{C}^1 . A fórmula que dá ξ' mostra então que ξ' é também de classe \mathbf{C}^1 , isto é, $\xi \in \mathbf{C}^2$. E assim por diante: se $f \in \mathbf{C}^k$, então $\xi \in \mathbf{C}^k$. ■

Vejamos agora o lema usado na demonstração.

Lema 125. *Sejam $X \subset \mathbb{R}^m$, $K \subset \mathbb{R}^k$ compacto, $f : X \times K \rightarrow \mathbb{R}^p$ contínua e $c \in \mathbb{R}^p$. Se $f^{-1}(c)$ é o gráfico de uma aplicação $\xi : X \rightarrow K$, [isto é, para cada $x \in X$ existe um único $y = \xi(x) \in K$ com $f(x, \xi(x)) = c$], então ξ é contínua.*

Demonstração: Dado $x_0 \in X$, seja $y_0 = \xi(x_0)$. Tomamos uma sequência de pontos $x_n \in X$, 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 $\xi(x_n) \in K$ 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 $y \in K$ pois K 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. ■

Contra-exemplo ao Lema, supondo apenas K limitado mas não compacto: seja $f : \mathbb{R} \times [0, 1) \rightarrow \mathbb{R}$ definida por $f(x, y) = (x^2 + y^2)(ye^{|x|} - 1)$. Então $f^{-1}(0)$ é o gráfico da função $\xi : \mathbb{R} \rightarrow [0, 1)$,

onde $\xi(x) = e^{-|x|}$ se $x \neq 0$ e $\xi(0) = 0$. Embora f seja contínua, ξ é descontínua.

Para interpretar geometricamente o Teorema da Função Implícita, vamos introduzir algumas definições.

Definição 126. *Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável no aberto $U \subset \mathbb{R}^n$. Diremos que o número real c é um valor regular de f quando não existirem pontos críticos de f no nível c , ou seja, $f(x) = c \Rightarrow \nabla f(x) \neq 0$.*

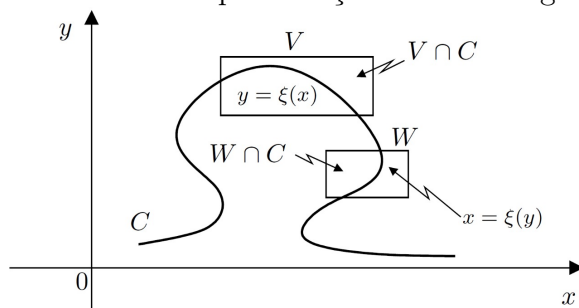
Note que esta definição está formulada de tal modo que $f^{-1}(c) = \emptyset$ implica automaticamente que c é um valor regular de f .

Definição 127. *Quando c é um valor regular de f , diz-se também que o nível c é regular. Quando existem pontos críticos $x \in U$ tais que $f(x) = c$, dizemos que c é um nível crítico de f .*

Exemplo 128. *Qualquer número real diferente de 3 é valor regular da função $f(x, y) = x^2 - y^2 + 3$ pois $\nabla f(x, y) = (2x, -2y)$ só se anula no ponto $(x, y) = (0, 0)$, no qual o valor de f é 3. Por outro lado $f(0, 0) = 3$ não é valor regular de f pois $\nabla f(0, 0) = (0, 0)$.*

Definição 129. *Um conjunto $C \subset \mathbb{R}^2$ chama-se uma curva de classe \mathcal{C}^k ($k \geq 0$) quando C é localmente o gráfico de uma função de classe \mathcal{C}^k . Isto quer dizer que cada ponto $p \in C$ está contido num aberto V tal que $V \cap C$ é o gráfico de uma função de classe \mathcal{C}^k .*

Figura 3.17: Curva C de classe \mathcal{C}^k com representação local como gráfico de funções em V e W .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Exemplo 130. *Vimos no Exemplo 123 que o círculo S^1 é uma curva de classe \mathcal{C}^∞ . Seja agora $C = \{(x, y) \in \mathbb{R}^2; x^2 - y^2 = 1\}$ (hipérbole). Afirmamos que C é uma curva (desconexa) de classe \mathcal{C}^∞ . Com efeito, sejam $V_1 = \{(x, y) \in \mathbb{R}^2; x > 0\}$ e $V_2 = \{(x, y) \in \mathbb{R}^2; x < 0\}$. Todo ponto $p \in C$ pertence a V_1 ou a V_2 . Além disso, $V_1 \cap C$ é o gráfico da função $x = \sqrt{1 + y^2}$, enquanto $V_2 \cap C$ é o gráfico de $x = -\sqrt{1 + y^2}$. Ambas são funções de classe \mathcal{C}^∞ na reta, logo C é uma curva de classe \mathcal{C}^∞ .*

O enunciado geométrico do Teorema da Função Implícita é o seguinte:

Teorema 131. *Seja $f : U \rightarrow \mathbb{R}$ uma função de classe \mathcal{C}^k ($k \geq 1$) no aberto $U \subset \mathbb{R}^2$. Para todo valor regular c da função f , o conjunto $f^{-1}(c)$ (se não for vazio) é uma curva de classe \mathcal{C}^k , chamada a curva de nível c da função f .*

Observação 132. *Quando c não é um valor regular de f , a imagem inversa $f^{-1}(c)$ pode ou não ser uma curva. Por exemplo, se $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ é dada por $f(x, y) = x^2 - y^2 + 3$ então $f^{-1}(3)$, reunião de duas retas que se cortam na origem, não é uma curva. Mas se $g(x, y) = x^2$ então 0 não é um valor regular de g , mas $g^{-1}(0)$, o eixo dos y , é uma curva.*

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 133. (Teorema da Função Implícita.) *Seja $f : U \rightarrow \mathbb{R}$ uma função de classe \mathcal{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 $J = (y_0 - \varepsilon, y_0 + \varepsilon)$ tais que $f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi : B \rightarrow J$, de classe \mathcal{C}^k . Para todo $x \in B$, tem-se*

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

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 $n = 1$. Vamos, entretanto, repetir a prova, por cortesia. Para fixar ideias, seja $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $\partial f / \partial y$ é contínua, existem $\delta > 0$, $\varepsilon > 0$ tais que, pondo $B = B(x_0; \delta)$ e $J = (y_0 - \varepsilon, y_0 + \varepsilon)$, temos $B \times J \subset U$ e $\frac{\partial f}{\partial y}(x, y) > 0$ para todo ponto $(x, y) \in B \times J$. Então, para todo $x \in B$, a função $y \mapsto f(x, y)$ é estritamente crescente no intervalo $J = [y_0 - \varepsilon, y_0 + \varepsilon]$. Como $f(x_0, y_0) = c$, segue-se que $f(x_0, y_0 - \varepsilon) < c$ e $f(x_0, y_0 + \varepsilon) > c$. Pela continuidade de f , podemos supor δ tão pequeno que, para todo $x \in B$, tenhamos $f(x, y_0 - \varepsilon) < c$ e $f(x, y_0 + \varepsilon) > c$. Em virtude do Teorema do Valor Intermediário, existe, para cada $x \in B$, um único $y = \xi(x) \in J$ tal que $f(x, y) = c$. Tem-se obrigatoriamente $y \in J$, logo $f^{-1}(c) \cap (B \times J) = f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi : B \rightarrow J$, a qual, pelo lema anterior, 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 $f(x + te_i, \xi(x) + k) = f(x, \xi(x)) = c$ para todo $t \in (-\delta, \delta)$.

Pelo Teorema do Valor Médio, 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 , temos $\lim_{t \rightarrow 0} k(t) = 0$. 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 $f \in \mathcal{C}^1$, resulta desta fórmula que as derivadas parciais de ξ são contínuas, logo $\xi \in \mathcal{C}^1$. Se $f \in \mathcal{C}^2$, então suas derivadas parciais são de classe \mathcal{C}^1 . Como já temos $\xi \in \mathcal{C}^1$, resulta ainda da fórmula acima que as derivadas parciais de ξ são de classe \mathcal{C}^1 , logo $\xi \in \mathcal{C}^2$. E assim por diante: $f \in \mathcal{C}^k$ implica $\xi \in \mathcal{C}^k$.

Corolário 134. *Seja $f : U \rightarrow \mathbb{R}$ de classe \mathcal{C}^k ($k \geq 1$) no aberto $U \subset \mathbb{R}^{n+1}$. Se $\xi : V \rightarrow \mathbb{R}$ é contínua no aberto $V \subset \mathbb{R}^n$ com $(x, \xi(x)) \in U$, $\frac{\partial f}{\partial y}(x, \xi(x)) \neq 0$ e $f(x, \xi(x)) = c$ para todo $x \in V$, então ξ é de classe \mathcal{C}^k .*

Evidentemente, não há nada especial a respeito da última variável. Dada $f : U \rightarrow \mathbb{R}$, de classe \mathcal{C}^k no aberto $U \subset \mathbb{R}^{n+1}$, se num ponto $p \in U$ tivermos $f(p) = c$ e $\frac{\partial f}{\partial x_i}(p) \neq 0$ para algum inteiro $i \in [1, n + 1]$, então existe um aberto $V \subset \mathbb{R}^{n+1}$, contendo p , tal que $f^{-1}(c) \cap V$ é o gráfico de uma função de n variáveis, de classe \mathcal{C}^k . Mais precisamente, as condições $x \in V$, $f(x) = c$ definem, de modo único, a i -ésima coordenada de x em função das n restantes: $x_i = \xi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$, sendo ξ uma função de classe \mathcal{C}^k .

Definição 135. *Um conjunto $M \subset \mathbb{R}^{n+1}$ chama-se uma hipersfície de classe \mathcal{C}^k quando é*

localmente o gráfico de uma função de n variáveis de classe \mathbf{C}^k . Isto significa que cada ponto $p \in M$ pertence a um aberto $V \subset \mathbb{R}^{n+1}$ tal que $V \cap M$ é o gráfico de uma função de classe \mathbf{C}^k definida num aberto do espaço \mathbb{R}^n . Quando $n = 1$ diz-se “curva” e, se $n = 2$, diz-se “superfície” em vez de “hiperfície”.

A definição acima inclui o caso $k = 0$. As hiperfícies de classe \mathbf{C}^0 têm interesse apenas topológico, não possuindo propriedades diferenciais. Podemos considerar também as *hiperfícies diferenciáveis* (caso intermediário entre \mathbf{C}^0 e \mathbf{C}^1), que são localmente gráficos de funções diferenciáveis. Evidentemente, toda hiperfície de classe \mathbf{C}^k ($k \geq 1$) é diferenciável.

Exemplo 136. Seja $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x \rangle = 1\}$ a esfera unitária n -dimensional. Indiquemos com $U \subset \mathbb{R}^n$ a bola aberta de raio 1, com centro na origem. Para cada $i = 1, \dots, n+1$, ponhamos $V_i = \{x \in \mathbb{R}^{n+1}; x_i > 0\}$ e $W_i = \{x \in \mathbb{R}^{n+1}; x_i < 0\}$. Escrevendo $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$, temos:

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

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

Logo, se considerarmos a função $\xi : U \rightarrow \mathbb{R}$, de classe \mathbf{C}^∞ , definida por

$$\xi(u) = \sqrt{1 - \langle u, u \rangle},$$

vemos que, para cada $i = 1, \dots, n+1$, $S^n \cap V_i$ é o gráfico da função $x_i = \xi(x^*)$, enquanto que $S^n \cap W_i$ é o gráfico de $x_i = -\xi(x^*)$. Como todo ponto $p \in S^n$ está contido em algum V_i ou em algum W_i , concluímos que S^n é uma hiperfície de classe \mathbf{C}^∞ em \mathbb{R}^{n+1} .

Observação 137. Seja $M \subset \mathbb{R}^{n+1}$. Dado $p \in M$, usaremos a notação $T_p M$ para indicar o conjunto dos vetores velocidade $\lambda'(0)$, dos caminhos $\lambda : (-\epsilon, \epsilon) \rightarrow M \subset \mathbb{R}^{n+1}$, contidos em M , diferenciáveis no ponto $t = 0$ e tais que $\lambda(0) = p$.

Definição 138. Quando M é uma hiperfície diferenciável, o conjunto $T_p M$ chama-se o espaço vetorial tangente a M no ponto p .

Esta denominação tem sua justificativa no

Teorema 139. Se a hiperfície $M \subset \mathbb{R}^{n+1}$ é diferenciável, então, para cada $p \in M$, o conjunto $T_p M$ é um subespaço vetorial de dimensão n do espaço euclidiano \mathbb{R}^{n+1} .

Demonstração: Dado $p = (a_1, \dots, a_{n+1})$ em M , existem abertos $V \subset \mathbb{R}^{n+1}$, $U \subset \mathbb{R}^n$, com $p \in V$, um inteiro $i \in [1, n+1]$ e uma função $\xi : U \rightarrow \mathbb{R}$, diferenciável, tais que $x \in V \cap M \Leftrightarrow$

$x_i = \xi(x^*)$, onde $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) \in U$. (Assim, nenhuma das variáveis de ξ tem índice i .) Escrevendo também $p^* = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_{n+1})$, afirmamos que

$$v = (\alpha_1, \dots, \alpha_{n+1}) \in T_p M \Leftrightarrow \alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \alpha_j. \quad (3.3)$$

De fato, em primeiro lugar, se $v \in T_p M$, então $v = \lambda'(0)$, onde, restringindo ε se necessário, podemos supor que $\lambda : (-\varepsilon, \varepsilon) \rightarrow M \cap V$, $\lambda(0) = p$. Então, para todo $t \in (-\varepsilon, \varepsilon)$, temos

$$\lambda_i(t) = \xi(\lambda_1(t), \dots, \lambda_{i-1}(t), \lambda_{i+1}(t), \dots, \lambda_{n+1}(t)).$$

Pela Regra da Cadeia,

$$\lambda'_i(0) = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \lambda'_j(0), \quad \text{isto é,} \quad \alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \alpha_j,$$

como foi afirmado.

Reciprocamente, se o vetor $v = (\alpha_1, \dots, \alpha_{n+1})$ é tal que $\alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \alpha_j$, então definimos um caminho $\lambda : (-\varepsilon, \varepsilon) \rightarrow M \cap V$ tomando $\varepsilon > 0$ tal que $|t| < \varepsilon \Rightarrow p^* + tv^* \in U$, pondo $\lambda_j(t) = a_j + t\alpha_j$ para $j \neq i$ e

$$\lambda_i(t) = \xi(a_1 + t\alpha_1, \dots, a_{i-1} + t\alpha_{i-1}, a_{i+1} + t\alpha_{i+1}, \dots, a_{n+1} + t\alpha_{n+1}) = \xi(p^* + tv^*).$$

Pela Regra da Cadeia, temos $\lambda'(0) = v$, logo $v \in T_p M$. A caracterização (3.3) acima obtida mostra que $T_p M$ é um subespaço vetorial de dimensão n de \mathbb{R}^{n+1} , gerado pelos n vetores linearmente independentes

$$e_1 + c_1 e_i, \dots, e_{i-1} + c_{i-1} e_i, e_{i+1} + c_{i+1} e_i, \dots, e_{n+1} + c_{n+1} e_i,$$

onde $c_j = (\partial \xi / \partial x_j)(p^*)$. Outra maneira de interpretar a afirmação (3.3) é dizer que ela caracteriza $T_p M$ como o núcleo do funcional linear não nulo $\varphi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, dado por $\varphi(v) = \alpha_i - \sum_{j \neq i} c_j \alpha_j$, onde $v = (\alpha_1, \dots, \alpha_{n+1})$ e c_j é a j -ésima derivada parcial de ξ no ponto p^* . Ou

ainda: $T_p M$ é o gráfico do funcional linear $d\xi(p^*)$. ■

Exemplo 140. Para hipersuperfícies $M \subset \mathbb{R}^{n+1}$ de classe \mathcal{C}^0 , $T_p M$ pode não ser um espaço vetorial de dimensão n ; se $X = \{(x, y, z) \in \mathbb{R}^3; z = x^2 + y^2\}$ (cone), $Y = \{(x, y, z) \in \mathbb{R}^3; z = |x|\}$ e $p = (0, 0, 0)$, então $T_p X$ consiste apenas do vetor 0, enquanto $T_p Y$ é o subespaço vetorial de dimensão 1 em \mathbb{R}^3 , formado pelos vetores $(0, \beta, 0)$, $\beta \in \mathbb{R}$.

Exemplo 141. Sabemos que a esfera $S^n \subset \mathbb{R}^{n+1}$ é uma hipersuperfície de classe C^∞ . Para cada ponto $p \in S^n$, o espaço vetorial tangente $T_p(S^n)$ é o conjunto $[p]^\perp$ dos vetores $v \in \mathbb{R}^{n+1}$ tais que $\langle p, v \rangle = 0$. Com efeito, se $\lambda : (-\varepsilon, \varepsilon) \rightarrow S^n$ é um caminho diferenciável no ponto $t = 0$, com $\lambda(0) = p$, então, como $|\lambda(t)| = 1$ para todo t , temos $\langle \lambda'(0), \lambda(0) \rangle = 0$, isto é, $\lambda'(0) \in [p]^\perp$. Isto mostra que $T_p(S^n) \subset [p]^\perp$. Como $[p]^\perp$ é um subespaço n -dimensional de \mathbb{R}^{n+1} , resulta que esta inclusão é, na realidade, uma igualdade.

O teorema abaixo permite obter um grande número de exemplos de hipersuperfícies. Ele diz que, se não há pontos críticos de f no nível c , então a “superfície de nível c ” da função f é, de fato, uma hipersuperfície.

Teorema 142. (Teorema Global da Função Implícita.) A imagem inversa $M = f^{-1}(c)$ de um valor regular c de uma função $f : U \rightarrow \mathbb{R}$, de classe C^k ($k \geq 1$) em um aberto $U \subset \mathbb{R}^{n+1}$, é uma hipersuperfície de classe C^k . Em cada ponto $p \in M$, o espaço vetorial tangente $T_p M$ é o núcleo da diferencial $df(p) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ ou, equivalentemente, o conjunto dos vetores $v \in \mathbb{R}^{n+1}$ perpendiculares ao vetor $\nabla f(p)$.

Demonstração: O fato de que $M = f^{-1}(c)$ é uma hipersuperfície de classe C^k decorre imediatamente do Teorema da Função Implícita. Quanto a $T_p M$, se o caminho $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$, com $\lambda(0) = p$, é diferenciável no ponto $t = 0$, então, sendo $f(\lambda(t)) = c$ para todo $t \in (-\varepsilon, \varepsilon)$, segue-se que $0 = (f \circ \lambda)'(0) = df(p) \cdot \lambda'(0) = \langle \nabla f(p), \lambda'(0) \rangle$. Assim, $T_p M$ está contido no conjunto dos vetores de \mathbb{R}^{n+1} que são perpendiculares a $\nabla f(p)$. Como este conjunto é também um subespaço vetorial de dimensão n , segue-se que ele coincide com $T_p M$. ■

Exemplo 143. Reexaminemos a esfera S^n à luz do teorema acima. Considerando a função $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ (de classe C^∞), definida por $f(x) = \langle x, x \rangle$, temos $\frac{\partial f}{\partial x_i}(x) = 2x_i$, logo $\nabla f(x) = 2x$ para todo $x \in \mathbb{R}^{n+1}$. Assim, $\nabla f(x) = 0 \Leftrightarrow x = 0 \Leftrightarrow f(x) = 0$. Em particular, 1 é um valor regular de f , donde $S^n = f^{-1}(1)$ é uma hipersuperfície em \mathbb{R}^{n+1} . Para cada $p \in S^n$, como $\nabla f(p) = 2p$, temos $T_p(S^n) = \{v \in \mathbb{R}^{n+1} \mid \langle v, 2p \rangle = 2\langle v, p \rangle = 0\} = [p]^\perp$.

Exemplo 144. A função $\det : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$, que associa a cada matriz $n \times n$, $X = (x_{ij})$, o seu determinante, é de classe C^∞ . A expansão de $\det X$ segundo os elementos da i -ésima linha nos dá $\det X = \sum_{j=1}^n (-1)^{i+j} x_{ij} X[ij]$, onde $X[ij]$ é o determinante da matriz $(n-1) \times (n-1)$ que se

obtem omitindo a i -ésima linha e a j -ésima coluna de X . Logo $\frac{\partial \det}{\partial x_{ij}}(X) = (-1)^{i+j} X[ij]$, para

cada $X \in \mathbb{R}^{n^2}$. Em particular, no ponto $X = I$ (= matriz identidade $n \times n$), temos $\frac{\partial \det}{\partial x_{ij}}(I) = \delta_{ij}$

(“delta de Kronecker”: $\delta_{ij} = 0$ se $i \neq j$ e $\delta_{ii} = 1$). Isto significa que o gradiente da função \det no ponto I é a matriz identidade. Seja agora $U \subset \mathbb{R}^{n^2}$ o subconjunto aberto formado pelas matrizes $n \times n$ que têm determinante $\neq 0$ (matrizes invertíveis). Então a restrição $\det : U \rightarrow \mathbb{R}$ é uma função sem pontos críticos. [Com efeito, basta que, para um dos índices i ou j fixado, todos os números $X[ij]$ sejam nulos para que ocorra $\det X = \sum_i (-1)^{i+j} x_{ij} X[ij] = \sum_j (-1)^{i+j} x_{ij} X[ij] = 0$, logo $X \notin U$.

Assim, todo número real é um valor regular para a função $\det : U \rightarrow \mathbb{R}$. Em particular, $M = (\det)^{-1}(1)$, ou seja, o conjunto das matrizes $n \times n$ que têm determinante igual a 1, é uma hipersuperfície em \mathbb{R}^{n^2} . [M é um grupo relativamente à multiplicação de matrizes, conhecido como o “grupo unimodular” de \mathbb{R}^n .] O espaço vetorial tangente $T_I M$, no ponto I (matriz identidade), consiste nas matrizes $X = (x_{ij})$ que são perpendiculares ao gradiente de \det nesse ponto, o qual sabemos que é $I = (\delta_{ij})$. Ora, temos $\langle X, I \rangle = \sum_{i,j} x_{ij} \delta_{ij} = \sum_i x_{ii}$, que é o traço de X .

Assim, o espaço vetorial tangente a M no ponto I é o conjunto das matrizes de traço nulo.

Observação 145. Toda hipersuperfície $M \subset \mathbb{R}^{n+1}$, sendo localmente o gráfico de uma função $x_i = \xi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) = \xi(x^*)$, de n variáveis, é também localmente a imagem inversa $f^{-1}(0)$ do valor regular 0 pela função $f(x) = x_i - \xi(x^*)$, definida no aberto V tal que $M \cap V$ é o gráfico de ξ . Não é verdade, porém, que toda hipersuperfície $M \subset \mathbb{R}^{n+1}$ seja globalmente do tipo $M = f^{-1}(c)$, imagem inversa de um valor regular, como no teorema acima. Se M é deste tipo, então a aplicação contínua $\varphi = \nabla f : M \rightarrow \mathbb{R}^{n+1}$ fornece o que se chama um “campo contínuo de vetores normais não nulos ao longo de M ”. Aqui, normal significa que, para cada $p \in M$, $\varphi(p) = \nabla f(p)$ é perpendicular a todo vetor tangente $v \in T_p M$. As hipersuperfícies que admitem um campo contínuo de vetores normais não nulos $\varphi : M \rightarrow \mathbb{R}^{n+1}$ chamam-se hipersuperfícies orientáveis. Um exemplo de superfície não-orientável em \mathbb{R}^3 é a faixa de Möbius.

Com o término desta seção, concluímos este capítulo, que antecede a discussão do capítulo seguinte, dedicado ao Teorema do Multiplicador de Lagrange.

Capítulo 4

Multiplicador de Lagrange

Finalmente, chega-se no Capítulo 4, onde será apresentado e demonstrado o Teorema do Multiplicador de Lagrange. Os conceitos e definições que foram mencionados nos capítulos anteriores serão utilizados para tal demonstração. A seguir, será feita uma contextualização para o prosseguimento com o estudo do teorema.

Sejam $M \subset \mathbb{R}^{n+1}$ uma hipersuperfície de classe \mathcal{C}^k ($k \geq 1$), contida num aberto $U \subset \mathbb{R}^{n+1}$ e $f : U \rightarrow \mathbb{R}$ uma função real de classe \mathcal{C}^k . Pretendemos determinar os máximos locais, mínimos locais e, mais geralmente, os pontos críticos da restrição $f|_M$. Antes de mais nada precisamos definir o que entendemos por um ponto crítico de $f|_M$.

Definição 146. *Os pontos críticos de f em U são os pontos $x \in U$ tais que $\nabla f(x) = 0$, isto é, $\frac{\partial f}{\partial v}(x) = 0$ para todo $v \in \mathbb{R}^{n+1}$.*

Isto quer dizer que, para todo caminho diferenciável $\lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^{n+1}$ com $\lambda(0) = x$, temos $(f \circ \lambda)'(0) = 0$. Usando esse fato como guia, apresentamos a definição abaixo.

Definição 147. *Um ponto crítico de $f|_M$ é um ponto $p \in M$ tal que $(f \circ \lambda)'(0) = 0$ para todo caminho diferenciável $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$, com $\lambda(0) = p$.*

Isto significa que $\frac{\partial f}{\partial v}(p) = 0$ para todo $v \in T_p M$, ou seja, $p \in M$ é um ponto crítico da restrição $f|_M$ se, e somente se, o vetor $\nabla f(p)$ é normal à superfície M no ponto p .

Se $p \in M$ é um ponto de máximo local (ou mínimo local) para a restrição $f|_M$ então, para todo caminho diferenciável $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$, com $\lambda(0) = p$, 0 é um ponto de máximo (ou de mínimo) local para a função real $f \circ \lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$, logo $(f \circ \lambda)'(0) = 0$, portanto p é um ponto crítico de $f|_M$ de acordo com a definição acima.

Evidentemente, os pontos críticos de f em U , que por acaso pertençam a M , se existirem, serão pontos críticos de $f|_M$. O interessante, porém, é que podem ocorrer pontos críticos de

$f|_M$ que não são pontos críticos de f , isto é, nos quais ∇f não se anula. Por exemplo, sejam $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, dada por $f(x, y) = y$, e $M = S^1 = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 = 1\}$. Evidentemente, f não possui pontos críticos, pois $\nabla f = (0, 1)$, em todos os pontos $(x, y) \in \mathbb{R}^2$. Mas $p = (0, 1)$ e $q = (0, -1)$ são pontos críticos de $f|_M$, inclusive porque são os pontos onde $f|_M$ atinge, respectivamente, seu máximo e seu mínimo. De um modo geral, se a superfície $M \subset \mathbb{R}^{n+1}$ é compacta, então $f|_M$ admite pelo menos dois pontos críticos, a saber, os pontos onde $f|_M$ assume seus valores máximo e mínimo.

A resposta ao nosso problema é dada pelo Teorema do Multiplicador de Lagrange, que será apresentado na próxima seção.

4.1 Teorema do Multiplicador de Lagrange

Teorema 148. (Teorema do Multiplicador de Lagrange.) *Consideremos $f : U \rightarrow \mathbb{R}$, uma função de classe \mathcal{C}^k ($k \geq 1$) no aberto $U \subset \mathbb{R}^{n+1}$, e $M = \varphi^{-1}(c)$ uma superfície contida em U , imagem inversa do valor regular $c \in \mathbb{R}$ por uma função $\varphi : U \rightarrow \mathbb{R}$, de classe \mathcal{C}^k . Um ponto $p \in M$ é ponto crítico de f se, e somente se, existe um número real λ tal que $\nabla f(p) = \lambda \cdot \nabla \varphi(p)$.*

Na próxima seção faremos a demonstração do Teorema.

4.2 Demonstração do Teorema do Multiplicador de Lagrange.

Demonstração: Queremos mostrar que, para uma função $f : U \rightarrow \mathbb{R}$ e uma restrição $M = \varphi^{-1}(c)$, onde $\varphi : U \rightarrow \mathbb{R}$, um ponto $p \in M$ é ponto crítico de f restrita a M se, e somente se, existe um número real λ tal que:

$$\nabla f(p) = \lambda \cdot \nabla \varphi(p).$$

Começaremos observando a estrutura dos elementos. A função f é definida em $U \subset \mathbb{R}^{n+1}$ e é de classe \mathcal{C}^k , o que significa que possui derivadas contínuas até a ordem k . A restrição é dada pelo conjunto de nível $M = \varphi^{-1}(c)$, onde φ é também uma função de classe \mathcal{C}^k . O valor $c \in \mathbb{R}$ é regular para φ , o que implica que $\nabla \varphi(p) \neq 0$ para todos os pontos $p \in M$.

Assim, M é uma superfície de dimensão n no espaço \mathbb{R}^{n+1} . Para que $p \in M$ seja um ponto crítico de f restrita a M , a derivada de f ao longo de qualquer direção tangente a M em p deve ser zero. Dizemos, portanto, que p é um ponto crítico de f restrita a M quando as variações

de f ao longo das direções permitidas em M não causam alterações no valor de f . Em termos do gradiente, essa condição equivale a dizer que

$$\nabla f(p) \perp T_p M,$$

onde $T_p M$ denota o espaço tangente a M no ponto p .

Considerando que $M = \varphi^{-1}(c)$ é um conjunto de nível de φ , segue-se que o vetor gradiente $\nabla \varphi(p)$ é ortogonal ao espaço tangente $T_p M$ em p . Em outras palavras, $\nabla \varphi(p)$ atua como um vetor normal a M em p , apontando em uma direção perpendicular a todas as direções tangentes à superfície M naquele ponto.

Para que p seja um ponto crítico de f em M , o gradiente $\nabla f(p)$ deve, então, ser ortogonal a todas as direções em $T_p M$. Isso significa que $\nabla f(p)$ deve pertencer ao espaço das direções ortogonais a $T_p M$. Como $T_p M \subset \mathbb{R}^{n+1}$ é de dimensão n , seu complemento ortogonal (ou seja, a direção normal a $T_p M$) será um espaço unidimensional. Esse espaço ortogonal é gerado por $\nabla \varphi(p)$, já que $\nabla \varphi(p) \neq 0$.

Portanto, para que $\nabla f(p)$ seja ortogonal a $T_p M$, ele deve ser um múltiplo escalar de $\nabla \varphi(p)$. Em outras palavras, existe um número real λ tal que:

$$\nabla f(p) = \lambda \cdot \nabla \varphi(p).$$

Esse número λ , que o chamamos de multiplicador de Lagrange no Capítulo 1, mede a intensidade com que a função f varia ao longo da direção de φ no ponto p . Em termos práticos, λ indica a “intensidade” com que f se opõe à restrição imposta por M , refletindo a relação entre o gradiente de f e a direção normal à superfície de nível de M .

Assim, concluímos que a condição

$$\nabla f(p) = \lambda \cdot \nabla \varphi(p)$$

é necessária e suficiente para que $p \in M$ seja ponto crítico de f restrita a M . Em outras palavras, p é um ponto crítico de f em M se, e somente se, o gradiente $\nabla f(p)$ é paralelo ao gradiente $\nabla \varphi(p)$, com a relação de proporcionalidade medida pelo multiplicador λ . ■

Vamos fazer algumas observações e apresentar alguns exemplos na seção a seguir.

4.3 Observações e Exemplos

Observação 149. Quando a superfície M não é dada como imagem inversa $\varphi^{-1}(c)$ de um valor regular, os pontos críticos de $f|_M$ são simplesmente os pontos $p \in M$ nos quais $\nabla f(p)$

é normal a M .

A pesquisa dos pontos críticos de $f | M$ reduz-se, portanto, a resolver o sistema de $n + 2$ equações

$$\begin{cases} \frac{\partial f}{\partial x_i}(p) = \lambda \cdot \frac{\partial \varphi}{\partial x_i}(p), & i = 1, 2, \dots, n + 1; \\ \varphi(p) = c, \end{cases}$$

nas $n + 2$ incógnitas $\lambda, x_1, \dots, x_{n+1}$, onde $p = (x_1, \dots, x_{n+1})$.

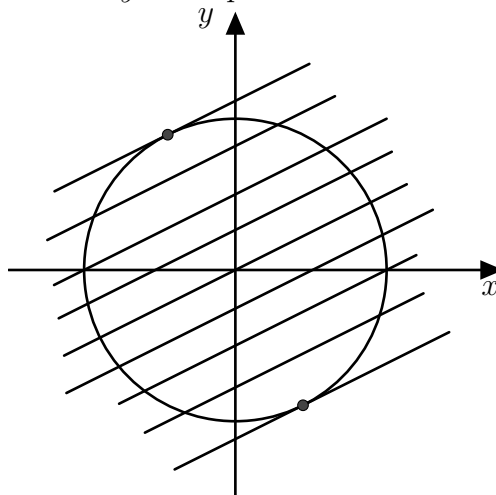
A presença de λ no sistema acima torna o número de equações igual ao número de incógnitas, o que muitas vezes ajuda a resolvê-lo.

A condição $\nabla f(p) = \lambda \cdot \nabla \varphi(p)$ significa que a hiperfície M é tangente à superfície de nível f que passa pelo ponto crítico p da função $f | M$. No caso em que se podem esboçar essas superfícies, esta observação auxilia a localizar os pontos críticos.

Vejam agora alguns problemas onde se aplica o método do multiplicador de Lagrange.

Exemplo 150. *Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = ax + by$, com $a^2 + b^2 \neq 0$. Quais são os pontos críticos da restrição de f ao círculo unitário S^1 ? Temos $\nabla f = (a, b)$, $S^1 = \varphi^{-1}(1)$, $\varphi(x, y) = x^2 + y^2$ e $\nabla \varphi = (2x, 2y)$. Os pontos críticos de $f | S^1$ são aqueles onde os vetores (a, b) e $(2x, 2y)$ são colineares. Como, além disso, deve-se ter $x^2 + y^2 = 1$, isto nos dá $x = \frac{a}{\sqrt{a^2 + b^2}}$, $y = \frac{b}{\sqrt{a^2 + b^2}}$ ou $x = \frac{-a}{\sqrt{a^2 + b^2}}$, $y = \frac{-b}{\sqrt{a^2 + b^2}}$. Nestes pontos, $f | S^1$ assume, respectivamente, seu valor máximo, igual a $\sqrt{a^2 + b^2}$, e seu valor mínimo, igual a $-\sqrt{a^2 + b^2}$.*

Figura 4.1: Retas $ax + by = c$ e pontos críticos no círculo unitário S^1 .



Fonte: LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2007.

Exemplo 151. Dados $a \in \mathbb{R}^{n+1}$ e uma hiperfície $M \subset \mathbb{R}^{n+1}$, que não contenha o ponto a , determinar o ponto $p \in M$ mais próximo a a . (Se M for um subconjunto fechado de \mathbb{R}^{n+1} , sabemos que um tal ponto p sempre existe.) Consideremos a função $f(x) = |x - a|$, que é de classe C^∞ em $\mathbb{R}^{n+1} - \{a\}$. Procuramos os pontos onde $f|_M$ assume seu valor mínimo. Eles estão entre os pontos críticos de $f|_M$, isto é, entre os pontos nos quais ∇f é normal a M . Ora, como $f(x) = \sqrt{\sum (x_i - a_i)^2}$, temos $\frac{\partial f}{\partial x_i} = \frac{x_i - a_i}{|x - a|}$, donde $\nabla f(x) = \frac{x - a}{|x - a|}$. Assim, os pontos críticos de $f|_M$ (entre os quais se encontram os pontos de M situados a uma distância mínima do ponto a) são os pontos $x \in M$ tais que $x - a$ é normal a M .

Exemplo 152. Seja (a_{ij}) uma matriz real $n \times n$ simétrica, isto é, $a_{ij} = a_{ji}$. A ela corresponde uma transformação linear $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$, definida por $A \cdot x = y$, com $y_i = \sum_j a_{ij}x_j$.

A condição $a_{ij} = a_{ji}$ é equivalente a $\langle Ax, y \rangle = \langle x, Ay \rangle$ para qualquer $x, y \in \mathbb{R}^n$. Diz-se então que a transformação linear A é auto-adjunta.

Definição 153. Um vetor $x \in \mathbb{R}^n$ chama-se um vetor próprio de A quando $x \neq 0$ e $A \cdot x = \lambda \cdot x$ para algum $\lambda \in \mathbb{R}$. O número λ chama-se então o valor próprio correspondente ao vetor próprio x .

Em geral, uma transformação linear $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ não precisa ter vetores próprios $x \in \mathbb{R}^n$ nem valores próprios reais. (Tome, por exemplo, uma rotação no plano, de ângulo θ , com $0^\circ < \theta < 180^\circ$). Mostraremos agora que se A é auto-adjunta então existe uma base ortonormal de \mathbb{R}^n formada por vetores próprios de A . Para isso, introduziremos a forma quadrática

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $f(x) = \langle A \cdot x, x \rangle$ ou, em termos de coordenadas, $f(x) = \sum_{i,j=1}^n a_{ij}x_ix_j$.

Estudaremos os pontos críticos de f na esfera unitária $S^{n-1} \subset \mathbb{R}^n$. Como $\frac{\partial f}{\partial x_i} = 2 \cdot \sum_j a_{ij}x_j$,

temos $\nabla f(x) = 2A \cdot x$. Por outro lado, $S^{n-1} = \varphi^{-1}(1)$, onde $\varphi(x) = \langle x, x \rangle$, donde $\nabla \varphi(x) = 2 \cdot x$. Logo, os pontos críticos de $f|_{S^{n-1}}$ são os pontos $u \in S^{n-1}$ tais que $A \cdot u = \lambda \cdot u$. Num tal ponto u , temos $f(u) = \langle Au, u \rangle = \langle \lambda u, u \rangle = \lambda$, pois $\langle u, u \rangle = 1$. Podemos, portanto, enunciar:

Proposição 154. Dada a forma quadrática $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \langle A \cdot x, x \rangle$, com $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ auto-adjunta, um ponto $u \in S^{n-1}$ é ponto crítico de $f|_{S^{n-1}}$ se, e somente se, $A \cdot u = \lambda \cdot u$, onde $\lambda = f(u)$.

Em particular, se λ_1 é o valor máximo de f no compacto S^{n-1} , atingido no ponto $u_1 \in S^{n-1}$, então λ_1 é o maior valor próprio de A . Consideremos agora o subespaço $(n - 1)$ -dimensional

$E = \{x \in \mathbb{R}^n; \langle x, u_1 \rangle = 0\}$, complemento ortogonal do vetor u_1 . Se $x \in E$ então $\langle A \cdot x, u_1 \rangle = \langle x, A \cdot u_1 \rangle = \langle x, \lambda_1 u_1 \rangle = \lambda_1 \cdot \langle x, u_1 \rangle = 0$. Logo $x \in E \Rightarrow A \cdot x \in E$. Por restrição, obtemos uma transformação linear auto-adjunta $A : E \rightarrow E$. Seja $f(u_2) = \lambda_2$ o valor máximo da forma quadrática f entre os vetores unitários pertencentes a E (isto é, perpendiculares a u_1). Então, $A \cdot u_2 = \lambda_2 \cdot u_2$. Prosseguindo analogamente, obtemos uma base ortonormal de \mathbb{R}^n , (u_1, u_2, \dots, u_n) , formada por vetores próprios de A .

Exemplo 155. Consideremos a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = x \cdot y \cdot z$, e, para cada $c > 0$, procuremos seu valor máximo na superfície $M = \{(x, y, z) \in \mathbb{R}^3; x + y + z = c, x > 0, y > 0, z > 0\}$. Esse máximo existe em \overline{M} porque \overline{M} é um conjunto compacto mas, na realidade, está em M pois f é positiva em M e nula em $\overline{M} - M$. Ora, $\text{grad } f = (yz, xz, xy)$ e M é uma parte de $\varphi^{-1}(c)$, com $\varphi(x, y, z) = x + y + z$, donde $\text{grad } \varphi = (1, 1, 1)$. Logo, num ponto de M onde $f|_M$ seja máxima devemos ter $yz = xz = xy = \lambda$ e $x + y + z = c$. Assim, $x = y = z = \frac{c}{3}$ e

$$f(x, y, z) = \frac{c^3}{27}. \text{ Sendo este valor máximo de } f \text{ em } M, \text{ devemos ter } xyz \leq \frac{c^3}{27} = \left(\frac{x+y+z}{3}\right)^3,$$

sempre que $x > 0, y > 0, z > 0$ e $x + y + z = c$. Mas c é arbitrário e três números positivos x, y, z têm sempre uma soma c . Podemos então afirmar que, dados 3 números positivos

$$x, y, z, \text{ tem-se } xyz \leq \left(\frac{x+y+z}{3}\right)^3, \text{ ou seja, } \sqrt[3]{xyz} \leq \frac{x+y+z}{3}: \text{ a média geométrica é menor}$$

do que ou igual à média aritmética. O mesmo raciocínio, aplicado à função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, dada por $f(x) = x_1 \cdot x_2 \dots x_n$, mostra que a média geométrica de n números positivos é menor do que ou igual à média aritmética desses números.

Exemplo 156. Usaremos o método do multiplicador de Lagrange para demonstrar a desigualdade de Hadamard: se X é uma matriz $n \times n$ cujas linhas são os vetores $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$ então $\det X \leq |X_1| \cdot |X_2| \dots |X_n|$, onde $|X_i|$ é a norma euclidiana de X_i . Isto é evidente se $\det X = 0$. Caso $\det X \neq 0$, então todos os vetores linha são $\neq 0$, logo, $X_i = |X_i| \cdot W_i$, com $|W_i| = 1$, para todo i . Então $\det X = |X_1| \cdot |X_2| \dots |X_n| \cdot \det W$, onde W é a matriz cujas linhas são os vetores unitários W_1, \dots, W_n . A desigualdade de Hamadard ficará provada se mostrarmos que $\det W \leq 1$. Mais geralmente, mostraremos que se $W = (w_{ij})$ é uma matriz $n \times n$ tal que $\sum_{i,j} w_{i,j}^2 = n$ então $\det W \leq 1$.

Demonstração: Definamos portanto $f, \varphi : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$ pondo $f(X) = \det X$ e $\varphi(X) = \sum_{i,j} x_{ij}^2$. Temos $\frac{\partial \varphi}{\partial x_{ij}}(X) = 2x_{ij}$ e $\frac{\partial f}{\partial x_{ij}}(X) = (-1)^{i+j} X_{[ij]}$, onde $X_{[ij]}$ é o determinante da matriz

$(n-1) \times (n-1)$, obtida de φ pela omissão da i -ésima linha e da j -ésima coluna. Todo número real $\neq 0$ é um valor regular de φ , logo $\varphi^{-1}(n)$ é uma hipersuperfície (compacta) de classe \mathcal{C}^k em \mathbb{R}^{n^2} . (Esfera de centro 0 e raio \sqrt{n} .)

Uma matriz $W = (w_{ij})$ é ponto crítico de $f|_M$ se, e somente se, $\sum_{i,j} w_{ij}^2 = n$ e $\text{grad } f(W) = \lambda \cdot \text{grad } \varphi(W)$ para algum λ real. Daí:

$$(-1)^{i+j} W_{[ij]} = 2\lambda \cdot W_{ij} \quad \text{para quaisquer } i, j \in [1, n]. \quad (4.1)$$

Multiplicando por w_{ij} , somando e levando em conta a regra de expansão de um determinante em relação aos elementos de uma linha, temos:

$$n \cdot \det. W = \sum_{i,j} (-1)^{i+j} w_{ij} W_{(i,j)} = 2\lambda \cdot \sum w_{ij}^2 = 2\lambda \cdot n,$$

donde $\det. W = 2\lambda$. Agora multipliquemos (4.1) por w_{ij} , fixemos i e somemos em relação a j . Resulta:

$$\det. W = \sum_j (-1)^{i+j} w_{ij} W_{ij} = \det. W \cdot \sum_j w_{ij}^2.$$

Suponhamos que W seja um ponto onde $f|_M$ assume seu valor máximo. Então $\det. W = f(W) \neq 0$ e da igualdade acima vem $|W_i|^2 = \sum_j w_{ij}^2 = 1$ para todo i .

Em seguida, multipliquemos (4.1) por w_{kj} , com $k \neq i$, e somemos em relação a j . Teremos

$$\sum_j (-1)^{i+j} w_{kj} W_{(i,j)} = 2\lambda \cdot \sum_j w_{kj} w_{(i,j)} = 2\lambda \cdot \langle W_k, W_i \rangle.$$

Ora, o primeiro somatório acima é zero, por ser o desenvolvimento, em relação aos elementos da i -ésima linha, do determinante de uma matriz com duas linhas (a i -ésima e a k -ésima) iguais a W_k . Logo $\langle W_k, W_i \rangle = 0$ para $k \neq i$.

Assim, todo ponto $W \in M$ onde $f|_M$ atinja o seu valor máximo é uma matriz cujas linhas são vetores unitários, 2 a 2 ortogonais, isto é, W é uma matriz ortogonal. Em particular, $\det. W = \pm 1$. Por ser $\det. W$ máximo, seu valor é evidentemente 1.

Concluimos que $\det. W \leq 1$ para toda $W \in M$, o que demonstra a desigualdade de Hadamard. ■

Observação 157. O valor absoluto de $\det. X$ é o volume do paralelepípedo n -dimensional

determinado pelos vetores X_1, \dots, X_n , que constituem as linhas da matriz X . A desigualdade de Hamadard significa, geometricamente, que se mantivermos constantes os comprimentos desses vetores, $\det. X$ se torna máximo quando eles forem 2 a 2 perpendiculares, caso em que o volume do paralelepípedo é o produto $|X_1| \cdot |X_2| \cdots |X_n|$ dos comprimentos das suas arestas.

4.4 Exemplo de uma aplicação: Ensemble Microcanônico

Nesta seção, apresentamos um exemplo de aplicação do Teorema do Multiplicador de Lagrange para n dimensões, que ocorre na Física, no contexto da Mecânica Estatística.

Nesse contexto, o método é empregado para determinar a distribuição de partículas em um sistema físico sujeito à restrições. O número total de partículas é denotado por N , e a energia total do sistema é representada por E . As partículas podem ocupar diferentes níveis de energia ε_j , onde j é um índice que identifica cada nível. Para cada nível de energia ε_j , o número de partículas que o ocupam é indicado por N_j . Assim, as restrições do sistema são expressas como

$$\sum_j N_j = N,$$

garantindo a conservação do número total de partículas, e

$$\sum_j N_j \varepsilon_j = E,$$

que assegura a conservação da energia total.

O objetivo é maximizar o número de microestados acessíveis ao sistema, denotado por $\Omega(N_j)$, que descreve o número de diferentes configurações microscópicas em que as partículas podem ser distribuídas entre os níveis de energia, obedecendo às restrições acima. Por meio da aplicação do Teorema do Multiplicador de Lagrange, encontra-se a distribuição N_j que maximiza $\Omega(N_j)$, determinando assim a configuração mais provável do sistema em equilíbrio. Esse procedimento conecta diretamente as propriedades microscópicas do sistema à entropia.

Capítulo 5

Considerações Finais

O objetivo principal desta monografia foi apresentar o Teorema do Multiplicador de Lagrange no contexto do espaço \mathbb{R}^n . Foram apresentados conceitos fundamentais de Topologia e Análise de funções reais em várias variáveis.

Este trabalho teve como referência principal o livro [4].

Nos cinco capítulos, foram discutidas as definições e exemplos dos tópicos abordados em cada um deles. No Capítulo 1, noções de máximos e mínimos condicionados. No Capítulo 2, noções de topologia no espaço euclidiano \mathbb{R}^n , conjuntos abertos, conjuntos fechados, sequências e compactos.

No Capítulo 3, abordamos funções reais de várias (n) variáveis e caminhos diferenciáveis, com as noções de derivadas parciais e direcionais, funções diferenciáveis, gradiente e o Teorema da Função Implícita. Finalmente, apresentamos e demonstramos o Teorema do Multiplicador de Lagrange, no Capítulo 4.

Dentre os próximos passos possíveis, podemos citar o estudo da aplicação dos multiplicadores de Lagrange em outras áreas, tais como em Física, no contexto do ensemble microcanônico, citado na última seção do Capítulo 4.

Essa e outras aplicações ficam como incentivo para estudos futuros. A dedicação ao estudo de conceitos, definições e ideias de Análise no \mathbb{R}^n não apenas contribuiu para o desenvolvimento deste trabalho, mas também ampliou o entendimento do autor sobre o tema. Este trabalho se encerra com a expectativa de que seus tópicos e abordagens possam ser aprofundados em futuras investigações.

Referências Bibliográficas

- [1] APOSTOL, Tom M. *Calculus: Volume II: Multivariable Calculus and Linear Algebra, with Applications to Differential Equations and Probability*. Pasadena, California: Blaisdell Publishing Company, 1968.
- [2] FIGUEIREDO, Djairo Guedes de. *Análise I*. 2. ed. Campinas: Unicamp, 1996.
- [3] LIMA, Elon Lages. *Curso de Análise: Volume 1*. 11. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2004.
- [4] LIMA, Elon Lages. *Curso de Análise, Volume 2*. 10. ed., Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 2007.
- [5] STEWART, James. *Cálculo, Volume 2*. 6. ed., Cengage Learning, São Paulo, 2013.
- [6] Vadlamani, S. K., Xiao, T. P., & Yablonovitch, E. (2020). *Physics successfully implements Lagrange multiplier optimization*. *Proceedings of the National Academy of Sciences*, 117(43), 26639–26650. National Acad Sciences.