



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
LICENCIATURA PLENA EM MATEMÁTICA

Alexandre Santana Leite Júnior

O Teorema da Função Inversa e o Lema de Morse

Recife - PE
Agosto de 2025

O Teorema da Função Inversa e o Lema de Morse

Trabalho de conclusão de curso submetido à Coordenação do Curso de licenciatura plena em Matemática da Universidade Federal Rural de Pernambuco, como parte dos requisitos para a obtenção do grau de licenciado em matemática.

Orientador: Prof. Dr. Gilson Mamede de Carvalho

Recife - PE
Agosto de 2025

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Gerada automaticamente, mediante os dados fornecidos pelo(a)
autor(a)

J95t Júnior, Alexandre Santana Leite.
O Teorema da Função inversa e o Lema de Morse /
Alexandre Santana Leite Júnior. – Recife, 2025.
60 f.

Orientador(a): Gilson Mamede de Carvalho.

Trabalho de Conclusão de Curso (Graduação) –
Universidade Federal Rural de Pernambuco, Licenciatura em
Matemática, Recife, BR-PE, 2025.

Inclui referências.

1. Teorema da Função Inversa. 2. Lema de Morse. 3.
Teorema de Schwarz. I. Carvalho, Gilson Mamede de, orient.
II. Título

CDD 510

Alexandre Santana Leite Júnior

O Teorema da Função Inversa e o Lema de Morse

Trabalho de conclusão de curso submetido à Coordenação do Curso de licenciatura plena em Matemática da Universidade Federal Rural de Pernambuco, como parte dos requisitos para a obtenção do grau de licenciado em matemática.

Aprovado em: 07/08/2025.

BANCA EXAMINADORA

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

Prof. Dr. Eberson Ferreira da Silva
Universidade Federal Rural de Pernambuco - UFRPE

Prof^ª. Dr^ª. Yane Lísley Ramos Araújo
Universidade Federal Rural de Pernambuco - UFRPE

Recife - PE
Agosto de 2025

Dedico este trabalho aos meus pais, que tanto fizeram e fazem por mim.

Agradecimentos

Agradeço primeiramente a Deus, por me conceder força, saúde e resiliência ao longo de toda a jornada.

À minha família, em especial aos meus pais, Alexandre Santana e Gilvânia Gomes, pelo apoio incondicional, emocional, financeiro e, sobretudo, por sempre acreditarem em mim. Amo vocês profundamente, e tenho plena certeza de que sem o amparo de vocês, tudo teria sido muito mais difícil. Ao meu irmão Alexsandro, pelo suporte constante durante essa caminhada. E ao meu avô Mizael Gomes Leão (vô Miza), por me ensinar a lição mais importante da vida: viver. Este trabalho também é de vocês.

Ao meu orientador, professor Gilson Mamede, pela dedicação, paciência, companheirismo e pelos valiosos ensinamentos ao longo da graduação. O senhor é, sem dúvida, um exemplo de profissional a ser seguido.

Aos professores do Departamento de Matemática da Universidade Federal Rural de Pernambuco, com os quais tive a honra e o privilégio de aprender tanto. Em especial, à professora Yane Araújo, pelo apoio, incentivo e ensinamentos; ao professor Thiago Tanaka, por acreditar em mim e me motivar a apresentar trabalhos em eventos acadêmicos; ao professor Thiago Dias (DK), por sua confiança, apoio constante e por sempre lembrar de mim nas confraternizações; ao professor Eberson Ferreira, por suas aulas de excelência e pela forma tão inspiradora de ensinar matemática; à professora Maité, pela orientação durante a monitoria, experiência que foi extremamente enriquecedora.

Estendo ainda minha gratidão aos demais professores do departamento, Filipe, Mendonça, Ross, Ângela, Clemente, Thamires, Bárbara, Jogli, Gilson Simões, Hinojosa, Lorena, Edgar, Gondão e Duque pelos ensinamentos, pela dedicação e pelo compromisso com a formação acadêmica.

Agradeço aos meus colegas, amigos e irmãos que o curso me presenteou. Em especial, à Jaqueline Mayara, pelo companheirismo, amizade e por estar sempre presente

nos momentos bons e ruins, tornando cada experiência mais leve e significativa. Obrigado por ser tão você. À Sophia Evelin, pela amizade sincera, pelas inúmeras conversas e pelos momentos únicos que tive a honra de compartilhar ao seu lado. À Beatriz Bento, pelo carinho, pela amizade, pela alegria contagiante e por sempre trazer palavras de incentivo quando mais precisei. A Lucas Kauan, pela amizade, pela ajuda sempre generosa e pela disposição em abrir as portas de sua casa para que pudéssemos compartilhar tantos momentos.

Agradeço também a Mateus Gomes e Natanael Oliveira, pelos momentos de descontração e pelos estudos que foram de grande importância nessa trajetória. A Carlos Bloise, que foi muito além de um amigo, tornando-se um verdadeiro irmão. A Emmanuel, pela disposição constante em ajudar, e a Hilário, pelas conversas profundas que sempre acrescentaram muito.

Por fim, estendo minha gratidão aos demais amigos e colegas da rural: Lívia, Laryssa, Tainá, Luiz Japonês, Emanuele, Marcos, Kewellen, Pedro Borges, Fábio, Paulo, Warsdley, Maria Júlia, Lucas (Jesus), Maria Laura, Karina, Thais, Myllene, Fabrício, Tom, Eduarda, Delleon Bar por ser um lugar de acolhimento e descontração e tantos outros que, de diferentes formas, fizeram parte dessa caminhada e contribuíram para tornar essa jornada inesquecível.

Resumo

Neste trabalho estudamos o Teorema da Função Inversa, uma das ferramentas centrais da análise matemática, e o utilizamos como base para apresentar e demonstrar o Lema de Morse. Para fundamentar esses resultados, desenvolvemos previamente os conceitos necessários, referentes à topologia do \mathbb{R}^n , a continuidade e a diferenciabilidade de funções, com ênfase nos pontos críticos, no gradiente e no Teorema de Schwarz. Em seguida, demonstramos o Teorema da Função Inversa, o qual garante a existência de difeomorfismos locais sob condições adequadas, e o utilizamos para construir uma demonstração rigorosa do Lema de Morse. Esse Lema mostra que funções de classe C^k ($k \geq 3$), em vizinhanças de pontos críticos não degenerados, podem ser localmente expressas como formas quadráticas por meio de mudanças de coordenadas de classe C^{k-2} .

Palavras-Chave: Teorema da Função Inversa, Lema de Morse, Teorema de Schwarz.

Abstract

In this work, we study the Inverse Function Theorem, one of the central tools of mathematical analysis, and use it as a basis to present and prove Morse's Lemma. To ground these results, we first develop the necessary concepts regarding the topology of \mathbb{R}^n , continuity, and differentiability of functions, with emphasis on critical points, the gradient, and Schwarz's Theorem. Subsequently, we prove the Inverse Function Theorem, which guarantees the existence of local diffeomorphisms under appropriate conditions, and employ it to construct a rigorous proof of Morse's Lemma. This Lemma shows that functions of class C^k ($k \geq 3$), in neighborhoods of non-degenerate critical points, can be locally expressed as quadratic forms via coordinate changes of class C^{k-2} .

KeyWords: Inverse Function Theorem, Morse Lemma, Schwarz's Theorem.

Sumário

Introdução	11
1 Conceitos e Resultados Auxiliares	12
1.1 Topologia do espaço \mathbb{R}^n	12
1.2 Aplicações contínuas	18
1.3 Funções diferenciáveis	23
1.4 Teorema de Schwarz	27
2 O Teorema da Função Inversa	38
3 O Lema da Morse	45
Referências Bibliográficas	60

Introdução

Marston Morse (1892–1977) foi um matemático americano cuja contribuição mais impactante foi o desenvolvimento da Teoria dos Pontos Críticos. Essa teoria estabelece uma relação profunda entre a topologia de uma variedade e o comportamento analítico de funções reais definidas sobre ela, ligando pontos críticos (classificados por um índice) a invariantes topológicos como os números de Betti por meio das Desigualdades de Morse. Seu trabalho, laureado com o Prêmio Bôcher (1933), é fundamental para áreas como geometria diferencial e cálculo variacional, e serve de base para resultados como o Lema de Morse, que explora a estrutura local de funções nas vizinhanças de pontos críticos não degenerados.

Um dos resultados de grande importância no contexto da análise em várias variáveis é o Teorema da Função Inversa. Seu desenvolvimento remonta a trabalhos de Joseph-Louis Lagrange no século XVIII, embora sua formulação moderna seja atribuída ao rigor matemático do século XIX. O Teorema fornece condições precisas sob as quais uma aplicação diferenciável entre espaços euclidianos admite uma inversa diferenciável em vizinhanças adequadas. Ele formaliza a ideia de que, localmente, uma função suficientemente regular pode ser “invertida”, desde que a matriz jacobiana no ponto seja invertível. O Teorema da Função Inversa é essencial não só em análise, mas também nas áreas de geometria diferencial e sistemas dinâmicos, por permitir a construção de coordenadas locais e a mudança de variáveis em contextos complexos.

Nosso objetivo principal neste trabalho é demonstrar o Lema de Morse e apresentar aplicações deste resultado. Para tanto, no primeiro capítulo definiremos conceitos fundamentais da topologia de \mathbb{R}^n , como abertos, fechados e compactos, bem como sequências, limites, continuidade e diferenciabilidade. Também apresentamos e demonstramos o Teorema de Schwarz que garante, sob certas condições, que as derivadas mistas são iguais. Finalizamos esse capítulo apresentando a matriz hessiana e alguns teoremas e resultados

que utilizamos ao longo do trabalho.

No segundo capítulo, traremos alguns resultados auxiliares os quais usamos na demonstração do Teorema da Função Inversa, que estabelecerá condições para que uma função diferenciável admita uma inversa localmente diferenciável.

E por fim, apresentamos e demonstramos o Lema de Morse. Segundo este lema, em uma vizinhança de um ponto crítico não degenerado de uma função f , será possível obter um sistema de coordenadas no qual f se expressará como uma forma quadrática com coeficientes constantes $f(y) = \sum_{i,j=1}^n a_{ij}y_iy_j$.

Capítulo 1

Conceitos e Resultados Auxiliares

Neste capítulo, apresentaremos algumas noções topológicas fundamentais do espaço euclidiano n -dimensional, como os conceitos de bola aberta e fechada, de conjuntos abertos, fechados e compactos, sequências, limite de funções reais de n variáveis e diferenciabilidade. Além disso, discutimos resultados clássicos, incluindo o Teorema de Schwarz. Esses conceitos fornecerão a base teórica necessária para a compreensão e a demonstração do Teorema da Aplicação Inversa e do Lema de Morse.

1.1 Topologia do espaço \mathbb{R}^n

Nesta seção apresentamos o espaço normado \mathbb{R}^n , a definição de produto interno, bem como suas principais estruturas topológicas e algumas de suas propriedades necessárias para o desenvolvimento deste trabalho. Para mais detalhes veja [4] e [6].

Definição 1.1. O espaço normado \mathbb{R}^n é o par $(\mathbb{R}^n, \|\cdot\|)$ em que

$$\mathbb{R}^n = \{(x_1, \dots, x_n); x_i \in \mathbb{R} \text{ para todo } i \in \mathbb{N} \text{ com } 1 \leq i \leq n\}$$

e $\|x\|$ é a norma euclidiana dada por

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Em algumas situações vamos precisar de outras normas no espaço \mathbb{R}^n , então relembremos a definição do que vem a ser uma norma em \mathbb{R}^n .

Definição 1.2. Uma norma em \mathbb{R}^n é uma função $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfazendo, para quaisquer $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$ as seguintes condições:

1. $\|x\| \geq 0$, valendo $\|x\| = 0$ apenas quando $x = 0$;
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$;
3. (Desigualdade triangular): $\|x + y\| \leq \|x\| + \|y\|$

Além da norma euclidiana, já apresentada, existem outras duas normas importantes que também podem ser utilizadas no espaço \mathbb{R}^n , conhecidas como norma do máximo e norma da soma e são elas:

1. Norma do máximo: $\|x\|_M = \max\{|x_1|, |x_2|, \dots, |x_n|\}$;
2. Norma da soma: $\|x\|_S = |x_1| + |x_2| + \dots + |x_n|$.

Observação 1.3. Vale a seguinte relação $\|x\|_M \leq \|x\| \leq \|x\|_S \leq n \cdot \|x\|_M$.

Observação 1.4. A norma euclidiana, do máximo e da soma são equivalentes; podemos considerar o espaço \mathbb{R}^n munido com qualquer uma das três normas.

Definição 1.5. Um produto interno no \mathbb{R}^n é uma aplicação bilinear

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

a qual satisfaz as seguintes propriedades para todos $x, y, z \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$:

1. $\langle x, x \rangle \geq 0$, valendo $\langle x, x \rangle = 0$ apenas quando $x = 0$.
2. $\langle x, y \rangle = \langle y, x \rangle$
3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
4. $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$

Observação 1.6. Temos que $\|x\| = \sqrt{\langle x, x \rangle}$.

A seguir vamos definir bolas, aberta e fechada, elementos importantes para definir o que são conjuntos abertos, fechados e compactos em \mathbb{R}^n .

Definição 1.7 (Bola aberta). Sejam $a \in \mathbb{R}^n$ e o número real $r > 0$, a bola aberta de centro a e raio r é o conjunto $B(a, r)$ dos pontos $x \in \mathbb{R}^n$ cuja distância ao ponto a é menor que r . Isto é

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

Definição 1.8 (Bola fechada). Sejam $a \in \mathbb{R}^n$ e o número real $r > 0$, a bola fechada de centro a e raio r é o conjunto $B[a, r]$ dos pontos $x \in \mathbb{R}^n$ cuja distância ao ponto a é menor ou igual a r . Ou seja

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

Definição 1.9 (Conjunto Limitado). Seja $X \subset \mathbb{R}^n$. Dizemos que X é limitado quando $X \subset B[a, r]$, para algum $a \in \mathbb{R}^n$ e algum $r > 0$. É possível observar que $X \subset \mathbb{R}^n$ é limitado se, e somente se, existir $k > 0$ tal que $\|x\| \leq k$ para todo $x \in X$.

Vejamos agora a definição de ponto interior a qual será usada para definir conjunto aberto.

Definição 1.10 (Ponto interior). Seja $X \subset \mathbb{R}^n$ não-vazio. Dizemos que $a \in X$ é um ponto interior a X se existir $r > 0$ tal que $B(a, r) \subset X$. O interior de X é um conjunto no qual denotamos $\text{int}.X$ e, claramente, $\text{int}.X \subseteq X$.

Definição 1.11 (Conjunto aberto). Seja $X \subset \mathbb{R}^n$. X é um conjunto aberto quando todos os seus pontos são interiores, ou seja, $\text{int}.X = X$.

Exemplo 1.12. Toda bola aberta $B = B(a, r)$ é um conjunto aberto

De fato, seja $x \in B$, então $\|x - a\| < r \Rightarrow r - \|x - a\| > 0$. Tome $s = r - \|x - a\|$ e considere $y \in B(x, s)$.

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

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

Definição 1.13 (Aberto relativo). Sejam $X \subset \mathbb{R}^n$ e $A \subset X$. Dizemos que o subconjunto A é aberto em X quando cada ponto $a \in A$ é centro de uma bola aberta $B(a; r)$, tal que

$B(a, r) \cap X \subset A$. Ou seja, os pontos de X que estão suficientemente próximos de cada $a \in A$, pertencem a A também.

Observação 1.14. Um conjunto $A \subset X$ é aberto em X se, e somente se, $A = U \cap X$, onde U é aberto em \mathbb{R}^n .

Exemplo 1.15. Seja $(\frac{1}{2}, 1]$. Note que $(\frac{1}{2}, 1] = (\frac{1}{2}, 2) \cap [0, 1]$. Logo, $(\frac{1}{2}, 1]$ é aberto em $[0, 1]$.

Em seguida, definimos o que são sequências em \mathbb{R}^n , bem como apresentamos as definições de subsequência, convergência e o teorema de Bolzano-Weierstrass. Esses conceitos e resultados serão fundamentais para o desenvolvimento de resultados posteriores.

Definição 1.16 (Sequência em \mathbb{R}^n). Definimos uma sequência em \mathbb{R}^n como uma aplicação $x : \mathbb{N} \rightarrow \mathbb{R}^n$, que associa a cada número natural k , um ponto $x_k \in \mathbb{R}^n$.

Notação: $(x_1, x_2, \dots, x_k, \dots)$, $(x_k)_{k \in \mathbb{N}}$ ou (x_k) .

Seja $(x_k)_{k \in \mathbb{N}}$ uma sequência em \mathbb{R}^n , com $x_k = (x_{k_1}, \dots, x_{k_n})$. Note que essa sequência gera n sequências reais (uma por componente). A sequência é limitada se existe $c > 0$ tal que $\|x_k\| \leq c$ para todo $k \in \mathbb{N}$, ou equivalentemente, se todos os seus termos estão contidos em alguma bola de \mathbb{R}^n .

Proposição 1.17. Seja $(x_k)_{k \in \mathbb{N}}$ uma sequência em \mathbb{R}^n , com $x_k = (x_{k_1}, \dots, x_{k_n})$. A sequência (x_k) é limitada se e somente se, para cada $i = 1, \dots, n$, a sequência de coordenadas $(x_{k_i})_{k \in \mathbb{N}}$ é limitada.

Demonstração: (\Rightarrow) Suponha (x_k) limitada. Então existe $c > 0$ tal que $\|x_k\| \leq c$ para todo k . Como $|x_{k_i}| \leq \|x_k\|$ para cada coordenada i , temos $|x_{k_i}| \leq c$, logo cada sequência de coordenadas é limitada.

(\Leftarrow) Suponha cada (x_{k_i}) limitada. Então existem constantes $c_i > 0$ tais que $|x_{k_i}| \leq c_i$ para todo k e cada i . Adotando a norma do máximo em \mathbb{R}^n , definimos $c = \max\{c_1, \dots, c_n\}$.

Para todo k vale:

$$\|x_k\|_M = \max\{|x_{k_1}|, \dots, |x_{k_n}|\} \leq c$$

Portanto, (x_k) é limitada. ■

Definição 1.18 (Subsequências em \mathbb{R}^n). Uma subsequência de $(x_k)_{k \in \mathbb{N}}$ é a restrição da sequência a um subconjunto infinito $\mathbb{N}' = \{k_1 < k_2 < \dots < k_m < \dots\}$.

Notações: $(x_k)_{k \in \mathbb{N}'}$, $(x_{k_m})_{m \in \mathbb{N}}$ ou $(x_{k_1}, \dots, x_{k_m}, \dots)$.

Definição 1.19 (Sequência convergente). Seja $a \in \mathbb{R}^n$. Dizemos que a é limite da sequência (x_k) se, para todo $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \implies \|x_k - a\| < \varepsilon$. Isso significa que $x_k \in B(a; \varepsilon)$ para $k > k_0$. A sequência (x_k) é convergente se existe $a \in \mathbb{R}^n$ com $\lim x_k = a$. Notações: $\lim_{k \rightarrow \infty} x_k = a$, $\lim x_k = a$ ou $x_k \rightarrow a$.

Teorema 1.20. A sequência (x_k) em \mathbb{R}^n converge para o ponto $a = (a_1, \dots, a_n)$ se, e somente se, $\forall i = 1, \dots, n$, tivermos que $\lim_{k \rightarrow \infty} x_{k_i} = a_i$. Ou seja, cada coordenada de x_k converge para a coordenada correspondente de a .

Demonstração: (\implies) Como $|x_{k_i}| \leq \|x_k\| \forall i = 1, \dots, n$, temos $|x_{k_i} - a_i| \leq \|x_k - a\|$, assim, se $\lim x_k = a$ então, $\lim_{k \rightarrow \infty} x_{k_i} = a_i$.

(\impliedby) Se $\lim x_{k_i} = a_i$, então dado $\varepsilon > 0$, existem $k_1, \dots, k_n \in \mathbb{N}$ tais que $k > k_i \implies |x_{k_i} - a_i| < \varepsilon; i = 1, \dots, n$. Seja $k_0 = \max\{k_1, \dots, k_n\}$, adotando a norma do máximo, temos que $k > k_0 \implies \|x_k - a\| < \varepsilon$, logo, $\lim x_k = a$. ■

A seguir, será apresentado um dos resultados centrais sobre sequências em \mathbb{R}^n o Teorema de Bolzano-Weierstrass.

Teorema 1.21 (Bolzano-Weierstrass). Toda sequência limitada em \mathbb{R}^n possui uma sub-sequência convergente.

Demonstração: Seja (x_k) uma sequência limitada em \mathbb{R}^n . Logo, existe $c > 0$ tal que $\|x_k\| < c$ para todo $k \in \mathbb{N}$. Além disso, $\|x_{k_i}\| \leq \|x_k\| < c \forall k \in \mathbb{N}$ e $\forall i = 1, \dots, n$. Logo, $(x_{k_i}) \subset \mathbb{R}$ é uma sequência limitada, então pelo teorema de Bolzano-Weierstrass na reta, $\exists \mathbb{N}_1 \subset \mathbb{N}$ infinito e $a_1 \in \mathbb{R}$ de modo que $\lim_{k \in \mathbb{N}_1} x_{k_1} = a_1$. Além disso, a sequência $(x_k)_{k \in \mathbb{N}_1}$ é limitada em \mathbb{R} , portanto existem $\mathbb{N}_2 \subset \mathbb{N}_1$ infinito e $a_2 \in \mathbb{R}$ tal que $\lim_{K \in \mathbb{N}_2} x_{k_2} = a_2$. Prosseguindo desta forma de modo que se obtenha n conjuntos $\mathbb{N} \supset \mathbb{N}_1 \supset \dots \supset \mathbb{N}_n$ e números reais a_1, \dots, a_n tais que $\lim_{k \in \mathbb{N}_i} x_{k_i} = a_i, \forall i = 1, \dots, n$. Considere $a = (a_1, \dots, a_n)$, então pelo Teorema 1.20 como $\lim_{k \in \mathbb{N}_i} x_{k_i} = a_i$, temos que $\lim_{k \in \mathbb{N}_n} x_k = a$. ■

De posse da definição de sequência convergente, podemos definir ponto aderente, o qual é um conceito importante para definir conjunto fechado em \mathbb{R}^n .

Definição 1.22 (Ponto aderente). Sejam $X \subset \mathbb{R}^n$ e $a \in \mathbb{R}^n$. Dizemos que a é ponto aderente a X se existe uma sequência $(x_k) \subset X$ com $\lim x_k = a$.

Definição 1.23 (Fecho). O fecho de $X \subset \mathbb{R}^n$, denotado por \overline{X} , é o conjunto de todos os pontos aderentes a X .

Definição 1.24 (Conjunto fechado). Um conjunto $F \subset \mathbb{R}^n$ é fechado se $F = \overline{F}$. Equivalentemente, F é fechado se o limite de toda sequência convergente de pontos de F pertence a F .

Observação 1.25. Um conjunto é fechado se seu complementar é aberto.

Exemplo 1.26. Toda bola fechada é um conjunto fechado.

Vamos utilizar o fato de que um conjunto F é fechado se, e somente se, seu complementar é aberto. Considere a bola fechada $B = B[a, r]$ cujo o complementar é

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

Basta mostra que A é aberto.

Seja $x_0 \in A$, então $\|x_0 - a\| > r$. Defina $\varepsilon = \|x_0 - a\| - r > 0$. Nosso objetivo agora é mostrar que existe $B(x_0, \frac{\varepsilon}{2}) \subset A$.

De fato, se $y \in B(x_0, \frac{\varepsilon}{2}) \Rightarrow \|y - x_0\| < \frac{\varepsilon}{2}$. Pela desigualdade triangular

$$\|x_0 - a\| \leq \|x_0 - y\| + \|y - a\| \Rightarrow \|y - a\| \geq \|x_0 - a\| - \|x_0 - y\|.$$

Assim,

$$\|y - a\| \geq \|x_0 - a\| - \frac{\varepsilon}{2}.$$

Substituindo $\varepsilon = \|x_0 - a\| - r$

$$\|y - a\| \geq \|x_0 - a\| - \frac{\|x_0 - a\| - r}{2} = \frac{\|x_0 - a\| + r}{2} > r,$$

pois $\|x_0 - a\| > r$. Portanto $y \in A$. Logo $B(x_0, \frac{\varepsilon}{2}) \subset A$.

Sabendo o que são conjuntos limitados e fechados, podemos definir o que é um conjunto compacto.

Definição 1.27 (Compacto). Um conjunto $X \subset \mathbb{R}^n$ é dito ser compacto quando for fechado e limitado.

Exemplo 1.28. Toda bola fechada $B[a, r]$ é compacta.

Claramente, sabemos que $B[a, r]$ é fechada pelo exemplo 1.26. Basta mostrar que esse conjunto é limitado.

Note que se $x \in B[a, r] \Rightarrow \|x - a\| \leq r$. Além disso, $\|x\| = \|x - a + a\| \leq r + \|a\|$.
Mostrando que é limitada.

Teorema 1.29. *As seguintes afirmações sobre o conjunto $K \subset \mathbb{R}^n$ são equivalentes:*

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

Demonstração: ($i \Rightarrow ii$) Como K é compacto, em particular, é limitado. Então, pelo Teorema de Bolzano-Weierstrass, existe uma subsequência $(x_k)_{k \in \mathbb{N}'}$ de (x_k) que converge para algum $a = \lim_{k \in \mathbb{N}'} x_k$. Como K é fechado, temos que $a \in K$.

($ii \Rightarrow i$) Mostraremos que K é limitado e, posteriormente, fechado, assim concluiremos que K é compacto.

Suponha, por contradição, que K não é limitado. Dessa forma, para cada $k \in \mathbb{N}$ existe $x_k \in K$ tal que $\|x_k\| > k$. Sendo assim, toda subsequência de (x_k) é também ilimitada. Logo, (x_k) não admite subsequência convergente, o que é uma contradição. Logo, K é limitado.

K é fechado, pois se $\lim x_k = a$, onde $x_k \in K, \forall k \in \mathbb{N}$, então por ii), uma subsequência de (x_k) convergiria para um ponto de K , mas toda subsequência de (x_k) converge para a . Logo $a \in K$. Portanto, K é compacto. ■

Definição 1.30 (Ponto de acumulação). Seja $X \subset \mathbb{R}^n$. Um ponto $a \in \mathbb{R}^n$ chama-se ponto de acumulação do conjunto X quando toda bola centrada em a tem algum ponto de X diferente de a . Ou ainda, para todo $\epsilon > 0$, existe $x \in X$ tal que $0 < \|x - a\| < \epsilon$. O conjunto dos pontos de acumulação de X será representado por X' .

1.2 Aplicações contínuas

Nesta seção, traremos definições de limite e continuidade de aplicações definidas em \mathbb{R}^n .

Definição 1.31 (Limite). Sejam $f: X \rightarrow \mathbb{R}^m$ uma aplicação definida num conjunto $X \subset \mathbb{R}^n$ e $a \in \mathbb{R}^n$ um ponto de acumulação de X . Diz-se que o ponto $b \in \mathbb{R}^m$ é o limite de $f(x)$ quando x tende para a , e escreve-se

$$b = \lim_{x \rightarrow a} f(x),$$

quando para todo $\varepsilon > 0$ dado arbitrariamente, existir um $\delta > 0$ tal que

$$x \in X, \quad 0 < |x - a| < \delta \quad \Rightarrow \quad |f(x) - b| < \varepsilon.$$

Teorema 1.32. *Sejam a um ponto de acumulação do conjunto X , $f : X \rightarrow \mathbb{R}^n$ uma aplicação e $f_1, f_2, \dots, f_n : X \rightarrow \mathbb{R}$ as funções-coordenadas de f . Então*

$$\lim_{x \rightarrow a} f(x) = b = (b_1, b_2, \dots, b_n)$$

se, e somente se,

$$\lim_{x \rightarrow a} f_i(x) = b_i \quad \text{para cada } i = 1, \dots, n.$$

Demonstração: (\Rightarrow) Suponha que $\lim_{x \rightarrow a} f(x) = b$. Fixe $i \in \{1, \dots, n\}$ e seja $\varepsilon > 0$ arbitrário.

Como $\lim_{x \rightarrow a} f(x) = b$, existe $\delta > 0$ tal que para todo $x \in X$ com $0 < \|x - a\| < \delta$, vale

$$\|f(x) - b\| < \varepsilon.$$

Pela propriedade da norma euclidiana, temos

$$|f_i(x) - b_i| \leq \sqrt{\sum_{j=1}^n (f_j(x) - b_j)^2} = \|f(x) - b\|.$$

Logo,

$$|f_i(x) - b_i| \leq \|f(x) - b\| < \varepsilon.$$

Portanto, $\lim_{x \rightarrow a} f_i(x) = b_i$ para cada $i = 1, \dots, n$.

(\Leftarrow) Suponha que $\lim_{x \rightarrow a} f_i(x) = b_i$ para cada $i = 1, \dots, n$. Seja $\varepsilon > 0$ arbitrário. Para cada i , existe $\delta_i > 0$ tal que se $x \in X$ e $0 < \|x - a\| < \delta_i$, então

$$|f_i(x) - b_i| < \frac{\varepsilon}{n}.$$

Tome $\delta = \min\{\delta_1, \dots, \delta_n\} > 0$. Se $x \in X$ e $0 < \|x - a\| < \delta$, então para todo i vale

$$|f_i(x) - b_i| < \frac{\varepsilon}{n}.$$

Pela desigualdade entre normas,

$$\|f(x) - b\| \leq \sum_{j=1}^n |f_j(x) - b_j| < \sum_{j=1}^n \frac{\varepsilon}{n} = \varepsilon.$$

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

Definição 1.33 (Continuidade). Seja $a \in X$, dizemos que f é contínua no ponto a quando, para todo $\varepsilon > 0$ arbitrariamente dado, pode-se obter $\delta > 0$ tal que $x \in X$, $\|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon$. De outra forma, para cada bola $B(f(a); \varepsilon)$ dada, existe uma bola $B(a; \delta)$ tal que $f(B(a; \delta) \cap X) \subset B(f(a); \varepsilon)$. Dizemos que $f : X \rightarrow \mathbb{R}^n$ é uma aplicação contínua no conjunto X , quando f é contínua em todos os pontos $a \in X$.

Proposição 1.34. *Sejam $X \subset \mathbb{R}^n$ e $f : X \rightarrow \mathbb{R}^n$ uma função. f é contínua em $a \in X$ se, e somente se, para toda sequência de pontos $(x_k) \subset X$ com $x_k \rightarrow a$ tem-se que $f(x_k) \rightarrow f(a)$.*

Demonstração: (\Rightarrow) Como f é contínua em $a \in X$, temos que para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $\|f(x) - f(a)\| < \varepsilon$ sempre que $\|x - a\| < \delta$.

Seja $(x_k) \subset X$ uma sequência de tal forma que $x_k \rightarrow a \in X$. Assim, para todo $\delta > 0$, existe $k_0 \in \mathbb{N}$ de modo que $\|x_k - a\| < \delta$ para todo $k \geq k_0$. Então $\|f(x_k) - f(a)\| < \varepsilon$ para todo $k \geq k_0$, ou seja $f(x_k) \rightarrow f(a)$.

(\Leftarrow) Suponha, por contradição, que f não é contínua em $a \in X$. Assim, existe $\varepsilon > 0$ tal que para todo $k \in \mathbb{N}$ existe $x_k \in X$ tal que $\|x_k - a\| < \frac{1}{k}$ e $\|f(x_k) - f(a)\| \geq \varepsilon$. Daí temos que $x_k \rightarrow a$ mas $f(x_k)$ não converge para $f(a)$, o que é um absurdo. ■

Teorema 1.35. *A imagem $f(K)$ do conjunto compacto $K \subset X$ pela aplicação contínua $f : X \rightarrow \mathbb{R}^n$ é também um conjunto compacto.*

Demonstração: Seja (y_k) uma sequência de pontos em $f(K)$. Para todo $k \in \mathbb{N}$, existe $x_k \in K$ tal que $f(x_k) = y_k$. Como K é compacto, pelo Teorema 1.29, existe uma subsequência de (x_k) , tal que $(x_k)_{k \in \mathbb{N}'}$ que converge para o ponto $a \in K$. Como f é uma aplicação contínua, em particular, f é contínua no ponto a . Pela Proposição 1.34

$\lim_{k \in \mathbb{N}'} f(x_k) = f(a)$. Portanto, toda sequência de pontos $y_k = f(x_k) \in f(K)$ possui uma subsequência $(y_k)_{k \in \mathbb{N}'}$ que converge para um ponto $f(a) \in f(K)$. Logo, $f(K)$ é compacto.

■

Corolário 1.36 (Weierstrass). *Seja $K \subset \mathbb{R}^m$ compacto. Se $f : K \rightarrow \mathbb{R}$ é contínua, então existem $x_0, x_1 \in K$ tais que $f(x_0) \leq f(x) \leq f(x_1)$ para todo $x \in K$. Em outras palavras, toda função real contínua definida num conjunto compacto K , atinge seus valores máximos e mínimos em pontos de K .*

Demonstração: Pelo Teorema 1.35, como f é contínua e K é compacto, a imagem $f(K) \subset \mathbb{R}$ é compacta. Como $f(K)$ é compacto em \mathbb{R} , ele é fechado e limitado. Sendo limitado, existem $m = \inf\{f(x) : x \in K\}$ e $M = \sup\{f(x) : x \in K\}$. Como $f(K)$ é fechado, temos que $m, M \in f(K)$. Portanto, existem $x_0, x_1 \in K$ tais que $f(x_0) = m$ e $f(x_1) = M$. Segue que para todo $x \in K$,

$$f(x_0) = m \leq f(x) \leq M = f(x_1),$$

como queríamos demonstrar. ■

Teorema 1.37. *Seja $X \subset \mathbb{R}^m$. A aplicação $f : X \rightarrow \mathbb{R}^n$ é contínua se, e somente se, a imagem inversa $f^{-1}(A)$ de todo conjunto aberto $A \subset \mathbb{R}^n$ é um subconjunto aberto em X .*

Demonstração: (\Rightarrow) Seja f contínua e $A \subset \mathbb{R}^n$ aberto. Então para todo $x \in f^{-1}(A)$, existe $\varepsilon > 0$ tal que $B(f(x), \varepsilon) \subset A$. Pela continuidade de f , x é centro de uma bola aberta B_x , tal que $f(B_x \cap X) \subset B(f(x), \varepsilon) \subset A$. Assim, $x \in B_x \cap X \subset f^{-1}(A)$. Considere $U = \bigcup B_x$, daí para todo $x \in f^{-1}(A)$, temos que $f^{-1}(A) \subset (U \cap X) \subset f^{-1}(A)$. Logo, $f^{-1}(A) = U \cap X$.

(\Leftarrow) Suponha que para todo $A \subset \mathbb{R}^n$ aberto, $f^{-1}(A)$ é aberto em X , ou seja, $f^{-1}(A) = M \cap X$, onde M é aberto em \mathbb{R}^m . Assim, dados $x \in X$ e $\varepsilon > 0$, tomamos $A = B(f(x), \varepsilon)$ e obtemos $M \subset \mathbb{R}^m$ aberto tal que $M \cap X = f^{-1}(A)$. Certamente $x \in M$, então existe $\delta > 0$ tal que $B(x; \delta) \subset M$ e assim, $f(B(x, \delta) \cap X) \subset B(f(x), \varepsilon)$. Logo, f é contínua em todos os pontos $x \in X$. ■

Teorema 1.38. *Seja $X \subset \mathbb{R}^m$. A aplicação $f : X \rightarrow \mathbb{R}^n$ é contínua se, e somente se, a imagem inversa $f^{-1}(F)$ de todo conjunto fechado $F \subset \mathbb{R}^n$ é um subconjunto fechado em X .*

Demonstração: (\Rightarrow) Seja $F \subseteq \mathbb{R}^n$ um conjunto fechado. Como F é fechado, $A = \mathbb{R}^n \setminus F$ é aberto. Pelo Teorema 1.37 $f^{-1}(A) = f^{-1}(\mathbb{R}^n \setminus F) = f^{-1}(\mathbb{R}^n) \setminus f^{-1}(F) = X \setminus f^{-1}(F)$ é aberto em X . Portanto, como o complementar é aberto, então $f^{-1}(F)$ é fechado.

(\Leftarrow) Seja $A \subset \mathbb{R}^n$ um conjunto aberto. Note que $F = \mathbb{R}^n \setminus A$ é fechado. Como $f^{-1}(F)$ é fechado em X . Mas, $f^{-1}(F) = f^{-1}(\mathbb{R}^n \setminus A) = X \setminus f^{-1}(A)$ é fechado em X . Portanto, como o complementar é fechado, então $f^{-1}(A)$ é aberto e pelo Teorema 1.37, f é contínua. ■

Definição 1.39 (Continuidade uniforme). Uma aplicação $f : X \rightarrow \mathbb{R}^n$ é uniformemente contínua em $X \subset \mathbb{R}^m$ se, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que:

$$\|x - y\| < \delta \implies \|f(x) - f(y)\| < \varepsilon$$

para quaisquer $x, y \in X$.

Diferentemente do caso da continuidade comum, aqui o número δ não depende do ponto do domínio, mas apenas do valor de ε dado. Em geral, uma função pode ser contínua em todos os pontos de X sem ser uniformemente contínua, pois o mesmo δ pode não servir para todos os pontos simultaneamente.

O teorema que caracteriza as aplicações uniformemente contínuas por meio de seqüências será um resultado utilizado na demonstração do Teorema de Leibniz. A seguir, enunciaremos o teorema que caracteriza as aplicações continuamente uniformes e apresentamos sua respectiva demonstração.

Teorema 1.40. *Para que $f : X \rightarrow \mathbb{R}^n$ seja uniformemente contínua no conjunto $X \subset \mathbb{R}^m$, é necessário e suficiente que, para todo par de seqüências $(x_k), (y_k)$ em X com $\lim \|x_k - y_k\| = 0$, se tenha $\lim \|f(x_k) - f(y_k)\| = 0$.*

Demonstração: (\Rightarrow) Seja $f : X \rightarrow \mathbb{R}^n$ uniformemente contínua em $X \subset \mathbb{R}^n$. Então, dado $\varepsilon > 0$, existe $\delta > 0$ tal que $\|x - y\| \leq \delta \implies \|f(x) - f(y)\| \leq \varepsilon$ para quaisquer $x, y \in X$. Seja $(x_k), (y_k) \subset X$ de modo que $\lim_{k \rightarrow \infty} \|x_k - y_k\| = 0$. Pela definição de limite, para todo $\delta > 0$ obtemos $k_0 \in \mathbb{N}$ tal que se $k > k_0$ implica $\|x - y\| < \delta$. Logo, segue que $\|f(x_k) - f(y_k)\| < \varepsilon$, ou seja, $\lim_{k \rightarrow \infty} \|f(x_k) - f(y_k)\| = 0$.

(\Leftarrow) Suponha, por contradição, que f não é uniformemente contínua. Logo, existe $\varepsilon > 0$ tal que para todo $\delta > 0$, encontramos $x_\delta, y_\delta \in X$ com $\|x_\delta - y_\delta\| < \delta$ e $\|f(x_\delta) - f(y_\delta)\| \geq \varepsilon$. Tome $\delta \in \{1, \frac{1}{2}, \dots, \frac{1}{k}, \dots\}$. Assim, existem $(x_k), (y_k) \subseteq X$ tais que $0 \leq \|x_k - y_k\| < \frac{1}{k}$ e

$\|f(x_k) - f(y_k)\| \geq \varepsilon$. Portanto, $0 \leq \lim_{k \rightarrow \infty} \|x_k - y_k\| \leq \lim_{k \rightarrow \infty} \frac{1}{k} = 0$. Ou seja, $\lim_{k \rightarrow \infty} \|x_k - y_k\| = 0$. Se $\lim_{k \rightarrow \infty} \|f(x_k) - f(y_k)\| = 0$, então $0 = \lim_{k \rightarrow \infty} \|f(x_k) - f(y_k)\| \geq \varepsilon$. O que é um absurdo, pois $\varepsilon > 0$. O que prova o teorema. ■

Definição 1.41 (Homeomorfismo). Um homeomorfismo entre o conjunto $X \subset \mathbb{R}^m$ e o conjunto $Y \subset \mathbb{R}^n$ é uma bijeção contínua $f : X \rightarrow Y$ cuja inversa $f^{-1} : Y \rightarrow X$ é também contínua.

Uma aplicação pode ser uma bijeção contínua, mas não necessariamente um homeomorfismo. De fato, considere $f : [0, 2\pi) \rightarrow S^1$, onde $f(t) = (\cos(t), \sin(t))$. Sabendo que S^1 é a circunferência de centro 0 e raio 1, note que a inversa de f , $f^{-1} : S^1 \rightarrow [0, 2\pi)$ aplica S^1 , que é compacto, sobre o intervalo $[0, 2\pi)$ não fechado, logo, não compacto. Portanto, f^{-1} não é contínua. De modo particular, perceba que f^{-1} é descontínua no ponto $a = (1, 0) = f(0) \in S^1$. Com efeito, se pusermos para cada $k \in \mathbb{N}$, $t_k = (1 - \frac{1}{k})2\pi$ e $z_k = (\cos(t_k), \sin(t_k))$, $\lim_{k \rightarrow \infty} z_k = (\cos(2\pi), \sin(2\pi)) = (1, 0) = a$, contudo, $\lim_{k \rightarrow \infty} f^{-1}(z_k) = \lim_{k \rightarrow \infty} t_k = 2\pi$. Assim, não vale $\lim_{k \rightarrow \infty} f^{-1}(z_k) = f^{-1}(a) = 0$.

Teorema 1.42. Se $K \subset \mathbb{R}^n$ é compacto, então toda aplicação contínua injetiva $f : K \rightarrow \mathbb{R}^n$ é um homeomorfismo sobre sua imagem compacta $L = f(K)$.

Demonstração: Note que $f : K \rightarrow L$ é uma bijeção. De fato, por hipótese f é injetiva e como $L = f(K) \subset \mathbb{R}^n$, então f é sobrejetiva sobre sua imagem. Logo, f é uma bijeção de K em L , e admite inversa $g = f^{-1} : L \rightarrow K$. Além disso, como K é compacto e f é contínua, segue do Teorema 1.35 que $L = f(K)$ é compacto. Para concluir que f é um homeomorfismo, resta mostrar que $g = f^{-1} : L \rightarrow K$ é contínua. Seja $F \subset K$ um conjunto fechado. Como K é compacto e F é fechado em K , segue que F é também compacto. Como f é contínua, a imagem $f(F)$ é compacta. Mas todo subconjunto compacto de \mathbb{R}^n é fechado, então $f(F)$ é fechado em \mathbb{R}^n . Assim $g = f^{-1} : L \rightarrow K$ é contínua. Portanto, f é uma bijeção contínua com inversa contínua, ou seja, f é um homeomorfismo entre K e $f(K)$. ■

1.3 Funções diferenciáveis

diferenciabilidade é uma noção central em Análise, pois garante que uma função pode ser descrita localmente por uma aproximação ou transformação linear. Esse conceito

estende a ideia de tangente das funções reais para funções em \mathbb{R}^n .

Definição 1.43. Seja $f : U \rightarrow \mathbb{R}$ uma função definida no aberto $U \subset \mathbb{R}^n$. Para cada $i = 1, \dots, n$, a i -ésima derivada parcial de f no ponto $a = (a_1, \dots, a_n) \in U$ é o número

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

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

Definição 1.44. Uma aplicação $f : U \rightarrow \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é dita de classe C^1 quando cada uma de suas funções-coordenadas $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ é de classe C^1 .

Definição 1.45. Uma função $f : U \rightarrow \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, é diferenciável no ponto $a \in U$ quando cumpre as seguintes condições:

- i. Existem as derivadas parciais $\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)$;
- ii. Para todo $v = (v_1, \dots, v_n)$ tal que $a + v \in U$, tem-se

$$f(a + v) - f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot v_i + r(v), \quad \text{com} \quad \lim_{\|v\| \rightarrow 0} \frac{r(v)}{\|v\|} = 0. \quad (1.1)$$

Observação 1.46. A definição da diferenciabilidade está na condição $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0$, pois a igualdade que define o “resto” $r(v)$ pode ser escrita para qualquer função que possua as n derivadas parciais. Com $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0$, temos que $\lim_{v \rightarrow 0} r(v) = 0$, pois

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

Assim, segue que $\lim_{v \rightarrow 0} [f(a + v) - f(a)] = 0$. Ou ainda, não basta que a função tenha derivadas parciais, para que f seja diferenciável, é necessário que o termo de erro (ou “resto”) seja desprezível em relação a $\|v\|$.

Proposição 1.47. Se $f : U \rightarrow \mathbb{R}$ é diferenciável em $a \in U$, então f é contínua em a .

Demonstração: Basta mostrar que

$$\lim_{\|v\| \rightarrow 0} f(a + v) = f(a).$$

Como f é diferenciável em a , existem as derivadas parciais $\frac{\partial f}{\partial x_i}(a)$ para todo $i \in \{1, \dots, n\}$ e existe $r(v)$ tal que

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

Daí, observe que

$$f(a + v) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot v_i + \frac{r(v)}{\|v\|} \cdot \|v\|.$$

Assim, temos que

$$\lim_{\|v\| \rightarrow 0} f(a + v) = \lim_{\|v\| \rightarrow 0} \left(f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot v_i + \frac{r(v)}{\|v\|} \cdot \|v\| \right) = f(a),$$

Ou seja, f é contínua em a . ■

Teorema 1.48. *Toda função $f : U \rightarrow \mathbb{R}$ de classe C^1 é diferenciável em todo ponto de U , onde $U \subset \mathbb{R}^n$ é aberto.*

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

$$\begin{aligned} r(v) &= f(a_1 + v_1, a_2 + v_2) - f(a_1, a_2) - \sum_{i=1}^2 \frac{\partial f}{\partial x_i}(a_1, a_2) \cdot v_i \\ &= f(a_1 + v_1, a_2 + v_2) - f(a_1, a_2) - \frac{\partial f}{\partial x_1}(a_1, a_2)v_1 - \frac{\partial f}{\partial x_2}(a_1, a_2)v_2. \end{aligned} \quad (1.2)$$

Agora defina,

$$g : [0, v_1] \rightarrow \mathbb{R} \quad \text{tal que} \quad g(t) = f(a_1 + t, a_2 + v_2)$$

e

$$h : [0, v_2] \rightarrow \mathbb{R} \quad \text{tal que} \quad h(t) = f(a_1, a_2 + t)$$

que são funções diferenciáveis e, pelo Teorema do Valor Médio na reta, temos que existe

$\theta_1 \in (0, 1), \theta_2 \in (0, 1)$ tais que $g(v_1) - g(0) = g'(v_1\theta_1)v_1$, isso implica em

$$f(a_1 + v_1, a_2 + v_2) - f(a_1, a_2 + v_2) = \frac{\partial f}{\partial x_1}(a_1 + \theta_1 v_1, a_2 + v_2)v_1 \quad (1.3)$$

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

$$f(a_1, a_2 + v_2) - f(a_1, a_2) = \frac{\partial f}{\partial x_2}(a_1, a_2 + \theta_2 v_2)v_2 \quad (1.4)$$

Utilizando (1.2), (1.3) e (1.4), obtemos

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

Com isto, observe que

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

Como $\frac{\partial f}{\partial x_i}$ são contínuas para $i = 1, 2$, por $\left| \frac{v_i}{|v|} \right| \leq 1$, para $i = 1, 2$, segue que

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

assim

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

Logo, f é diferenciável em todo ponto de U . ■

Definição 1.49 (Gradiente). Sejam $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \rightarrow \mathbb{R}$ uma função diferenciável. O gradiente de f no ponto $a \in U$ é um vetor dado por

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right).$$

Definição 1.50. Sejam $f : U \rightarrow \mathbb{R}$ diferenciável e $a \in U$. Dizemos que a é um ponto crítico de f quando $\nabla f(a) = 0$.

Definição 1.51. Seja v um vetor de \mathbb{R}^n . A derivada direcional de f no ponto a , na direção de v é

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

caso este limite exista.

1.4 Teorema de Schwarz

O Teorema de Schwarz é um dos principais resultados na Análise. Ele afirma que, sob condições adequadas de regularidade, se a função f é de classe C^2 em um aberto de \mathbb{R}^n , então as derivadas parciais mistas de segunda ordem coincidem. Como consequência, a matriz Hessiana de f , que é a matriz quadrada formada pelas derivadas parciais de segunda ordem, é simétrica. Esse resultado é essencial para a demonstração do Lema de Morse.

Definição 1.52. Seja $f : U \rightarrow \mathbb{R}$ uma função que possui as derivadas parciais $\frac{\partial f}{\partial x_1}(x)$, $\frac{\partial f}{\partial x_2}(x)$, \dots , $\frac{\partial f}{\partial x_n}(x)$ em todo ponto x do aberto $U \subset \mathbb{R}^n$. A j -ésima derivada parcial da função $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ no ponto $x \in U$ é indicada por

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

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

A demonstração do Teorema de Schwarz baseia-se em um princípio atribuído a Leibniz, que permite a diferenciação sob o sinal de integral, desde que o integrando resultante seja uma função contínua. Por sua vez, a demonstração do Teorema de Leibniz utiliza o seguinte Lema

Lema 1.53. *Sejam $X \subset \mathbb{R}^n$ um conjunto arbitrário e $K \subset \mathbb{R}^n$ compacto. Fixemos $x_0 \in X$. Se $f : X \times K \rightarrow \mathbb{R}^p$ é contínua, então para todo $\varepsilon > 0$ existe $\delta > 0$ tal que:*

$$\|x - x_0\| < \delta \implies \|f(x, t) - f(x_0, t)\| < \varepsilon$$

para quaisquer $x \in X$ e $t \in K$.

Demonstração: Fixe $x_0 \in X$. Supondo por contradição, que existe $\varepsilon > 0$ tal que, para todo $\delta > 0$, existem $x \in X$ e $t \in K$ com $\|x - x_0\| < \delta$ e $\|f(x, t) - f(x_0, t)\| \geq \varepsilon$. Em

particular, tomando $\delta = \frac{1}{k}$, e seqüências de pontos $x_k \in X$ e $t_k \in K$ tais que

$$\|x_n - x_0\| < \frac{1}{k} \quad \text{e} \quad \|f(x_k, t_n) - f(x_0, t_k)\| \geq \varepsilon.$$

Assim, $x_k \rightarrow x_0$ quando $k \rightarrow \infty$, e $(t_k) \subset K$, que é compacto. Portanto, existe uma subseqüência (t_{k_n}) que converge para algum $t_0 \in K$. Como $x_{k_n} \rightarrow x_0$ e $t_{k_n} \rightarrow t_0$, temos $(x_{k_n}, t_{k_n}) \rightarrow (x_0, t_0)$. Como f é contínua em $X \times K$, segue que

$$f(x_{k_n}, t_{k_n}) \rightarrow f(x_0, t_0) \quad \text{e} \quad f(x_0, t_{k_n}) \rightarrow f(x_0, t_0).$$

$$\varepsilon \leq \|f(x_{k_n}, t_{k_n}) - f(x_0, t_{k_n})\| = 0,$$

o que é um absurdo. ■

Com a demonstração do Lema anterior podemos agora enunciar e demonstrar o Teorema de Leibniz. O qual garante quando podemos derivar sob o sinal da integral desde que o integrando resultante seja uma função contínua.

Teorema 1.54 (Derivação sob o sinal da integral.). *Sejam $U \subseteq \mathbb{R}^n$ aberto e $f : U \times [a, b] \rightarrow \mathbb{R}$ uma função tal que $\frac{\partial f}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ existem e são contínuas para todo $i = 1, \dots, n$ e f é contínua. Daí, defina a função $\varphi : U \rightarrow \mathbb{R}$, dada por $\varphi(x) = \int_a^b f(x, t) dt$.*

Então, $\frac{\partial \varphi}{\partial x_i} : U \rightarrow \mathbb{R}$ existe e vale que $\frac{\partial \varphi}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x, t) dt$.

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

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

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

$$\begin{aligned} & \frac{\varphi(x_0 + se_i) - \varphi(x_0)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t) dt \\ & = \frac{1}{s} \int_a^b \left[\frac{\partial f}{\partial x_i}(x_0 + \theta e_i, t) \cdot s \right] dt - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t) dt \\ & = \int_a^b \left[\frac{\partial f}{\partial x_i}(x_0 + \theta e_i, t) - \frac{\partial f}{\partial x_i}(x_0, t) \right] dt. \end{aligned} \tag{1.5}$$

Pelo Lema 1.53, dado $\varepsilon > 0$, podemos exibir $\delta > 0$ tal que se $\|\theta e_i\| < \delta$, então

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

Retomando a (1.5) com $|s| < \delta$ temos

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

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

Teorema 1.55 (Schwarz). *Se $f : U \rightarrow \mathbb{R}$ é de classe C^2 no aberto $U \subset \mathbb{R}^n$, então para quaisquer $i, j = 1, \dots, n$ e $x \in U$, tem-se*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

Demonstração: Suponha, sem perda de generalidade, que $U = I \times J$ seja um retângulo em \mathbb{R}^2 . Fixando $b \in J$, o Teorema Fundamental do Cálculo garante que, para todo $(x, y) \in U$, tem-se

$$f(x, y) = f(x, b) + \int_b^y \frac{\partial f}{\partial y}(x, t) dt.$$

Como $\frac{\partial^2 f}{\partial x \partial y}$ é contínua, pelo Teorema 1.54, podemos derivar sob o sinal da integral.

Assim, $\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial x}(x, b) + \int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t) dt$.

Logo depois, derivando em relação a y

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x}(x, y) \right) \\ &= \frac{\partial}{\partial y} \left[\frac{\partial f}{\partial x}(x, b) + \int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t) dt \right] \\ &= \frac{\partial}{\partial y} \left[\frac{\partial f}{\partial x}(x, b) \right] + \frac{\partial}{\partial y} \left[\int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t) dt \right]. \end{aligned} \tag{1.6}$$

Note que $\frac{\partial f}{\partial x}(x, b)$ não depende de y portanto é constante. Logo a derivada em relação a y é zero, isto é $\frac{\partial}{\partial y} \left[\frac{\partial f}{\partial x}(x, b) \right] = 0$.

Seja H de tal forma que $H'(t) = \frac{\partial^2 f}{\partial x \partial y}$, então

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

Derivando em relação a y

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

Retornando a equação (1.6) temos

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x}(x, y) &= 0 + \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{\partial^2 f}{\partial x \partial y}(x, y) \end{aligned}$$

■

Exemplo 1.56. O caso mais conhecido de uma função que possui derivadas de segunda ordem em todos os pontos do plano mas $\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$ na origem é $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{se } x^2 + y^2 \neq 0, \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

. Note que nos pontos de $\mathbb{R}^2 - \{0\}$, f possui derivadas parciais repetidas de todas as ordens. Na origem, existem (e são nulas) $\frac{\partial^2 f}{\partial x^2}$ e $\frac{\partial^2 f}{\partial y^2}$. Quanto às derivadas mistas, começamos observando que, para todo $y \neq 0$, temos $f(0, y) = 0$, logo

$$\frac{\partial f}{\partial x}(0, y) = \lim_{x \rightarrow 0} \frac{xy(x^2 - y^2)}{x(x^2 + y^2)} = -y,$$

e daí

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

De forma análoga, observe que $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$. Logo, as duas derivadas mistas de f na origem são diferentes.

Falaremos brevemente da fórmula de Taylor no qual nos diz que é possível aproximar uma função diferenciável por um polinômio, construído a partir das derivadas da função em um ponto fixado.

A fórmula de Taylor para uma função $f: U \rightarrow \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, é dado por

$$f(a+v) = f(a) + df(a) \cdot v + \frac{1}{2} d^2 f(a) \cdot v^2 + \cdots + \frac{1}{p!} d^p f(a) \cdot v^p + r_p(v).$$

Vamos apresentar três principais situações da fórmula de Taylor. São elas:

1. **Fórmula de Taylor Infinitesimal.** Se f é p vezes diferenciável no ponto a , então $\lim_{v \rightarrow 0} \frac{r_p(v)}{\|v\|^p} = 0$.
2. **Resto de Lagrange.** Supondo $[a, a+v] \subset U$, f de classe C^p , $p+1$ vezes diferenciável no segmento aberto $(a, a+v)$, então existe $\theta \in (0, 1)$ tal que

$$r_p(v) = \frac{1}{(p+1)!} d^{p+1} f(a + \theta v) \cdot v^{p+1}.$$

3. **Resto Integral.** Se f é de classe C^{p+1} e $[a, a+v] \subset U$ então

$$r_p(v) = \frac{1}{p!} \int_0^1 (1-t)^p d^{p+1} f(a+tv) \cdot v^{p+1} dt.$$

Dai, para $v = (\alpha_1, \dots, \alpha_n)$, temos que

$$d^2 f(a) \cdot v^2 = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \alpha_i \alpha_j,$$

$$d^3 f(a) \cdot v^3 = \sum_{i,j,k=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a) \alpha_i \alpha_j \alpha_k,$$

e assim por diante.

Para cada inteiro $p > 0$, a forma $d^p f(a): \mathbb{R}^n \rightarrow \mathbb{R}$ chama-se a p -ésima diferencial da função f no ponto a . O valor de $d^p f(a)$ num vetor $v \in \mathbb{R}^n$ é indicado pela notação $d^p f(a) \cdot v^p$, para sugerir que se trata de um polinômio homogêneo de grau p nas coordenadas de v .

Teorema 1.57 (Fórmula de Taylor). *Seja $f: U \rightarrow \mathbb{R}$ de classe C^2 no aberto $U \subset \mathbb{R}^n$. Fixado $a \in U$, para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ tal que $a + v \in U$, escrevamos*

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

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

Demonstração: Veja em [4]. ■

Agora falaremos sobre a Hessiana e a forma quadrática, definição e resultados que serão utilizados nos próximos capítulos.

Definição 1.58 (Hessiana). *Seja $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função. Se todas as derivadas parciais de segunda ordem de f existem, então a matriz Hessiana \mathbf{H} de f é uma matriz quadrada $n \times n$, onde*

$$Hf(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

Se $t \in \mathbb{R}$ então $H \cdot (tv)^2 = t^2(\cdot H \cdot v^2)$

Outra importante definição é o que chamamos de forma quadrática.

Definição 1.59. Uma forma quadrática $H: \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função que associa um vetor $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ a um número real, por meio de uma matriz $H = (h_{ij})$ simétrica $n \times n$. Ela é dada por

$$H \cdot v^2 = \sum_{i,j=1}^n h_{ij} \alpha_i \alpha_j.$$

Exemplo 1.60. Considere a matriz simétrica 3×3 dada por:

$$H = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \text{ com } a, b, c, d, e, f \in \mathbb{R}$$

e o vetor $v = (\alpha_1, \alpha_2, \alpha_3) = (x, y, z) \in \mathbb{R}^3$.

A forma quadrática associada à matriz H é definida por:

$$H \cdot v^2 = \sum_{i=1}^3 \sum_{j=1}^3 h_{ij} \alpha_i \alpha_j.$$

Substituindo os elementos da matriz e do vetor, temos:

$$\begin{aligned} H \cdot v^2 &= h_{11}x^2 + h_{12}xy + h_{13}xz \\ &\quad + h_{21}yx + h_{22}y^2 + h_{23}yz \\ &\quad + h_{31}zx + h_{32}zy + h_{33}z^2 \\ &= ax^2 + bxy + cxz + byx + dy^2 + eyz + czx + ezy + fz^2 \\ &= ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz. \end{aligned}$$

Portanto, a forma quadrática associada à matriz H é:

$$H(x, y, z) = ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz$$

A forma Hessiana da função duas vezes diferenciável $f: U \rightarrow \mathbb{R}$ no ponto $x \in U$ será indicada com $H(x)$, ou $Hf(x)$. Sabemos que $H(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)$, portanto

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

O Teorema de Schwarz 1.55 garante que a matriz $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)$, chamada *matriz Hessiana* de f no ponto x , é simétrica, desde que as derivadas parciais de segunda ordem sejam contínuas.

Exemplo 1.61. Seja $H: \mathbb{R}^n \rightarrow \mathbb{R}$ uma forma quadrática dada por $H \cdot v^2 = \sum_{i,j=1}^n h_{ij} \alpha_i \alpha_j$ para $v = (\alpha_1, \dots, \alpha_n)$. A forma quadrática pode ser:

1. Positiva quando $H \cdot v^2 > 0$ para todo $v \neq 0$ em \mathbb{R}^n ;
2. Negativa quando $H \cdot v^2 < 0$ para todo $v \neq 0$ em \mathbb{R}^n ;
3. Indefinida quando existirem vetores $u, v \in \mathbb{R}^n$ de modo que $H \cdot v^2 > 0$ e $H \cdot u^2 < 0$.

Observação 1.62. Se H for positiva ou negativa, diremos que H é uma forma definida.

Observação 1.63. Se uma forma H é definida então sua matriz (h_{ij}) é necessariamente invertível. De fato, se indicarmos com $H_0: \mathbb{R}^n \rightarrow \mathbb{R}^n$ a transformação linear cuja matriz relativamente à base canônica é (h_{ij}) , temos $H \cdot v^2 = \langle H_0 \cdot v, v \rangle$. Assim, H definida $\Rightarrow \langle H_0 \cdot v, v \rangle \neq 0$ para todo $v \neq 0 \Rightarrow H_0 \cdot v \neq 0$ para todo $v \neq 0 \Rightarrow H_0$ é invertível.

Definição 1.64. Seja $a \in U \subset \mathbb{R}^n$ um ponto crítico. Diremos que um ponto crítico a é não degenerado quando a matriz Hessiana nesse ponto é invertível, isto é, $\det \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right) \neq 0$.

Definição 1.65. Um ponto $a \in U$ é dito um ponto de máximo (mínimo) local da função $f: U \rightarrow \mathbb{R}$ quando existe $\delta > 0$ tal que $f(x) \leq f(a)$ para todo $x \in U \cap B(a; \delta)$ (respectivamente, $f(x) \geq f(a)$ para todo $x \in U \cap B(a; \delta)$).

Podemos relacionar o conceito de ponto crítico com Hessiana da seguinte forma

Teorema 1.66. *Sejam $f: U \rightarrow \mathbb{R}$ uma função de classe C^2 , $a \in U$ um ponto crítico de f e H a forma quadrática Hessiana de f no ponto a . Então:*

(i) *Se H é positiva, então a é um mínimo local de f .*

(ii) *Se H é negativa, então a é um máximo local de f .*

(iii) *Se H é indefinida, então a não é máximo e nem mínimo local de f .*

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

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

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

$$f(a + v) - f(a) = \frac{1}{2}H \cdot v^2 + r(v), \tag{1.7}$$

onde H é a forma quadrática Hessiana de f em a . Uma vez que H é contínua e S^{n-1} é compacta, existe $v_0 \in S^{n-1}$ tal que $H \cdot v_0^2 = 2c \leq H \cdot v^2$, para todo $v \in S^{n-1}$. Como H é

forma positiva, temos que $c > 0$. Retornando à equação (1.7) observamos que

$$\begin{aligned} f(a+v) - f(a) &= \frac{\|v\|^2}{2} \cdot H \cdot \left(\frac{v}{\|v\|}\right)^2 + \|v\|^2 \cdot \left(\frac{r(v)}{\|v\|^2}\right) \\ &\geq \frac{\|v\|^2}{2} \cdot (2c) + \|v\|^2 \cdot \left(\frac{r(v)}{\|v\|^2}\right) \\ &= \|v\|^2 \cdot \left(c + \frac{r(v)}{\|v\|^2}\right). \end{aligned} \quad (1.8)$$

Mas lembre-se que $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^2} = 0$. Assim, dado $\varepsilon = \frac{c}{2} > 0$, existe $\delta > 0$ tal que se $\|v\| < \delta$ implica

$$\left| \frac{r(v)}{\|v\|^2} \right| < \frac{c}{2},$$

ou seja,

$$-\frac{c}{2} < \frac{r(v)}{\|v\|^2} < \frac{c}{2}.$$

Voltando para a equação (1.8), segue que para $\|v\| < \delta$:

$$f(a+v) - f(a) > \|v\|^2 \left(c - \frac{c}{2}\right) = \|v\|^2 \cdot \frac{c}{2} > 0$$

para todo $v \neq 0$. Ou seja, $f(a+v) > f(a)$ para todo $v \in B_\delta(0) \setminus \{0\}$. Logo, a é um mínimo local estrito de f .

(ii) Similar ao item (i).

(iii) Como H é indefinida, existem $v, w \in \mathbb{R}^n$ tais que $H \cdot v^2 > 0$ e $H \cdot w^2 < 0$. Considerando $t \in \mathbb{R}^*$ suficientemente pequeno de modo que $(a+tv), (a+tw) \in U$, por (1.7) segue que:

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

Então $f(a + tw) > f(a)$. Por outro lado:

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

Então $f(a + tw) < f(a)$, a não é mínimo local nem máximo local de f . ■

Definição 1.67. Seja (a_{ij}) uma matriz real $n \times n$ simétrica, isto é, $a_{ij} = a_{ji}$. A ela corresponde uma transformação linear $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que é dita auto-adjunta, o que é equivalente a $\langle Ax, y \rangle = \langle x, Ay \rangle$ para todo $x, y \in \mathbb{R}^n$.

Definição 1.68. Um autovetor de A é um vetor não nulo $x \in \mathbb{R}^n$ tal que $A \cdot x = \lambda x$ para algum escalar $\lambda \in \mathbb{R}$, o qual é chamado de autovalor correspondente.

Lema 1.69 (Teorema Espectral em \mathbb{R}^n). *Se $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é auto-adjunta, então existe uma base ortonormal de \mathbb{R}^n formada por autovetores de A .*

Demonstração: Para provar a existência dos autovetores, introduzimos a forma quadrática $f : \mathbb{R}^n \rightarrow \mathbb{R}$ associada a A , dada por $f(x) = \langle A \cdot x, x \rangle$. Buscaremos os pontos críticos de f restrita à esfera unitária $S^{n-1} = \{x \in \mathbb{R}^n \mid \langle x, x \rangle = 1\}$. Uma vez que S^{n-1} é um conjunto compacto e f é contínua, f atinge seu valor máximo, denotado por λ_1 , em um ponto $u_1 \in S^{n-1}$. Pelo método dos Multiplicadores de Lagrange, u_1 é um ponto crítico de $f|_{S^{n-1}}$ se satisfaz $\nabla f(u_1) = \lambda' \nabla \varphi(u_1)$, onde $\varphi(x) = \langle x, x \rangle$. Com o cálculo dos gradientes $\nabla f(x) = 2A \cdot x$ e $\nabla \varphi(x) = 2 \cdot x$, a condição dos multiplicadores se torna $2A \cdot u_1 = \lambda'(2 \cdot u_1)$, o que implica $A \cdot u_1 = \lambda' u_1$. Assim, o vetor u_1 que maximiza a forma quadrática f sobre a esfera unitária é, na verdade, um autovetor de A com autovalor $\lambda_1 = \lambda'$. A existência de pelo menos um autovetor e autovalor real está garantida. A seguir, consideramos o subespaço $(n - 1)$ -dimensional $E = \{x \in \mathbb{R}^n \mid \langle x, u_1 \rangle = 0\}$, o complemento ortogonal de u_1 . A propriedade auto-adjunta de A garante que E é um subespaço invariante sob A . De fato, para qualquer $x \in E$:

$$\langle A \cdot x, u_1 \rangle = \langle x, A \cdot u_1 \rangle = \langle x, \lambda_1 u_1 \rangle = \lambda_1 \langle x, u_1 \rangle = 0.$$

Portanto, se $x \in E$, temos que $A \cdot x \in E$. Restringindo a transformação A ao subespaço

E , obtemos uma nova transformação auto-adjunta $A|_E: E \rightarrow E$. Repetindo o processo de otimização sobre a esfera unitária de E , encontramos um novo autovetor $u_2 \in E$ com autovalor λ_2 . Por estar em E , u_2 é ortogonal a u_1 . Prosseguindo este processo de forma indutiva nos subespaços ortogonais subsequentes, obtemos uma sequência de n autovetores (u_1, u_2, \dots, u_n) que são mutuamente ortogonais. Como são n vetores não nulos ortogonais em \mathbb{R}^n , eles formam automaticamente uma base de \mathbb{R}^n . Normalizando-os, concluímos que existe uma base ortonormal de \mathbb{R}^n formada por autovetores de A . ■

Teorema 1.70. *Seja $M \in \mathbb{R}^{n \times n}$ uma matriz, e vetores $u, v \in \mathbb{R}^n$. Então, vale a seguinte igualdade:*

$$\langle Mu, v \rangle = \langle u, M^\top v \rangle,$$

onde M^\top denota a transposta de M .

Demonstração: Vamos usar a definição padrão de produto interno em \mathbb{R}^n :

$$\langle x, y \rangle = x^\top y.$$

Com isto, temos

$$\langle Mu, v \rangle = (Mu)^\top v.$$

Usando que $(M \cdot u)^\top = u^\top \cdot M^\top$, obtemos

$$\langle Mu, v \rangle = u^\top M^\top v = \langle u, M^\top v \rangle.$$

■

Capítulo 2

O Teorema da Função Inversa

A formulação rigorosa do Teorema da Função Inversa é atribuída ao matemático italiano Ulisse Dini, em 1876. Em sua obra *Lezioni di Analisi Infinitesimale*, Dini apresentou uma abordagem inovadora ao demonstrar, inicialmente, o Teorema da Função Implícita por meio de indução matemática. A partir desse resultado, deduziu o Teorema da Função Inversa como uma consequência direta. Sua demonstração destacou-se por empregar apenas ferramentas elementares, como o Teorema do Valor Intermediário e o Teorema do Valor Médio, sem recorrer a conceitos mais sofisticados, como compacidade ou princípios de ponto fixo.

No século XX, duas abordagens alternativas se consolidaram. A primeira, comum em livros como os de Apostol e Rudin, adota uma estratégia inversa à de Dini: inicia com a demonstração do Teorema da Função Inversa, utilizando compacidade no espaço \mathbb{R}^n e o Teorema de Weierstrass, e depois conclui o Teorema da Função Implícita. A segunda abordagem baseia-se no Princípio da Contração de Banach, o que permite generalizar o resultado para espaços de dimensão infinita.

Antes de enunciar e demonstrar o Teorema da função inversa, traremos mais algumas definições e resultados necessários.

Definição 2.1. Sejam $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ abertos. Uma aplicação $f: U \rightarrow V$ é dita um *difeomorfismo* entre U e V quando é uma bijeção diferenciável, cuja inversa $g = f^{-1}: V \rightarrow U$ também é diferenciável.

Exemplo 2.2. Sejam $f, g: \mathbb{R} \rightarrow \mathbb{R}$ definidas por:

$$f(x) = 2x \quad \text{e} \quad g(y) = \frac{1}{2}y.$$

E note que

$$(g \circ f)(x) = g(f(x)) = g(2x) = \frac{1}{2} \cdot 2x = x,$$

$$(f \circ g)(y) = f(g(y)) = f\left(\frac{1}{2}y\right) = 2 \cdot \frac{1}{2}y = y.$$

Logo, f e g são homeomorfos. Para qualquer $x, h \in \mathbb{R}$

$$f(x+h) - f(x) = 2(x+h) - 2x = 2h.$$

Como $f'(x) = 2$, temos

$$f(x+h) - f(x) = f'(x) \cdot h + 0,$$

onde o resto $r(h) = 0$. Portanto

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

Analogamente para g

$$g(y+k) - g(y) = \frac{1}{2}(y+k) - \frac{1}{2}y = \frac{1}{2}k,$$

com resto $s(k) = 0$ e:

$$\lim_{k \rightarrow 0} \frac{s(k)}{|k|} = 0.$$

Portanto, f e g são difeomorfismos de \mathbb{R} em \mathbb{R} .

Observação 2.3. Nem todo homeomorfismo diferenciável é um difeomorfismo. Ou seja, nem todo homeomorfismo diferenciável possui inversa diferenciável.

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

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

Logo, $g'(0)$ não existe.

Definição 2.4. Seja $U \subset \mathbb{R}^m$ um aberto. Uma aplicação diferenciável $f : U \rightarrow \mathbb{R}^m$, é dita

um *difeomorfismo local*, quando para cada $x \in U$ existe uma bola aberta $B = B(x; \delta) \subset U$ tal que f aplica B difeomorficamente sobre um aberto V contendo $f(x)$.

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

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

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

Observação 2.5. O difeomorfismo local $f : U \rightarrow \mathbb{R}^m$ é um difeomorfismo global de U sobre o aberto $V = f(U) \subset \mathbb{R}^m$, se, e somente se, f é uma aplicação injetiva.

Teorema 2.6. *Seja $f : U \rightarrow \mathbb{R}^n$ de classe C^1 no aberto $U \subset \mathbb{R}^n$. Se, para algum $a \in U$, a derivada $f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é injetiva, então existem $\delta > 0$ e $c > 0$ tais que $B = B(a, \delta) \subset U$ e, para quaisquer $x, y \in B$ tem-se $\|f(x) - f(y)\| \geq c\|x - y\|$. Em particular, a restrição $f|_B$ é injetiva.*

Demonstração: Perceba que $u \mapsto \|f'(a) \cdot u\|$ é positiva em todos os pontos u da esfera unitária S^{m-1} , que é compacta. Pelo Corolário 1.36, existe $c > 0$ tal que $\|f'(a) \cdot u\| \geq 2c$ para todo $u \in S^{m-1}$. Pela linearidade, segue que $\|f'(a) \cdot v\| \geq 2c \cdot \|v\|$ para todo $v \in \mathbb{R}^m$. Além disso, para todo $x \in U$, escrevamos

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

Então, para $x, y \in U$ quaisquer, temos $f(x) - f(y) = f'(a) \cdot (x - y) + r(x) - r(y)$. Pela desigualdade triangular, segue que

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

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

Teorema 2.7 (Diferenciabilidade do Homeomorfismo Inverso). *Seja $f : U \rightarrow V$ um homeomorfismo de classe C^1 entre os abertos $U, V \subset \mathbb{R}^n$. Se para algum $x \in U$, a derivada $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um operador invertível, então o homeomorfismo inverso $g = f^{-1} : V \rightarrow U$ é diferenciável no ponto $f(x)$, com $g'(f(x)) = [f'(x)]^{-1}$.*

Demonstração: Sejam $x \in U$ e $x + v \in U$, escrevamos $f(x) = y$ e $f(x + v) = y + w$. Então, $w = f(x + v) - f(x) = f'(x) \cdot v + r(v)$, $\lim_{v \rightarrow 0} r(v)/\|v\| = 0$ e $v = g(f(x + v)) - g(f(x)) = g(y + w) - g(y)$. Para provarmos que $[f'(x)]^{-1}$ é a derivada de g no ponto y , escrevemos

$$s(w) = g(y + w) - g(y) - f'(x)^{-1} \cdot w, \quad (2.1)$$

e basta-nos mostrar que $\lim_{w \rightarrow 0} \frac{s(w)}{\|w\|} = 0$. Substituindo as expressões de v e w na equação (2.1) temos $v = f'(x)^{-1}[f'(x) \cdot v + r(v)] + s(w)$. Assim

$$\begin{aligned} v &= f'(x)^{-1} \cdot f'(x) \cdot v + f'(x)^{-1} \cdot r(v) + s(w) \\ &= I \cdot v + f'(x)^{-1} \cdot r(v) + s(w) \\ &= v + f'(x)^{-1} \cdot r(v) + s(w). \end{aligned}$$

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

$$\frac{s(w)}{\|w\|} = -f'(x)^{-1} \cdot \frac{r(v)}{\|v\|} \cdot \frac{\|v\|}{\|w\|}. \quad (2.2)$$

Substituindo a expressão de w na equação (2.2), obtemos

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

Quando $w \rightarrow 0$, temos que $v \rightarrow 0$ pela continuidade de g , visto que f é um homeomorfismo. Assim, $\frac{r(v)}{\|v\|} \rightarrow 0$. Além disso, como $f \in C^1$ no aberto $U \subset \mathbb{R}^n$ e $f'(x)^{-1}$ é injetiva, pelo Teorema 2.6, existem $\delta > 0$ e $c > 0$ tais que $\|v\| < \delta$ implica $\|f(x + v) - f(x)\| \geq c\|v\|$,

portanto

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

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

Corolário 2.8. *Se $f : U \rightarrow V$ é um homeomorfismo de classe C^k cuja derivada $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é invertível para todo $x \in U$, então seu inverso $g = f^{-1} : V \rightarrow U$ é de classe C^k .*

Demonstração: De fato, a derivada $g' : V \rightarrow \mathcal{L}(\mathbb{R}^n)$, dada por $g'(y) = [f'(g(y))]^{-1}$ para cada $y \in V$, pode ser escrita como $g' = \text{inv} \circ f' \circ g$, onde a aplicação inv , de classe C^∞ , é a inversão de transformações lineares bijetivas e $f' \in C^{k-1}$. Admitindo, por indução, que $g \in C^{k-1}$, resulta que $g' \in C^{k-1}$, logo $g \in C^k$. ■

Lema 2.9. *Sejam $U \subset \mathbb{R}^m$ aberto e $g : U \rightarrow \mathbb{R}^n$ diferenciável no ponto $a \in U$, com $g'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ sobrejetiva. Se a é um ponto de mínimo local de $\|g(x)\|$, com $x \in U$, então $g(a) = 0$.*

Demonstração: Se a é um ponto de mínimo local para $\|g(x)\|$, a também será um ponto de mínimo local para $\varphi : U \rightarrow \mathbb{R}$, dada por $\varphi(x) = \langle g(x), g(x) \rangle$. Note que

$$\varphi'(x) = \langle g'(x), g(x) \rangle + \langle g(x), g'(x) \rangle = 2\langle g'(x), g(x) \rangle.$$

Então, $\varphi'(a) = 2\langle g'(a), g(a) \rangle$. Como $\|g(x)\| \leq \|g(a)\|$ para todo $x \in U$ suficientemente próximo de a , temos que $\varphi(a) \leq \varphi(x)$ para todo $x \in U$ suficientemente próximo de a . Logo, a é um mínimo local para φ . Por φ ser diferenciável, a é ponto crítico, então $\varphi'(a) = 0$. Então, para todo $v \in \mathbb{R}^m$, temos

$$\varphi'(a) \cdot v = 2\langle g'(a) \cdot v, g(a) \rangle$$

$$0 \cdot v = 2\langle g'(a) \cdot v, g(a) \rangle$$

$$0 = 2\langle g'(a) \cdot v, g(a) \rangle.$$

Por $g'(a)$ ser sobrejetiva, ela não pode ser a matriz nula. Escolhendo v_0 de modo que $g'(a) \cdot v_0 = g(a)$, temos que

$$0 = \varphi'(a) \cdot v_0 = 2\langle g(a), g(a) \rangle = 2\|g(a)\|^2.$$

Portanto, $g(a) = 0$. ■

Agora temos tudo o que precisaremos para demonstrar o Teorema da Função Inversa.

Teorema 2.10 (Teorema da Função Inversa). *Seja $f : U \rightarrow \mathbb{R}^m$ de classe C^k ($k \geq 1$) no aberto $U \subset \mathbb{R}^m$. Se $a \in U$ é tal que $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é invertível, então existe uma bola aberta $B = B(a, \delta) \subset U$ tal que a restrição $f|_B$ é um difeomorfismo, cuja inversa é de classe C^k , sobre um aberto $V \ni f(a)$.*

Demonstração: Por U ser aberto, existe $\delta_1 > 0$ tal que $B(a, \delta_1) \subset U$ e, pelo Teorema 2.6, existe $\delta_2 > 0$ de modo que $f|_{B(a, \delta_2)}$ é injetiva. Tomando $\delta_3 = \min\{\delta_1, \delta_2\}$ e $\delta = \frac{\delta_3}{2}$, obtemos que $f : B_1 \rightarrow f(B_1)$ é bijetiva, onde $B_1 = B[a, \delta] = \overline{B(a, \delta)}$. Por f ser contínua, usando o Teorema 1.42, obtemos que $f : B_1 \rightarrow f(B_1)$ é um homeomorfismo de classe C^1 e como $f'(a)$ é invertível por hipótese, segue que $f'(x)$ é um operador invertível para todo $x \in B(a, \delta)$ “próximo” de a .

Mostremos agora que $f(B(a, \delta)) \subset \mathbb{R}^m$ é aberto. Seja $p \in B(a, \delta)$ tal que $f(p) = q$. Considere $S = S[a; \delta]$ que é a fronteira de $\overline{B(a; \delta)}$. Como $f|_{\overline{B(a; \delta)}}$ é injetiva, $q = f(p) \notin f(S)$, pois se $q \in f(S)$, existiria $x_1 \in S$ tal que $f(x_1) = q$. Então $f(p) = q = f(x_1)$, mas como f é injetiva, $p = x_1$, o que é um absurdo, pois p pertence ao interior da bola e x_1 está na esfera.

Portanto, existe $\varepsilon > 0$ tal que $\|f(x) - q\| \geq 2\varepsilon$ para todo $x \in S$, pois $f(S)$ é compacto. Agora, afirmamos que $B(q, \varepsilon) \subset f(B(a, \delta))$. Com efeito, seja $y \in B(q, \varepsilon)$. Considerando $g(x) = f(x) - y$, temos que o mínimo de $\|g(x)\|$ não é atingido num ponto $x \in S$ quando x varia no compacto $\overline{B(a; \delta)}$, pois se $x_1 \in S$, teríamos $h(x_1) \leq h(x)$ para todo $x \in \overline{B(a, \delta)}$ (onde $h(x) = \|g(x)\|$). Mas, note que

$$\|f(x_1) - y\| \geq \|f(x_1) - q\| - \|q - y\| > 2\varepsilon - \varepsilon = \varepsilon,$$

enquanto que, se $x = p$, teríamos $\|f(p) - y\| = \|q - y\| < \varepsilon$. Portanto, $\|f(p) - y\| < \varepsilon < \|f(x_1) - y\|$, o que contradiz a suposição de que o mínimo ocorre em x_1 .

Assim, o mínimo de $\|g(x)\| = \|f(x) - y\|$ é atingido num ponto $x_0 \in B(a; \delta)$. Diante disso, note que x_0 é ponto de mínimo de $\|g(x)\|$ e g é de classe C^1 . Como $f'(x_0)$ é sobrejetiva (pois é invertível, já que x_0 está próximo de a), então pelo Lema 2.9 temos que $g(x_0) = 0$, ou seja, $f(x_0) - y = 0$, portanto $f(x_0) = y$. Assim, $y \in f(B(a, \delta))$, logo

$$B(q, \varepsilon) \subset f(B(a, \delta)).$$

Tendo feito uso do Teorema 2.7, concluímos que $f : B(a, \delta) \rightarrow f(B(a, \delta))$ é um difeomorfismo, e mais, pelo Corolário 2.8, tal difeomorfismo é de classe C^k .



Capítulo 3

O Lema da Morse

Neste capítulo enunciaremos e demonstraremos o Lema de Morse, além de discutir suas principais consequências. O Lema de Morse é um resultado central na Análise e na Topologia Diferencial, com papel importante na compreensão do comportamento local de funções diferenciáveis em torno de seus pontos críticos. De modo geral, o lema afirma que, sob certas condições de regularidade, uma função real de várias variáveis que possui um ponto crítico não degenerado, ou seja, um ponto em que o gradiente é nulo e a matriz Hessiana é invertível, pode ser escrita, após uma mudança adequada de coordenadas, como uma soma de quadrados com sinais positivos e negativos. Essa forma simplificada permite identificar o tipo do ponto crítico, indicando se ele é um mínimo, um máximo ou um ponto de sela.

O resultado leva o nome do matemático norte-americano Harold Calvin Marston Morse (1892–1977), responsável por desenvolver uma teoria profunda sobre o comportamento de funções suaves, hoje conhecida como Teoria de Morse. Morse teve uma carreira acadêmica notável, com passagens por instituições como Harvard e Princeton, e dedicou grande parte de suas pesquisas ao estudo da topologia de variedades diferenciáveis. Seu trabalho visava compreender como os pontos críticos de uma função determinam, de forma surpreendentemente precisa, aspectos da estrutura topológica do espaço onde essa função está definida. Essa abordagem abriu caminho para conexões entre análise, geometria e topologia, influenciando áreas modernas como física matemática, geometria simplética e sistemas dinâmicos. O lema que leva seu nome é uma das ferramentas mais fundamentais dentro desse contexto teórico. Inicialmente, traremos alguns resultados necessários para demonstração desse Lema.

Lema 3.1. *Seja $f: \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ definida por $f(X) = X^2$ onde X^2 é o quadrado da matriz X , de n linhas e n colunas. A aplicação f é de classe C^∞ e sua derivada, em cada ponto X , é a transformação linear $f'(X): \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$, dada por:*

$$f'(X) \cdot V = V \cdot X + X \cdot V.$$

Demonstração: Com efeito, note que

$$\begin{aligned} f(X + tV) - f(X) &= (X + tV)^2 - X^2 \\ &= (X + tV) \cdot (X + tV) - X^2 \\ &= X^2 + tX \cdot V + tV \cdot X + t^2V^2 - X^2 \\ &= t(X \cdot V + V \cdot X + tV^2) \end{aligned}$$

De onde obtemos

$$\begin{aligned} f'(X) \cdot V &= \lim_{t \rightarrow 0} \frac{f(X + tV) - f(X)}{t} \\ &= \lim_{t \rightarrow 0} [X \cdot V + V \cdot X + tV^2] = X \cdot V + V \cdot X. \end{aligned}$$

■

Observação 3.2. Note que se $X = I$, então $f'(I) \cdot V = 2V$. Além disso, temos que $f'(I)V = 0$ se, e somente se, $V = 0$. Ou seja, $f'(I): \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ é injetiva e pelo teorema do núcleo e da imagem $f'(I)$ é um isomorfismo.

Observação 3.3. Pelo Teorema da Aplicação Inversa 2.10, existem abertos U, W em \mathbb{R}^{n^2} , ambos contendo $I = f(I)$, tais que $f: U \rightarrow W$ é um difeomorfismo C^∞ . Isto significa que toda matriz Y suficientemente próxima da identidade ($Y \in W$) possui uma raiz quadrada X (isto é, $X^2 = Y$), a qual é única se impusermos que ela esteja suficientemente próxima da identidade ($X \in U$). Além disso, X é uma função de classe C^∞ de Y .

Teorema 3.4 (Lema de Morse). *Seja a um ponto crítico não degenerado de uma função $f: U \rightarrow \mathbb{R}$ de classe C^k ($k \geq 3$) num aberto $U \subset \mathbb{R}^n$. Existe um sistema de coordenadas $\xi: V \rightarrow W$ de classe C^{k-2} com $a \in W \subset U$, $0 \in V$ e $\xi(0) = a$ tal que*

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

para todo $y = (y_1, \dots, y_n) \in V$, onde $a_{ij} = \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$.

Demonstração: Afim de simplificar a notação, consideramos $a = 0$ e $f(a) = 0$, caso contrário, defina $g(x) = f(x) - a$ e o resultado seguirá de maneira análoga.

Seja W uma bola aberta de centro 0, contida em U . Pela fórmula de Taylor com resto integral, temos que se $x \in W$

$$f(x) = \sum_{i,j=1}^m a_{ij}(x) x_i x_j,$$

com

$$a_{ij}(x) = \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) dt. \quad (3.1)$$

De fato, pela fórmula de Taylor com resto integral

$$r_p(v) = \frac{1}{p!} \int_0^1 (1-t)^p f^{(p+1)}(a+tv) \cdot v^{p+1} dt.$$

Considerando $a = 0$ e $p = 1$ (p está relacionado ao grau do polinômio de Taylor e, nesse caso, com a ordem da derivada. $p = \text{ordem da derivada} - 1$).

Daí

$$\begin{aligned} r_1(x) &= \frac{1}{1!} \int_0^1 (1-t) f^{(1+1)}(tx) x^{1+1} dt \\ &= \int_0^1 (1-t) f^{(2)}(tx) x^2 dt. \end{aligned} \quad (3.2)$$

Note que $f^{(2)}(tx) = \frac{\partial^2 f}{\partial x_i \partial x_j}(tx)$. Representando na sua forma matricial

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(tx) = \begin{bmatrix} \frac{\partial^2 f}{x_1^2}(tx) & \frac{\partial^2 f}{x_1 x_2}(tx) & \cdots & \frac{\partial^2 f}{x_1 x_n}(tx) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{x_n x_1}(tx) & \frac{\partial^2 f}{x_n x_2}(tx) & \cdots & \frac{\partial^2 f}{x_n^2}(tx) \end{bmatrix}.$$

Substituindo em (3.2)

$$\begin{aligned} r_1(x) &= \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j} x^2 dt. \\ &= \int_0^1 (1-t) \begin{bmatrix} \frac{\partial^2 f}{x_1^2}(tx) & \frac{\partial^2 f}{x_1 x_2}(tx) & \cdots & \frac{\partial^2 f}{x_1 x_n}(tx) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{x_n x_1}(tx) & \frac{\partial^2 f}{x_n x_2}(tx) & \cdots & \frac{\partial^2 f}{x_n^2}(tx) \end{bmatrix} x^2 dt. \end{aligned}$$

Além disso, podemos reescrever x^2 de forma conveniente. Assim,

$$r_1(x) = \int_0^1 (1-t) \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{x_1^2}(tx) & \frac{\partial^2 f}{x_1 x_2}(tx) & \cdots & \frac{\partial^2 f}{x_1 x_n}(tx) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{x_n x_1}(tx) & \frac{\partial^2 f}{x_n x_2}(tx) & \cdots & \frac{\partial^2 f}{x_n^2}(tx) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} dt. \quad (3.3)$$

Note que a multiplicação dessas matrizes está bem definida. Daí

$$\begin{aligned} r_1(x) &= \int_0^1 (1-t) \left[\sum_{i=1}^n x_i \cdot \frac{\partial^2 f}{\partial x_i \partial x_1}(tx) \quad \cdots \quad \sum_{i=1}^n x_i \cdot \frac{\partial^2 f}{\partial x_i \partial x_n}(tx) \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} dt \\ &= \int_0^1 (1-t) \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) x_i x_j dt \\ &= \sum_{i,j=1}^n x_i x_j \left[\int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) dt \right] \end{aligned} \quad (3.4)$$

Veja que de (3.1) temos que $\int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) dt = a_{ij}(x)$ substituindo em (3.4) temos

$$\sum_{i,j=1}^n a_{ij} x_i x_j$$

Note que cada $a_{ij}(x)$ é uma função de classe C^{k-2} , definida na bola W . Assim, pelo

Teorema 1.55 a matriz $A(x) = (a_{ij}(x))$ é simétrica. Logo, podemos escrever

$$f(x) = \langle A(x) \cdot x, x \rangle \quad \forall x \in W \quad (3.5)$$

Seja $A_0 = A(0)$, daí

$$\begin{aligned} A_0 &= \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(0) dt \\ &= \frac{\partial^2 f}{\partial x_i \partial x_j}(0) \int_0^1 (1-t) dt \\ &= \frac{\partial^2 f}{\partial x_i \partial x_j}(0) \left[t - \frac{t^2}{2} \right]_0^1 = \frac{1}{2} \\ &= \frac{1}{2} \cdot \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(0) \right). \end{aligned}$$

Sendo 0 um ponto crítico não degenerado de f temos que a matriz simétrica A_0 é invertível. Com isto, usando o Teorema 2.10 concluímos que existe $\delta > 0$ de modo que $f : B_\delta \rightarrow f(B_\delta)$ é um difeomorfismo para cada $x \in B_\delta$ com $B_\delta = B_\delta(0)$. Assim, defina $C(x) = A_0^{-1}A(x)$, como A é de classe C^{k-2} em B_δ , C também é. Note que $C(0) = A_0^{-1}A_0 = Id$ e pela Observação 3.3 existe uma única matriz de classe C^{k-2} $B(x)$ tal que $B(x)^2 = C(x)$.

Daí, para cada $x \in B_\delta$

$$B(x)^2 = C(x) = A_0^{-1}A(x) \Rightarrow A(x) = A_0B(x)^2. \quad (3.6)$$

Note também que

$$A(0) = A_0B(0)^2 \Rightarrow B(0) = Id.$$

Assim, como A_0 e $A(x)$ são simétricas, ou seja $A_0^\top = A_0$ e $A(x)^\top = A(x)$, com isto

$$A_0B(x)^2 = (A_0B(x)^2)^\top = [B(x)^2]^\top A_0^\top = [B(x)^\top]^2 A_0.$$

Por simplicidade omitiremos a variável x . Portanto

$$A_0B^2 = (B^\top)^2 A_0. \quad (3.7)$$

Multiplicando ambos os lados de (3.7) à esquerda por A_0^{-1}

$$B^2 = A_0^{-1}(B^\top)^2 A_0. \quad (3.8)$$

O lado direito de (3.8) pode ser reescrito como um quadrado, pois

$$A_0^{-1}(B^\top)^2 A_0 = A_0^{-1} B^\top B^\top A_0 = (A_0^{-1} B^\top A_0)(A_0^{-1} B^\top A_0) = (A_0^{-1} B^\top A_0)^2.$$

Substituindo em (3.8)

$$B^2 = (A_0^{-1} B^\top A_0)^2. \quad (3.9)$$

Denotando $D(x) := A_0^{-1} B(x)^\top A_0$, por (3.9), temos

$$B(x)^2 = D(x)^2.$$

Como B e D são contínuas, para x suficientemente próximo de 0; $B(x)$ e $D(x)$ estão ambas próximas de I .

Além disso, pela observação 3.3, a função $A \mapsto A^2$ é injetiva numa vizinhança de I . Portanto, $B(x) = D(x)$, ou seja $B = A_0^{-1} B^\top A_0$, equivalentemente

$$A_0 B = B^\top A_0. \quad (3.10)$$

Substituindo (3.10) em (3.6)

$$A = A_0 B^2 = (B^\top A_0) B = B^\top A_0 B. \quad (3.11)$$

Temos que (3.5) é dado por

$$f(x) = \langle A(x) \cdot x, x \rangle.$$

Substituindo pela expressão do $A(x)$ obtido em (3.11) obtemos

$$f(x) = \langle (B(x)^\top \cdot A_0 \cdot B(x)) \cdot x, x \rangle$$

Associando e utilizando o Teorema 1.70

$$f(x) = \langle A_0 \cdot (B(x) \cdot x), B(x) \cdot x \rangle$$

Tomando $\varphi(x) = B(x) \cdot x$ obtemos

$$f(x) = \langle A_0 \varphi(x), \varphi(x) \rangle \quad (3.12)$$

Nosso objetivo agora é mostrar que se tomarmos um raio suficientemente pequeno da bola B_δ a aplicação $\varphi: B_\delta \rightarrow \mathbb{R}^n$, definida por $\varphi(x) = B(x) \cdot x$, é um difeomorfismo de classe C^{k-2} sobre sua imagem. Com efeito, para todo $x \in B_\delta$ e todo $v \in \mathbb{R}^n$. A derivada direcional de φ no ponto x na direção v é definida por

$$\varphi'(x) \cdot v = \frac{\partial \varphi}{\partial v}(x) = \lim_{t \rightarrow 0} \frac{\varphi(x + tv) - \varphi(x)}{t}.$$

Substituindo $\varphi(x) = B(x) \cdot x$

$$\varphi'(x) \cdot v = \lim_{t \rightarrow 0} \frac{B(x + tv) \cdot (x + tv) - B(x) \cdot x}{t}.$$

Expandindo o numerador

$$\begin{aligned} B(x + tv) \cdot (x + tv) - B(x) \cdot x &= B(x + tv) \cdot x + tB(x + tv) \cdot v - B(x) \cdot x \\ &= [B(x + tv) \cdot x - B(x) \cdot x] + tB(x + tv) \cdot v. \end{aligned}$$

Portanto

$$\begin{aligned} \varphi'(x) \cdot v &= \lim_{t \rightarrow 0} \frac{[B(x + tv) \cdot x - B(x) \cdot x] + tB(x + tv) \cdot v}{t} \\ &= \lim_{t \rightarrow 0} \left[\frac{B(x + tv) \cdot x - B(x) \cdot x}{t} + B(x + tv) \cdot v \right]. \end{aligned}$$

Separando em dois limites:

$$\varphi'(x) \cdot v = \underbrace{\lim_{t \rightarrow 0} \frac{B(x + tv) \cdot x - B(x) \cdot x}{t}}_{(I)} + \underbrace{\lim_{t \rightarrow 0} B(x + tv) \cdot v}_{(II)}.$$

Do termo (II)

Como B é contínua (pois é de classe C^{k-2} com $k \geq 3$), temos:

$$\lim_{t \rightarrow 0} B(x + tv) \cdot v = B(x) \cdot v. \quad (3.13)$$

Do termo (I)

$$\lim_{t \rightarrow 0} \frac{[B(x + tv) - B(x)] \cdot x}{t} = \left[\lim_{t \rightarrow 0} \frac{B(x + tv) - B(x)}{t} \right] \cdot x.$$

O limite entre colchetes é a derivada direcional da função matricial B no ponto x na direção v ,

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

Assim

$$\frac{\partial B}{\partial v}(x) \cdot x. \quad (3.14)$$

Juntando os termos (3.13) e (3.14)

$$\varphi'(x) \cdot v = \frac{\partial B}{\partial v}(x) \cdot x + B(x) \cdot v.$$

Se $x = 0$

$$\varphi'(0) \cdot v = \frac{\partial B}{\partial v}(0) \cdot 0 + B(0) \cdot v = 0 + Id \cdot v = v,$$

pois $B(0) = I$. Como $\varphi'(0) \cdot v = v$ para todo $v \in \mathbb{R}^n$, temos que $\varphi'(0) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um isomorfismo linear. Sendo φ de classe C^{k-2} com $k \geq 3$, pelo Teorema da Função Inversa 2.10, se o raio de B_δ for tomado suficientemente pequeno existem vizinhanças W de 0 e $V \subset \mathbb{R}^n$ de $\varphi(0) = 0$ tais que

$$\varphi : B_\delta \rightarrow V$$

é um difeomorfismo de classe C^{k-2} .

Assim, se $x \in B_\delta$, de (3.12) temos

$$f(x) = \langle A_0 \cdot \varphi(x), \varphi(x) \rangle.$$

Definimos então o sistema de coordenadas $\xi : V \rightarrow W$ como $\xi = \varphi^{-1}$. Este satisfaz:

1. ξ é de classe C^{k-2} ,
2. $\xi(0) = 0$,

3. Para todo $y \in V$,

$$f(\xi(y)) = \langle A_0 \cdot \varphi(\xi(y)), \varphi(\xi(y)) \rangle = \langle A_0 \cdot y, y \rangle = \sum_{i,j=1}^n a_{ij} y_i y_j.$$

■

Corolário 3.5. *Seja a um ponto crítico não degenerado de uma função $f : U \rightarrow \mathbb{R}$ de classe C^k , com $k \geq 3$, num aberto $U \subset \mathbb{R}^n$. Então, existe um sistema de coordenadas $\varphi : V \rightarrow W$, de classe C^∞ , com $W \subset U$, $0 \in V$, $\varphi(0) = a$ e tal que*

$$f(\varphi(y)) - f(a) = -y_1^2 - \dots - y_i^2 + y_{i+1}^2 + \dots + y_n^2.$$

Demonstração: Uma vez que $f \in C^k$, com $k \geq 3$, pelo Teorema de Schwarz 1.55, temos que a matriz $A = (a_{ij})$, onde cada a_{ij} é dado no Teorema 3.4, é uma matriz simétrica. Em particular, temos que A_0 é um operador auto-adjunto e, pelo Lema 1.69, existe uma base de ortonormal para o \mathbb{R}^n formada por autovetores de A_0 , isto é, existem números reais $\lambda_1, \dots, \lambda_n$ e vetores u_1, \dots, u_n , com $\|u_j\| = 1$ e $A_0 \cdot u_j = \lambda_j u_j$ para todo $j = 1, \dots, n$. Note que a menos de uma reordenação, podemos considerar, sem perda de generalidade, que $\lambda_1 < \lambda_2 < \dots < \lambda_n$. Além disso, como a é um ponto não degenerado de f temos que A_0 é um operador invertível, em particular, temos que cada $\lambda_j \neq 0$ para todo $j = 1, \dots, n$. Caso exista, considere $j_0 \in \{1, \dots, n\}$ de modo que, $\lambda_i < 0$ para todo $i \leq j_0$ e $\lambda_i > 0$ para todo $i > j_0$ e daí defina $v_j = u_j / \sqrt{-\lambda_j}$ se $j \leq j_0$ e $v_j = u_j / \sqrt{\lambda_j}$ se $j > j_0$. Com isto temos que $\{v_1, \dots, v_n\}$ é uma base para \mathbb{R}^n cumprindo as seguintes condições:

$$\langle A_0 \cdot v_j, v_k \rangle = C_{j,k} \langle A_0 \cdot u_j, u_k \rangle = \lambda_j C_{j,k} \langle u_j, u_k \rangle = 0$$

para todo $j \neq k$, onde $C_{j,k} = (\lambda_j \cdot \lambda_k)^{-\frac{1}{2}}$ se $j, k \leq j_0$ ou se $j, k > j_0$ e $C_{j,k} = (-\lambda_j \cdot \lambda_k)^{-\frac{1}{2}}$ se $j \leq j_0$ e $k > j_0$ ou $k \leq j_0$ e $j > j_0$. Também obtemos que

$$\langle A_0 \cdot v_k, v_k \rangle = \left(\frac{1}{\sqrt{-\lambda_k}} \right)^2 \langle A_0 \cdot u_k, u_k \rangle = \frac{-1}{\lambda_k} \lambda_k \langle u_k, u_k \rangle = -1 \quad \text{se } k \leq j_0$$

e

$$\langle A_0 \cdot v_k, v_k \rangle = \left(\frac{1}{\sqrt{\lambda_k}} \right)^2 \langle A_0 \cdot u_k, u_k \rangle = \frac{1}{\lambda_k} \lambda_k \langle u_k, u_k \rangle = 1 \quad \text{se } k > j_0$$

Agora considere a transformação linear invertível $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $T(e_k) = v_k$ para

todo $k = 1, \dots, n$ e $V_0 = T^{-1}(V)$ onde V é o aberto obtido no Teorema 3.4 e, por fim, defina

$$\psi : V_0 \rightarrow W; \quad \psi(y) = \xi(T(y)).$$

Note que ψ é um difeomorfismo satisfazendo

$$\begin{aligned} f(\psi(y)) &= f(\xi(T \cdot y)) = \langle A_0 \cdot T \cdot y, T \cdot y \rangle = \langle A_0 \cdot \sum_{j=1}^n y_j v_j, \sum_{k=1}^n y_k v_k \rangle \\ &= \sum_{j,k=1}^n y_j y_k \langle A_0 \cdot v_j, v_k \rangle = -y_1^2 - \dots - y_{j_0}^2 + y_{j_0+1}^2 + \dots + y_n^2 \end{aligned}$$

■

O número i que aparece no Corolário acima chama-se o índice do ponto crítico e determina sua natureza (mínimo, máximo ou ponto de sela). Onde i é o número de autovalores negativos da Hessiana em a .

- Se todos os autovalores são negativos, então a forma quadrática associada é negativa definida, e o ponto crítico a é um máximo local.
 - Nesse caso, o índice é $i = n$ (todos os autovalores negativos).
- Se todos os autovalores são positivos, então a forma quadrática é positiva definida, e o ponto crítico a é um mínimo local.
 - Nesse caso, o índice é $i = 0$ (nenhum autovalor negativo).
- Se alguns autovalores são negativos e outros positivos, então a forma quadrática é indefinida, e o ponto crítico a é um ponto de sela.
 - Nesse caso, o índice satisfaz $0 < i < n$ (mistura de sinais).

Vejamos dois exemplos.

Exemplo 3.6. Considere a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por:

$$f(x, y) = x^2 - y^2 + xy^2.$$

Note que a função é um polinômio, portanto de classe C^∞ , e em particular de classe C^k para qualquer $k \geq 3$. Vamos agora encontrar o ponto crítico, para isso vamos calcular o

gradiente e igualar a 0.

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (0, 0).$$

As derivadas parciais são

$$\frac{\partial f}{\partial x} = 2x + y^2, \quad \frac{\partial f}{\partial y} = -2y + 2xy.$$

Resolvemos o sistema

$$\begin{cases} 2x + y^2 = 0, \\ -2y + 2xy = 0. \end{cases}$$

Da segunda equação:

$$-2y + 2xy = 2y(x - 1) = 0 \Rightarrow y = 0 \quad \text{ou} \quad x = 1.$$

Se $y = 0$, da primeira equação temos $x = 0$. Portanto, $(0, 0)$ é um ponto crítico.

Agora vamos verificar que o ponto crítico é não degenerado.

Calculamos a matriz Hessiana de f

$$H_f(x, y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}.$$

As derivadas de segunda ordem são

$$f_{xx} = 2, \quad f_{xy} = \frac{\partial}{\partial y}(2x + y^2) = 2y, \quad f_{yy} = \frac{\partial}{\partial y}(-2y + 2xy) = -2 + 2x.$$

Avaliando em $(0, 0)$, temos

$$Hf(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Essa matriz é diagonal com determinante

$$\det(Hf(0, 0)) = 2 \cdot (-2) = -4 \neq 0.$$

Portanto, o ponto crítico $(0, 0)$ é não degenerado.

Utilizamos a definição

$$H \cdot v^2 = \sum_{i,j=1}^2 h_{ij} \alpha_i \alpha_j,$$

onde $v = (\alpha_1, \alpha_2) = (x, y)$, e $H = H_f(0, 0)$. Portanto

$$H \cdot v^2 = h_{11}x^2 + h_{12}xy + h_{21}yx + h_{22}y^2.$$

Como $H = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$, temos:

$$H \cdot v^2 = 2x^2 + 0 \cdot xy + 0 \cdot yx - 2y^2 = 2x^2 - 2y^2.$$

Dividindo por 2, temos

$$Q(x, y) = x^2 - y^2.$$

Como $Q(x, y)$ é indefinida, pela Observação 1.61 e pelo item (iii) do Teorema 1.66 como $(0, 0)$ é ponto crítico de f , $(0, 0)$ não é ponto de máximo e nem de mínimo local (ponto de sela).

O índice de Morse é 1, pois há um termo negativo.

Pelo Lema de Morse, existe uma mudança de coordenadas local $\varphi : V \rightarrow \mathbb{R}^2$, de classe C^{k-2} , com $\varphi(0) = (0, 0)$, tal que

$$f(\varphi(y)) = f(0) + Q(y) = x^2 - y^2.$$

Ou seja

$$f(\varphi(y)) - f(0) = -y_1^2 + y_2^2.$$

Este exemplo mostra que, sob as hipóteses do Lema de Morse, uma função suave pode ser transformada localmente em uma forma quadrática canônica via uma mudança de coordenadas suave.

Exemplo 3.7. Seja a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por

$$f(x, y, z) = -x^2 - y^2 + z^2 + xyz.$$

Note que f é uma função polinomial, logo $f \in C^\infty$,

$$\begin{aligned}\frac{\partial f}{\partial x} &= -2x + yz \\ \frac{\partial f}{\partial y} &= -2y + xz \\ \frac{\partial f}{\partial z} &= 2z + xy\end{aligned}$$

Resolvemos o sistema

$$\begin{cases} -2x + yz = 0 \\ -2y + xz = 0 \\ 2z + xy = 0 \end{cases}$$

O ponto $(0, 0, 0)$ satisfaz todas as equações. Logo, $(0, 0, 0)$ é ponto crítico.

Calculamos as segundas derivadas

$$\begin{aligned}f_{xx} &= -2, & f_{xy} &= z, & f_{xz} &= y \\ f_{yy} &= -2, & f_{yz} &= x, & f_{zz} &= 2\end{aligned}$$

No ponto $(0, 0, 0)$, temos

$$Hf(0, 0, 0) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Note que

$$\det(Hf(0, 0, 0)) \neq 0$$

. Portanto, $(0, 0, 0)$ é um ponto crítico não degenerado.

Seja $v = (x, y, z) \in \mathbb{R}^3$. A forma quadrática é

$$H \cdot v^2 = -2x^2 - 2y^2 + 2z^2.$$

Dividindo por 2

$$Q(x, y, z) = -x^2 - y^2 + z^2.$$

A matriz Hessiana é simétrica e não degenerada. Os autovalores são $-2, -2, 2$, então o índice de Morse é 2.

Logo, pelo Lema de Morse, existe uma mudança de coordenadas φ de classe C^{k-2}

tal que:

$$f(\varphi(y)) - f(0) = -y_1^2 - y_2^2 + y_3^2.$$

Considerações Finais

Este trabalho teve como objetivo principal estudar o Teorema da Função Inversa e, a partir dele, demonstrar o Lema de Morse em \mathbb{R}^n . Para isso, construímos as ferramentas necessárias, como os conceitos de topologia do espaço n -dimensional, a definição de limites, derivadas, o Teorema de Schwarz, a matriz Hessiana e a forma quadrática associada.

Mostramos que, sob hipóteses adequadas, uma função C^k , com $k \geq 3$, pode ser reescrita localmente como uma forma quadrática por meio de uma mudança de coordenadas de classe C^{k-2} . Essa reformulação, garantida pelo Lema de Morse, revela a estrutura do ponto crítico e sua relação com o índice da matriz Hessiana.

O trabalho evidenciou a importância do Teorema da Função Inversa como ponto de partida para a compreensão do comportamento local de funções diferenciáveis.

Referências Bibliográficas

- [1] BEZERRA, B. V. B. **O teorema da função inversa e sua relação com as superfícies regulares**. Recife: Universidade Federal Rural de Pernambuco, 2023. (Trabalho de Conclusão de Curso – Licenciatura em Matemática).
- [2] BOLDRINI, J. L. ET AL. **Álgebra Linear**. 3 ed. São Paulo: Harbra, 1986.
- [3] LIMA, E. L. **Análise Real, vol. 1**. 9 ed. Rio de Janeiro: IMPA, 2007. (Coleção Matemática Universitária).
- [4] LIMA, E. L. **Análise Real, vol. 2**. 6 ed. Rio de Janeiro: IMPA, 2012. (Coleção Matemática Universitária).
- [5] LIMA, E. L. **Curso de Análise, vol. 1**. 15 ed. Rio de Janeiro: IMPA, 2022. (Coleção Projeto Euclides).
- [6] LIMA, E. L. **Curso de Análise, vol. 2**. 9 ed. Rio de Janeiro: IMPA, 2006. (Coleção Projeto Euclides).
- [7] PITCHER, E. **Marston Morse (1892–1977): Biographical Memoir**. Washington, D.C.: National Academy of Sciences, 1994.