

Análise em \mathbb{R}^d

Hugo Cattarucci Botós (IME - USP)
hugocbotos@gmail.com

Última atualização em 26 de maio de 2024

Copyright © 2024 de Hugo Cattarucci Botós. Autorizo reprodução e distribuição do texto para fins não-lucrativos desde que a autoria seja citada. Se encontrar erros, agradeço se me notificar via e-mail: hugocbotos@gmail.com.

Sumário

| | |
|--|-----------|
| 1 Beabá topológico | 3 |
| 1.1 Revisão de sequências na reta | 3 |
| 1.2 Normas e métricas | 6 |
| 1.3 Topologia de espaço métrico | 18 |
| 1.4 Sequência em espaços métricos | 21 |
| 1.5 Continuidade | 24 |
| 1.6 Homeomorfismos | 28 |
| 1.7 Conexidade | 30 |
| 1.8 Compacidade | 36 |
| 1.9 Sequências de Cauchy | 38 |
| 2 Diferenciabilidade | 41 |
| 2.1 Curvas diferenciáveis | 41 |
| 2.2 Funções diferenciáveis | 46 |
| 2.3 Fórmula de Taylor | 55 |
| 2.4 Convexidade e pontos críticos | 59 |
| 3 Variedades em \mathbb{R}^p | 67 |
| 3.1 Teorema da função inversa | 67 |
| 3.2 Variedades mergulhadas em \mathbb{R}^p e espaço tangente | 70 |
| 3.3 Multiplicadores de Lagrange | 80 |
| 4 Convergência uniforme | 85 |
| 5 Integral de Riemann | 93 |
| 5.1 Propriedades básicas da integral de Riemann. | 93 |
| 5.2 Volume e Mudança de variável | 101 |
| 5.2.1 Volume de paralelogramos | 103 |
| 5.3 Teorema de mudança de variável | 105 |
| 5.4 Teorema da divergência | 111 |

Capítulo 1

Beabá topológico

The introduction of the cipher 0 or the group concept was general nonsense too, and mathematics was more or less stagnating for thousands of years because nobody was around to take such childish steps ...

Alexander Grothendieck

1.1 Revisão de sequências na reta.

Uma sequência em um conjunto X é uma função $\mathbb{Z}_{\geq 1} \rightarrow X$ associa a cada inteiro positivo n um ponto x_n de X . Uma sequência x_n em \mathbb{R} é formada por números $x_1, x_2, x_3, \dots, x_n, \dots$ e converge para $x \in \mathbb{R}$ se temos $|x - x_n|$ tão pequeno quanto quisermos para n suficientemente grande. Mais precisamente, dizemos que x_n converge para x se para cada $\varepsilon > 0$ existe n_0 tal que $|x_n - x| < \varepsilon$ para $n \geq n_0$. Denotamos essa convergência por $x_n \rightarrow x$ ou $\lim_{n \rightarrow \infty} x_n = x$.

Exercício 1.1 (Desigualdade de Bernoulli). Prove que para $n \in \mathbb{Z}_{\geq 0}$ e $x > -1$ temos

$$(1 + x)^n \geq 1 + nx.$$

A desigualdade de Bernoulli é bastante útil quando mostrando limites numéricos básicos.

Exemplo 1.2. Considere $a \in \mathbb{R}_{>1}$ e $x_n = a^{\frac{1}{n}}$. Temos que $x_n \rightarrow 1$. De fato, considere $h_n = a^{\frac{1}{n}} - 1$ e note que $h_n \geq 0$. Pela desigualdade de Bernoulli 1.1,

$$a = (1 + h_n)^n \geq 1 + nh_n,$$

ou seja,

$$\begin{aligned} 0 \leq h_n &\leq \frac{a-1}{n} \rightarrow 0, \\ a^{\frac{1}{n}} &= 1 + h_n \rightarrow 1. \end{aligned}$$

Por outro lado, se $0 < a < 1$, temos

$$a^{\frac{1}{n}} = \frac{1}{(a^{-1})^{\frac{1}{n}}} \rightarrow 1,$$

pois $a^{-1} > 1$.

Exercício 1.3. Mostre que $n^{1/n} \rightarrow 1$.

Vamos supor que o leitor saiba que conjuntos limitados superiormente em \mathbb{R} possuem supremo, e conjuntos limitados inferiormente possuem ínfimo. Esses fatos são normalmente vistos em um curso básico de análise na reta.

Exercício 1.4. Mostre que

- se x_n é crescente e limitada, então $x_n \rightarrow \sup\{x_i : i \geq 1\}$;
- se x_n é decrescente e limitada, então $x_n \rightarrow \inf\{x_i : i \geq 1\}$.

Exemplo 1.5. Considere $a_n = (1 + \frac{1}{n})^n$. Mostremos que a_n é crescente. A seguir, fatoraremos a_{n+1}/a_n até termos uma expressão apropriada para aplicar a desigualdade de Bernoulli 1.1, assim mostrando que $\frac{a_{n+1}}{a_n} \geq 1$.

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^n} \\ &= \left(1 + \frac{1}{n}\right) \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^{n+1}} \\ &= \left(1 + \frac{1}{n}\right) \left(\frac{(n+2)n}{(n+1)^2}\right)^{n+1} \\ &= \left(1 + \frac{1}{n}\right) \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \\ &\geq \left(1 + \frac{1}{n}\right) \left(1 - (n+1)\frac{1}{(n+1)^2}\right) \\ &= \frac{n+1}{n} \frac{n}{n+1} = 1.\end{aligned}$$

Portanto, $a_{n+1} \geq a_n$, ou seja, a sequência a_n é crescente. Se essa for limitada superiormente, teremos que $a_n \rightarrow \sup\{a_i : i \geq 1\}$, que é o número e .

Mostremos que a_n é limitada superiormente. Pelo binômio de Newton

$$\begin{aligned}a_n &= \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} \\ &= \sum_{k=0}^n \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{1}{n^k} \\ &= \sum_{k=0}^n \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \frac{1}{k!} \\ &\leq \sum_{k=0}^n \frac{1}{k!} \\ &= 1 + 1 + \frac{1}{2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{2 \cdot 3 \cdots n} \\ &\leq 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} \\ &\leq 3.\end{aligned}$$

Logo, a_n converge para $\sup\{a_i : i \geq 1\}$.

Exercício 1.6. Sejam x_n e y_n sequências em \mathbb{R} convergentes para x e y e seja $\alpha \in \mathbb{R}$ uma constante. Mostre que

$$x_n + y_n \rightarrow x + y, \quad \alpha x_n \rightarrow \alpha x, \quad x_n y_n \rightarrow xy.$$

Mostre que para $y \neq 0$, temos que y_n é não nulo par n suficientemente grande e

$$\frac{x_n}{y_n} \rightarrow \frac{x}{y}.$$

Proposição 1.7. Considere uma função $f : I \rightarrow \mathbb{R}$, onde I é um intervalo. São equivalentes:

- f é contínua em x_0 , i.e., dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$|f(x) - f(x_0)| < \varepsilon \quad \text{quando} \quad |x - x_0| < \delta.$$

- Para toda sequência convergente $x_n \rightarrow x_0$, temos $f(x_n) \rightarrow f(x_0)$.

Exercício 1.8. Prove a Proposição 1.7.

Definição 1.9. Uma subsequência de x_n é uma sequência

$$x_{n_1}, x_{n_2}, x_{n_3}, \dots$$

com

$$n_1 < n_2 < n_3 < \dots$$

Exercício 1.10. Considere uma sequência x_n e um ponto x . Mostre que $x_n \rightarrow x$ se as subsequências x_{2k} e x_{2k+1} convergem para x .

Extra: Utilize tal resultado para provar o teste de convergência de Leibniz para séries: se a_k é uma sequência decrescente de termos não-negativos e $a_k \rightarrow 0$, então

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n (-1)^{k+1} a_k$$

converge.

Teorema 1.11 (Teorema de Bolzano-Weierstrass). Toda sequência limitada admite subsequência convergente.

Demonstração. Considere a sequência x_n limitada. Há $R > 0$ tal que $x_n \in [-R, R]$. Seja $I_0 = [-R, R]$. Dividindo I_0 ao meio, obtemos dois intervalos fechados. Escolha I_1 como sendo uma das metades possuindo x_n para infinitos n 's. Temos que $\{n \in \mathbb{Z}_{\geq 1} : x_n \in I_1\}$ é infinito e o comprimento de I_1 , denotado por $|I_1|$, vale $|I_1| = |I_0|/2$. Dividindo I_1 ao meio, escolhemos I_2 tal que $\{n \in \mathbb{Z}_{\geq 1} : x_n \in I_2\}$ é infinito e $|I_2| = |I_1|/2 = |I_0|/2^2$. Seguindo tal algoritmo, obtemos os intervalos fechados e limitados I_0, I_1, I_2, \dots tais que

- $\{n \in \mathbb{Z}_{\geq 1} : x_n \in I_i\}$ é infinito,
- $I_0 \supset I_1 \supset I_2 \supset \dots \supset I_i \supset I_{i+1} \supset \dots$,
- $|I_i| = |I_0|/2^i$.

Escolha $x_{n_1} \in I_1$. Existe $n_2 > n_1$ tal que $x_{n_2} \in I_2$, pois $\{n \in \mathbb{Z}_{\geq 1} : x_n \in I_2\}$ é infinito. Similarmente, existe $n_3 > n_2$ tal que $x_{n_3} \in I_3$. Seguindo tal procedimento, existem $n_1 < n_2 < n_3 < \dots < n_i < n_{i+1} < \dots$ tais que $x_{n_i} \in I_i$.

Como I_i é um intervalo fechado e limitado, temos que esse possui extremos a_i, b_i , isto é, $I_i = [a_i, b_i]$. Adicionalmente, temos que $I_{i+1} \subset I_i$ e, consequentemente, $a_i \leq a_{i+1}$ e $b_i \geq b_{i+1}$. Além disso, a sequência a_i é limitada superiormente por R e a sequência b_i é limitada inferiormente por $-R$, ou seja, os números

$$\alpha := \sup\{a_i : i \in \mathbb{Z}_{\geq 1}\} \quad \text{e} \quad \beta := \inf\{b_i : i \in \mathbb{Z}_{\geq 1}\}$$

existem.

Assim, $a_n \rightarrow \alpha$ e $b_n \rightarrow \beta$, por serem sequências monótonas. Como $b_n - a_n = |I_0|/2^n \rightarrow 0$, temos que $\alpha = \beta$, ou seja, $a_n, b_n \rightarrow \alpha$.

Por outro lado, $a_k \leq x_{n_k} \leq b_k$, pois $x_{n_k} \in I_k$. Deste modo,

$$a_k - \alpha \leq x_{n_k} - \alpha \leq b_k - \alpha,$$

ou seja,

$$\lim_{k \rightarrow \infty} x_{n_k} = \alpha.$$

■

Exemplo 1.12. A sequência $x_n = (-1)^n$ não converge, mas é limitada. Temos as subsequências convergentes $x_{2k} \rightarrow 1$ e $x_{2k+1} \rightarrow -1$.

Exercício 1.13. Considere $f : [a, b] \rightarrow \mathbb{R}$ contínua. Mostre que:

- f é limitada;
- existe $x_0 \in [a, b]$ tal que $f(x_0) = \sup\{f(x) : x \in [a, b]\}$, ou seja, f atinge máximo.

1.2 Normas e métricas

Comecemos com \mathbb{R}^d . Temos o produto interno canônico

$$\langle x, y \rangle = \sum_i x^i y^i,$$

onde $x = (x^1, \dots, x^d)$ e $y = (y^1, \dots, y^d)$.

Esse produto nos fornece uma norma definida pela fórmula

$$|x| = \sqrt{\langle x, x \rangle}.$$

Mais explicitamente,

$$|x| = \sqrt{\sum_i (x^i)^2}.$$

Uma norma, se o leitor não viu ou não lembra, tem de satisfazer as seguintes propriedades:

- $|kx| = |k||x|$ para $k \in \mathbb{R}$ e $x \in V$;
- $|x| \geq 0$ para qualquer $x \in V$ e somente vale zero se $x = 0$;
- Vale a desigualdade triangular: Para $x, y \in \mathbb{R}^d$, temos que

$$|x + y| \leq |x| + |y|.$$

Geometricamente, essa última propriedade é interpretada da seguinte forma. Se temos um triângulo com lados $x, y, x + y$, então a soma do comprimento de dois lados é maior ou igual que o comprimento do terceiro lado. De jeito mais simplista. Se queremos ir de 0 a $x + y$, então podemos ir direto, percorrendo a distância $|x + y|$, ou ir passando por x . Isto é, vamos de 0 a x percorrendo a distância $|x|$ e depois de x a $x + y$ percorrendo a distância $|y|$, porque $x + y - x = y$.

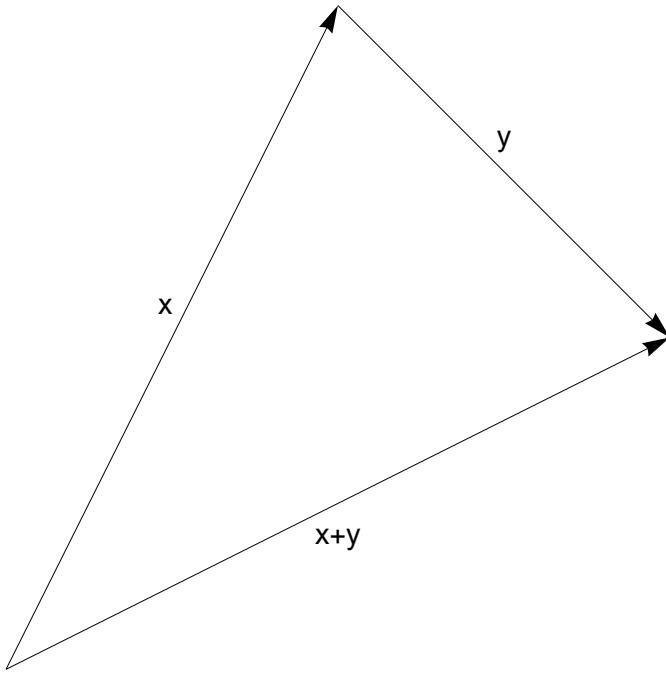


Figura 1.1: Triângulo com vetores x , y e $x + y$.

De forma mais geral, temos:

Definição 1.14. Uma norma em um espaço vetorial V sobre $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} é uma função $p : V \rightarrow \mathbb{R}$ que satisfaz as três propriedades listadas acima.

- $p(kx) = |k|p(x)$ para $k \in \mathbb{K}$ e $x \in V$;
- $p(x) \geq 0$ para qualquer $x \in V$ e somente vale zero se $x = 0$;
- Vale a desigualdade triangular: Para $x, y \in V$, temos que

$$p(x + y) \leq p(x) + p(y).$$

O espaço V munido da p é chamado de espaço vetorial normado e é denotado por $(V, |\cdot|)$.

Voltemos a \mathbb{R}^d . Repare que as duas primeiras propriedades de normas são evidentes, mas a terceira não é. A fim de prova-la precisamos da desigualdade de Cauchy-Schwarz. Primeiro provemos o “teorema” de Pitágoras.

Definição 1.15. Dois vetores $x, y \in \mathbb{R}^d$ são ortogonais (ou perpendiculares) se $\langle x, y \rangle = 0$.

Repare que o vetor nulo é ortogonal a todo vetor.

Lema 1.16 (Teorema de Pitágoras). Considere $x, y \in \mathbb{R}^d$. Temos que x, y são ortogonais se, e só se,

$$|x + y|^2 = |x|^2 + |y|^2.$$

Demonstração. Basta notar que

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

■

Teorema 1.17 (Desigualdade de Cauchy-Schwartz). Se $x, y \in \mathbb{R}^d$, então

$$|\langle x, y \rangle| \leq |x||y|.$$

A igualdade vale se x, y são linearmente dependentes.

Demonstração. Se $x = 0$ ou $y = 0$, então $\langle x, y \rangle = 0$ e o resultado é trivialmente válido. Podemos desta forma assumir que x e y são não nulos. Considere o vetor

$$\Pi[y]x := \frac{\langle x, y \rangle}{\langle y, y \rangle} y,$$

que se chama projeção de x sobre a reta $\mathbb{R}y$.

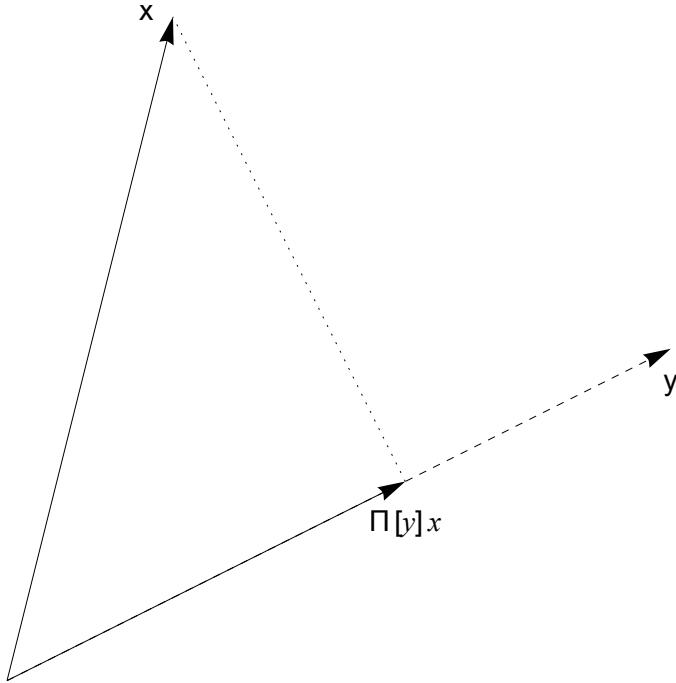


Figura 1.2: Projeção de x sobre y .

Note que $x - \Pi[y]x$ e $\Pi[y]x$ são ortogonais:

$$\langle x - \Pi[y]x, \Pi[y]x \rangle = \left\langle x - \frac{\langle x, y \rangle}{\langle y, y \rangle} y, \frac{\langle x, y \rangle}{\langle y, y \rangle} y \right\rangle = 0.$$

Assim, pelo Teorema de Pitágoras 1.16, obtemos

$$|x|^2 = |x - \Pi[y]x|^2 + |\Pi[y]x|^2.$$

Temos que

$$\begin{aligned} |x - \Pi[y]x|^2 &= |x|^2 - |\Pi[y]x|^2 \\ &= |x|^2 - \langle x, y \rangle^2 |y|^2 \\ &= \frac{1}{|y|^2} \left(|x|^2 |y|^2 - \langle x, y \rangle^2 \right). \end{aligned}$$

Como $|x - \Pi[y]x| \geq 0$ e somente é zero quando x, y são proporcionais, obtemos o resultado. ■

Observação 1.18. Todos esses resultados valem para espaços vetoriais reais ou complexos com produto interno. As provas são essencialmente as mesmas com a tecnicidade de que no caso complexo você obtém $|\operatorname{Re} \langle x, y \rangle| \leq |x||y|$ no argumento acima. Nesse cenário se usa o seguinte truque: o número $\xi = \frac{\langle x, y \rangle}{|\langle x, y \rangle|}$ é complexo de norma 1. Assim,

$$\langle \bar{\xi}x, y \rangle = |\langle x, y \rangle|$$

e obtemos

$$|\langle x, y \rangle| = |\operatorname{Re} \langle \bar{\xi}x, y \rangle| \leq |\bar{\xi}x||y| = |\bar{\xi}||x||y| = |x||y|,$$

que é o resultado desejado.

Da desigualdade de Cauchy-Schwarz 1.17, temos que para $x, y \in \mathbb{R}^d$ não nulos vale

$$-1 \leq \frac{\langle x, y \rangle}{|x||y|} \leq 1$$

e assim há $\theta \in [0, \pi]$ tal que $\langle x, y \rangle = |x||y| \cos(\theta)$.

O número θ é unicamente determinado, porque a função cosseno $\cos : [0, \pi] \rightarrow [-1, 1]$ é bijetora. Além disso, dizemos que θ é o ângulo entre x e y . Note também que para x, y não nulos, temos x, y ortogonais se, e só se, $\theta = \frac{\pi}{2}$.

Exercício 1.19. Mostre a Lei dos Cossenos. Se x, y são vetores não-nulos e θ é o ângulo entre eles, então

$$|x - y|^2 = |x|^2 + |y|^2 - 2|x||y| \cos(\theta).$$

Finalmente podemos provar a terceira propriedade da definição de norma para \mathbb{R}^d , a desigualdade triangular. Basta usarmos a desigualdade de Cauchy-Schwarz 1.17.

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

$$|x + y| \leq |x| + |y|.$$

Exercício 1.20. Considere o espaço M_d das matrizes quadradas reais $d \times d$. Mostre que

- $\langle A, B \rangle = \operatorname{tr}(AB^T)$ é um produto interno e $\mathbf{p}(A) = \sqrt{\operatorname{tr}(AA^T)}$ é uma norma em M_d ;
- $\mathbf{q}(A) = \sup\{|Ax| : x \in \mathbb{R}^d, |x| = 1\}$ é uma norma em M_d .

Agora definimos métrica (que é a mesma coisa que distância).

Definição 1.21. Uma métrica $d : X \times X \rightarrow \mathbb{R}$ em um conjunto X é uma função satisfazendo as seguintes propriedades:

- **Positividade:** $d(x, y) \geq 0$ para quaisquer $x, y \in X$ e vale zero quando $x = y$;
- **Simetria:** $d(x, y) = d(y, x)$ para quaisquer $x, y \in X$;
- **Desigualdade triangular:** $d(x, z) \leq d(x, y) + d(y, z)$ para quaisquer $x, y, z \in X$.

O conjunto X munido de uma métrica d é chamado de **espaço métrico**, que denotamos por (X, d) .

O autor desse texto as vezes esquece de escrever coisas. Ao invés de escrever (X, d) , muitas vezes escreverei X para denotar um espaço métrico, quando não houver confusão. As vezes denotamos a métrica de X por d_X quando for conveniente.

Se V é um espaço vetorial normado sobre \mathbb{R} ou \mathbb{C} com norma p , então $d(x, y) := p(x - y)$ é uma métrica. Isso segue direto das propriedades de norma.

Para \mathbb{R}^d , temos que a distância entre x e y é

$$d(x, y) = \sqrt{\sum_i (x^i - y^i)^2}.$$

Temos assim que todo espaço normado é naturalmente um espaço métrico.

Definição 1.22. Seja X um espaço métrico com métrica $d_X : X \times X \rightarrow \mathbb{R}$ e $A \subset X$. O conjunto A é espaço métrico com métrica $d_A : A \times A \rightarrow \mathbb{R}$ dada por $d_A(x, y) := d_X(x, y)$. Em outras palavras, a métrica de A é só a métrica de X (a métrica ambiente) restrita a A . Essa métrica em A é conhecida como métrica induzida de X em A .

Assim, todo subconjunto de \mathbb{R}^d tem é um espaço métrico.

Temos a reta $\mathbb{R} \times \{0\}$ em \mathbb{R}^2 . A distância é dada por

$$d_{\mathbb{R} \times \{0\}}((x, 0), (y, 0)) = \sqrt{(x - y)^2 + (0 - 0)^2} = |x - y|$$

que é a distância entre x, y na reta real \mathbb{R} . Assim, a métrica induzida nos eixos coordenados é simplesmente a métrica vinda de \mathbb{R} .

Pela mesma lógica, se temos $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^n \times \mathbb{R}^{d-n}$, onde $n < d$, temos que a métrica em $\mathbb{R}^n \times \{0\}$ é exatamente a que se espera, a vinda de \mathbb{R}^n .

O círculo unitário

$$\mathbb{S}^1 := \{(x^1, x^2) \in \mathbb{R}^2 : (x^1)^2 + (x^2)^2 = 1\}$$

tem métrica induzida de \mathbb{R}^2 . A distância entre $x, y \in \mathbb{S}^1$ é o comprimento do segmento de reta ligando x a y em \mathbb{R}^2 , o que não é exatamente a métrica que vem à mente quando pensamos em um círculo. A distância mais óbvia seria o comprimento $d'(x, y)$ do menor arco entre x, y .

Essa distância $d'(x, y)$ é ângulo entre x, y , porque o círculo tem raio 1:

$$d'(x, y) = \arccos(\langle x, y \rangle) \in [0, \pi],$$

onde $\arccos : [-1, 1] \rightarrow [0, \pi]$ é inversa da função $\cos : [0, \pi] \rightarrow [-1, 1]$.

Proposição 1.23. Considere $x, y \in \mathbb{S}^1$.

$$|x - y| \leq d'(x, y) \leq \frac{\pi}{2}|x - y|.$$

Demonstração. Se $x = y$ ou $x = -y$ a desigualdade é válida. Nesse segundo temos $|x - y| = 2$ e $d'(x, y) = \pi$.

Agora suponha $x \neq y$ e $x \neq -y$. Temos o triângulo $0, x, y$. Seja θ o ângulo entre x e y , isto é, $\theta = d'(x, y)$. Portanto, lembrando que $\cos(\theta) = 1 - 2\sin^2(\frac{\theta}{2})$, temos que

$$|x - y|^2 = 2 - 2\langle x, y \rangle = 2(1 - \cos(\theta)) = 4\sin^2(\theta/2).$$

Provemos que para $s \in [0, \frac{\pi}{2}]$ valem as desigualdades

$$\frac{2}{\pi}s \leq \sin(s) \leq s.$$

A função $s \mapsto s - \sin(s)$ vale 0 em $s = 0$ e tem derivada $1 - \cos(s)$, que é não negativa, ou seja, a função é crescente. Daí segue que $s - \sin(s) \geq 0$. Por outro lado, a função $\sin(s)$ é concava (sua segunda derivada é não positiva) no intervalo $[0, \pi]$, ou seja, a parte inferior de seu gráfico é convexa. Assim, a reta ligando $(0, 0)$ a $(\frac{\pi}{2}, 1)$, dada por $s \mapsto \frac{2}{\pi}s$, está abaixo do seu gráfico para $0 \leq s \leq \frac{\pi}{2}$.

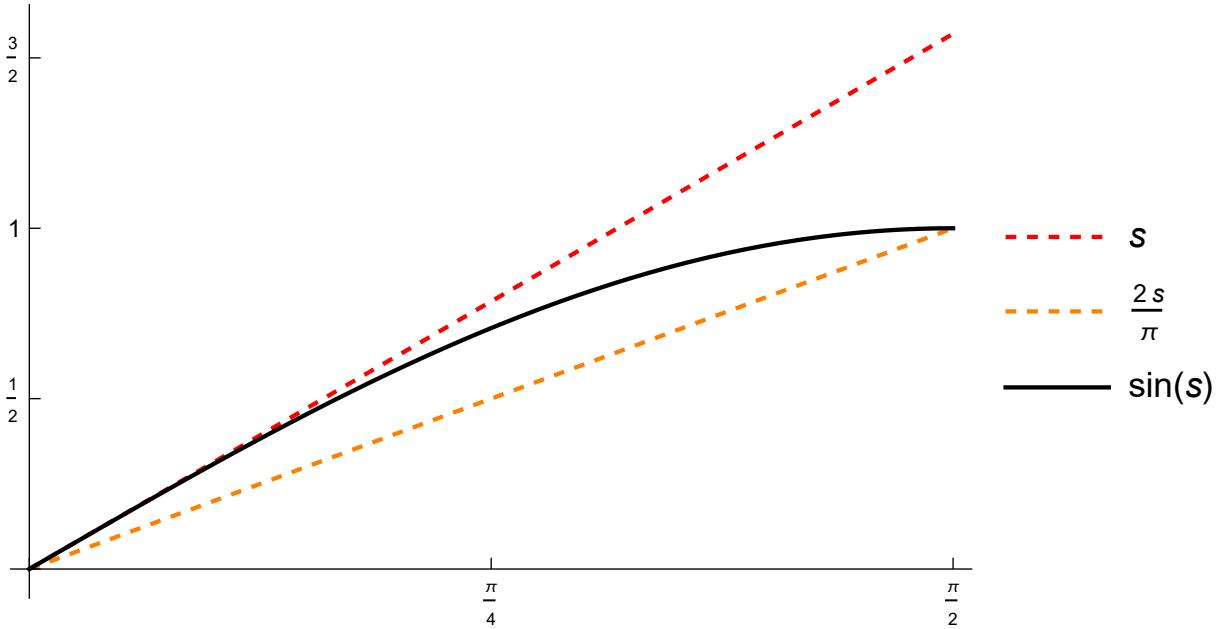


Figura 1.3: $\frac{2}{\pi}s \leq \sin(s) \leq s$ para $0 \leq s \leq \frac{\pi}{2}$.

Tomando $s = \theta/2$ temos:

$$\frac{4}{\pi^2}\theta^2 \leq |x - y|^2 \leq \theta^2,$$

$$\frac{2}{\pi}d'(x, y) \leq |x - y| \leq d'(x, y),$$

$$|x - y| \leq d'(x, y) \leq \frac{\pi}{2}|x - y|.$$

■

Observação 1.24. A esfera unitária d -dimensional $\mathbb{S}^d = \{x \in \mathbb{R}^{d+1} : |x| = 1\}$ satisfaz a mesma propriedade. Temos a métrica $d'(x, y) = \arccos(\langle x, y \rangle)$ que calcula o comprimento do arco ligando x a y . Mais precisamente, se $x \neq y$, temos círculo $\mathbb{S}^d \cap (\mathbb{R}x + \mathbb{R}y)$ e $d'(x, y)$ computa o comprimento de arco nesse círculo. A comparação entre essa métrica intrínseca d' e a métrica vinda do ambiente é

$$|x - y| \leq d'(x, y) \leq \frac{\pi}{2}|x - y|,$$

a mesma que provamos acima. A prova é igual.

O que não é óbvio é que d' em \mathbb{S}^d é de fato uma métrica. Mais precisamente, não é óbvio que ela satisfaz a desigualdade triangular. Provemos tal fato.

Proposição 1.25. Para $x, y, z \in \mathbb{S}^d$, temos

$$d'(x, y) \leq d'(x, y) + d'(y, z)$$

Demonstração. Suponha que não vale. Há x, y, z tais que

$$\arccos\langle x, z \rangle > \arccos\langle x, y \rangle + \arccos\langle y, z \rangle.$$

Como a função $\cos : [0, \pi] \rightarrow [-1, 1]$ é decrescente, temos

$$\langle z, x \rangle < \cos(\arccos\langle x, y \rangle + \arccos\langle y, z \rangle),$$

e, usando que $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$, obtemos

$$\langle z, x \rangle < \cos(\arccos\langle x, y \rangle)\cos(\arccos\langle y, z \rangle) - \sin(\arccos\langle x, y \rangle)\sin(\arccos\langle y, z \rangle),$$

que pode ser simplificado para

$$\langle z, x \rangle < \langle x, y \rangle \langle y, z \rangle - \sqrt{1 - \langle x, y \rangle^2} \sqrt{1 - \langle y, z \rangle^2},$$

onde usamos que $\sin(\alpha) = \sqrt{1 - \cos(\alpha)}$ para $0 \leq \alpha \leq \pi$. Assim

$$\sqrt{1 - \langle x, y \rangle^2} \sqrt{1 - \langle y, z \rangle^2} < \langle x, y \rangle \langle y, z \rangle - \langle z, x \rangle,$$

$$(1 - \langle x, y \rangle^2)(1 - \langle y, z \rangle^2) < \langle x, y \rangle^2 \langle y, z \rangle^2 - 2\langle x, y \rangle \langle y, z \rangle \langle z, x \rangle + \langle z, x \rangle^2,$$

$$1 - \langle x, y \rangle^2 - \langle y, z \rangle^2 - \langle z, x \rangle^2 - 2\langle x, y \rangle \langle y, z \rangle \langle z, x \rangle < 0.$$

Por outro lado, a matriz de Gram

$$G := \begin{bmatrix} \langle x, x \rangle & \langle x, y \rangle & \langle x, z \rangle \\ \langle y, x \rangle & \langle y, y \rangle & \langle y, z \rangle \\ \langle z, x \rangle & \langle z, y \rangle & \langle z, z \rangle \end{bmatrix}$$

é igual a $A^T A$, onde

$$A = \begin{bmatrix} x^1 & y^1 & z^1 \\ x^2 & y^2 & z^2 \\ x^3 & y^3 & z^3 \end{bmatrix}.$$

Em particular, $\det(G) = \det(A)^2 \geq 0$, ou seja, obtemos

$$1 - \langle x, y \rangle^2 - \langle y, z \rangle^2 - \langle z, x \rangle^2 - 2\langle x, y \rangle \langle y, z \rangle \langle z, x \rangle \geq 0,$$

contradizendo a desigualdade que deduzimos acima. ■

Portanto, temos a métrica de \mathbb{R}^{d+1} restrita a esfera \mathbb{S}^d , a esfera intrínseca d' em \mathbb{S}^d e elas ainda se comparam, o que quer dizer que o conceito de pequeno para uma métrica é pequeno para outra.

Tal fenômeno também ocorre em \mathbb{R}^d , onde se compara entre métricas oriundas de normas. Se temos uma norma \mathbf{p} em \mathbb{R}^d , então sempre há constantes $C_2, C_1 > 0$ para as quais $|x|C_1 \leq \mathbf{p}(x) \leq C_2|x|$ (veja Proposição 1.31). Em termos de métricas, temos

$$C_1|x - y| \leq \mathbf{p}(x - y) \leq C_2|x - y|.$$

Por exemplo, a função $|x|_1 := |x^1| + |x^2| + \cdots + |x^d|$ define uma norma. Por um lado,

$$|x|_1 = \langle (|x^1|, \dots, |x^d|), (1, 1, \dots, 1) \rangle \leq |x|(1, 1, \dots, 1) = \sqrt{d}|x|.$$

Por outro lado, $|x|_1^2$ tem entre seus temos $(x^1)^2 + (x^2)^2 + \cdots + (x^d)^2$, ou seja, $|x|_1^2 \geq |x|^2$. Logo,

$$|x| \leq |x|_1 \leq \sqrt{d}|x|.$$

Exemplo similar vem da norma $|x|_\infty := \max_i |x^i|$.

Temos

$$\frac{1}{\sqrt{d}}|x| \leq |x|_\infty \leq |x|.$$

Exercício 1.26. Mostre que:

- $|\cdot|_1, |\cdot|_\infty$ são normas;
- $\sum_i |x^i||y^i| \leq |x|_1|y|_\infty$ para $x, y \in \mathbb{R}^d$.

Essas três normas, $|\cdot|$, $|\cdot|_1$ e $|\cdot|_\infty$, fazem parte de uma família maior de normas, as normas l^p . Para $p \in [1, \infty)$, definimos

$$|x|_p := \left(\sum_i |x^i|^p \right)^{\frac{1}{p}}.$$

Para todo $p \in [1, \infty)$, a expressão acima define uma norma. A parte difícil é mostrar a desigualdade triangular, que faremos a seguir (veja Proposição 1.30). Além disso, repare que $|x|_\infty = \lim_{p \rightarrow \infty} |x|_p$ e que $|x| = |x|_2$.

Exercício 1.27. Mostre que $|x|_\infty = \lim_{p \rightarrow \infty} |x|_p$.

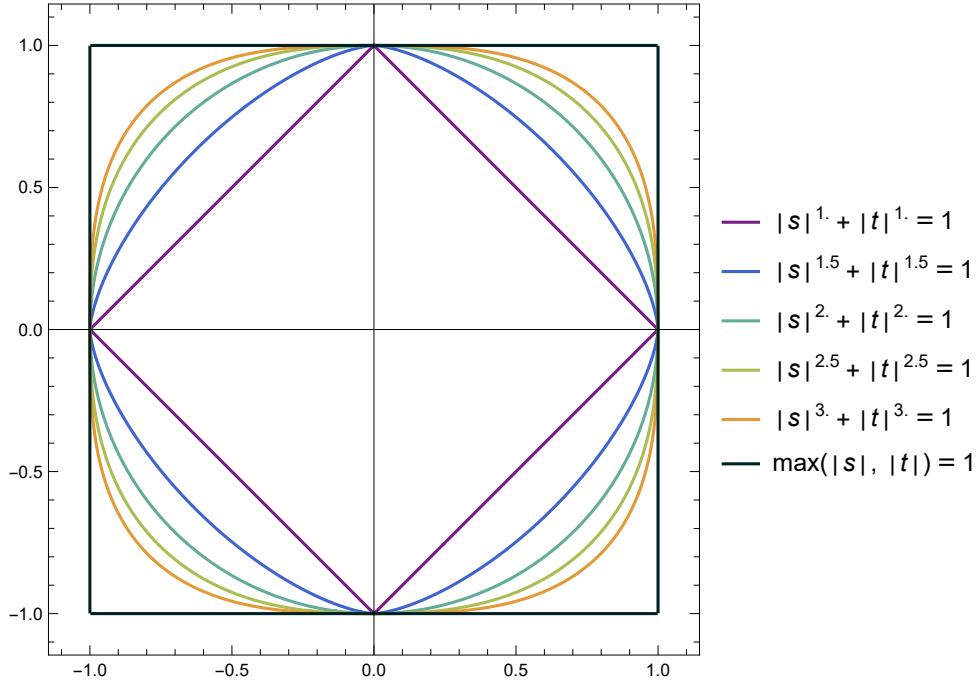


Figura 1.4: Círculo $|(s, t)|_p = 1$ para vários p 's.

Mostremos a desigualdade triangular para as normas l^p .

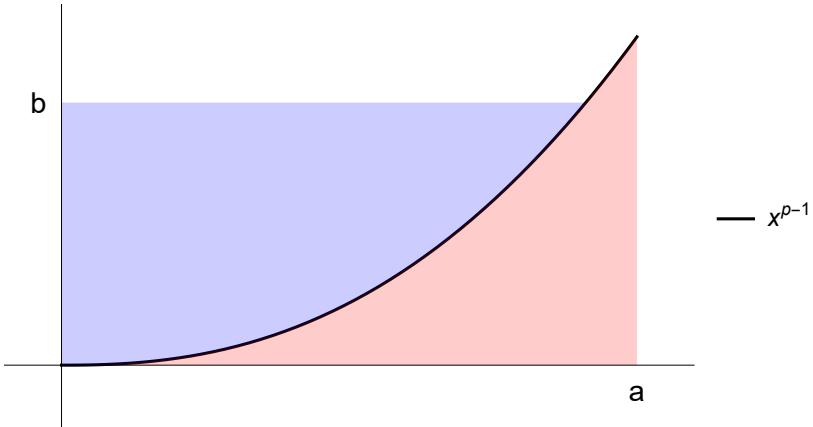


Figura 1.5: Gráfico da função x^{p-1} .

Lema 1.28 (Desigualdade de Young). Considere $p \in (1, \infty)$ e tome q tal que $\frac{1}{p} + \frac{1}{q} = 1$, isto é, $q = \frac{p}{p-1}$. Para $a, b \geq 0$ temos

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Demonação. Considere a função $f(x) = x^{p-1}$. A área abaixo do gráfico para $x \in [0, a]$ é $\frac{a^p}{p}$, em vermelho na Figura 1.5. A inversa dessa função é $g(y) = y^{q-1}$, pois se temos $y = x^{p-1}$, então $x = y^{\frac{1}{p-1}} = y^{q-1}$. A área em azul na Figura 1.5 é a integral de $g(y)$ de 0 a b , isto é, $\frac{b^q}{q}$.

O quadrado definido pelos pontos $(0, 0), (a, 0), (a, b), (0, b)$ tem área ab . Além disso, esse se encontra dentro da região conjunta formada pela região vermelha com a azul. Assim,

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

■

Lema 1.29 (Desigualdade de Hölder). Se $p \in (1, \infty)$ e $q = \frac{p}{p-1}$, então para quaisquer $x, y \in \mathbb{R}^d$

$$\sum_i |x^i| |y^i| \leq |x|_p |y|_q.$$

Repare que para $p = q = 2$, temos que

$$\sum_i |x^i| |y^i| \leq |x|_2 |y|_2,$$

(que segue da desigualdade de Cauchy-Schwarz 1.17). Além disso, podemos extender o resultado para $p = 1$ e $q = \infty$. Pois nesse caso ainda temos $\frac{1}{p} + \frac{1}{q} = 1$ e

$$\sum_i |x^i| |y^i| \leq |x|_1 |y|_\infty.$$

Demonstração. Se $x = 0$ ou $y = 0$ a desigualdade é trivialmente verdadeira. Podemos assumir $x \neq 0$ e $y \neq 0$. Considere $u = x/|x|_p$ e $v = y/|y|_q$. Temos $|u|_p = |v|_q = 1$. Pela desigualdade de Young 1.28, temos

$$\sum_i |u^i| |v^i| \leq \sum_i \frac{|u^i|^p}{p} + \frac{|v^i|^q}{q} = \frac{\sum_i |u^i|^p}{p} + \frac{\sum_i |v^i|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Como $u^i = x^i/|x|_p$ e $v^i = y^i/|y|_q$, temos

$$\sum_i |x^i| |y^i| \leq |x|_p |y|_q.$$

■

Finalmente, provamos a desigualdade triangular.

Proposição 1.30 (Desigualdade de Minkowski). Considere $p \in [1, \infty]$. Para todo $x, y \in \mathbb{R}^d$ temos

$$|x + y|_p \leq |x|_p + |y|_p$$

Demonstração. Podemos supor que $1 < p < \infty$, pois os casos extremos $p = 1, \infty$ são fáceis e são deixados ao leitor. Se $x + y, x$ ou y são nulos, o resultado é obviamente verdadeiro. Assim, podemos supor $x + y \neq 0, x \neq 0$ e $y \neq 0$. Temos que

$$\begin{aligned} |x + y|_p^p &= \sum_i |x^i + y^i|^p |x^i + y^i|^{p-1} \\ &\leq \sum_i (|x^i| + |y^i|) |x^i + y^i|^{p-1} \\ &= \sum_i |x^i| |x^i + y^i|^{p-1} + \sum_i |y^i| |x^i + y^i|^{p-1} \\ &\leq \left(\sum_i |x^i|^p \right)^{1/p} \left(|x^i + y^i|^{(p-1)q} \right)^{1/q} + \left(\sum_i |y^i|^p \right)^{1/p} \left(|x^i + y^i|^{(p-1)q} \right)^{1/q} \\ &= |x|_p \left(|x^i + y^i|^p \right)^{\frac{p-1}{p}} + \left(\sum_i |y^i|^p \right)^{1/p} \left(|x^i + y^i|^p \right)^{\frac{p-1}{p}} \\ &= |x|_p |x + y|_p^{p-1} + |y|_p |x + y|_p^{p-1} \\ &= (|x|_p + |y|_p) |x + y|_p^{p-1} \end{aligned}$$

onde a desigualdade de Hölder 1.29 entra na passagem da terceira para quarta linha.

Portanto,

$$|x + y|_p^p \leq (|x|_p + |y|_p) |x + y|_p^{p-1},$$

$$|x + y|_p \leq |x|_p + |y|_p.$$

■

Proposição 1.31. Para quaisquer normas \mathbf{p} em \mathbb{R}^d , existem constantes $C_1, C_2 > 0$ tais que

$$C_1|x| \leq \mathbf{p}(x) \leq C_2|x|.$$

Demonstração. Na base canônica temos $x = x^1 e_1 + \cdots + x^d e_d$. Logo, pela desigualdade de Cauchy-Schwarz 1.17, temos

$$\mathbf{p}(x) \leq |x^1|\mathbf{p}(e_1) + \cdots + |x^d|\mathbf{p}(e_d) \leq C_2|x|$$

onde $C_2 = \sqrt{\mathbf{p}(e_1)^2 + \cdots + \mathbf{p}(e_d)^2}$.

Agora mostremos que existe uma constante positiva C_1 pequena suficiente de modo que $C_1|x| \leq \mathbf{p}(x)$. Suponha que não existe tal constante. Para cada n existe $x_n = (x_n^1, \dots, x_n^d)$ tal que

$$\mathbf{p}(x_n) < \frac{1}{n}|x_n|.$$

Tome $y_n = x_n/|x_n|$. Temos $p(y_n) < \frac{1}{n}$, ou seja, $p(y_n) \rightarrow 0$.

Temos que y_n está na esfera de raio 1. Em particular, para cada índice i temos $y_n^i \in [-1, 1]$.

Uma sequência limitada admite subsequência convergente (veja Teorema 1.11). A menos de passar para uma subsequência, podemos assumir que y_n^1 converge. Novamente passando para subsequência, podemos agora supor que y_n^2 também converge. Fazendo o processo novamente, podemos supor que y_n^3 também converge. Assim, a menos de passar para uma subsequência, podemos supor que y_n^1, \dots, y_n^d convergem em \mathbb{R} para y^1, \dots, y^d . Como $(y_n^1)^2 + \cdots + (y_n^d)^2 = 1$ para todo n , temos $|y| = 1$, onde $y = (y^1, \dots, y^d)$.

Por um lado,

$$p(y) \leq p(y_n - y) + p(y_n) \leq C_2|y_n - y| + p(y_n) \rightarrow 0,$$

ou seja, $p(y) = 0$.

Por outro lado, $y \neq 0$, ou seja, $p(y) \neq 0$, porque \mathbf{p} é uma norma. Uma contradição. ■

A menos que seja dito o contrário, sempre que considerarmos uma norma/métrica em \mathbb{R}^d , estamos considerando a canônica.

Definição 1.32. Duas métricas d_1, d_2 em X são Lipschitz equivalentes se há constantes $C_1, C_2 > 0$ tais que

$$C_1 d_1(x, y) \leq d_2(x, y) \leq C_2 d_1(x, y).$$

Exercício 1.33. Prove que a relação “ser Lipschitz equivalente” em métricas sobre X define uma relação de equivalência.

Como vimos, as duas métricas consideradas na esfera \mathbb{S}^d são Lipschitz equivalentes e pela Proposição 1.31 temos que métricas oriundas de normas distintas em \mathbb{R}^d são Lipschitz equivalentes.

Além disso, observe que como todo espaço vetorial V de dimensão finita é isomorfo a algum \mathbb{R}^d , temos que duas normas nesse espaço vetorial são Lipschitz equivalentes. Em particular, as normas sobre matrizes quadradas vistas no Exercício 1.20 são equivalentes, o que não é imediato.

Definição 1.34. Seja (X, d) um espaço métrico. A bola aberta de centro x e raio r é o conjunto $B_{(X,d)}(x, r) = \{y \in X : d_X(x, y) < r\}$. Se estamos trabalhando com apenas a métrica d , então escrevemos apenas $B(x, r)$ para designar a bola, pois não há ambiguidade.

Em \mathbb{R}^d , uma bola centrada em x com raio r é da forma

$$B_{\mathbb{R}^d}(x, r) = \{y \in \mathbb{R}^d : \sum_i |y^i - x^i|^2 < r^2\}.$$

Para $d = 1$, obtemos

$$B_{\mathbb{R}}(x, r) = \{y \in \mathbb{R} : |y - x|^2 < r^2\} = (x - r, x + r),$$

o intervalo centrado em x com comprimento $2r$.

Se temos duas métricas d_1, d_2 Lipschitz equivalentes com constantes $C_1, C_2 > 0$ satisfazendo

$$C_1 d_1(x, y) \leq d_2(x, y) \leq C_2 d_1(x, y),$$

então podemos comparar as bolas de (X, d_1) e (X, d_2) :

$$B_{(X, d_1)}(x, r) \subset B_{(X, d_2)}(x, C_2 r) \quad \text{e} \quad B_{(X, d_2)}(x, s) \subset B_{(X, d_1)}\left(x, \frac{s}{C_1}\right).$$

Definição 1.35. Seja (X, d_X) um espaço métrico. Um aberto de X é um subconjunto U com a seguinte propriedade: para cada $x \in X$, existe $r > 0$ tal que $B(x, r) \subset U$.

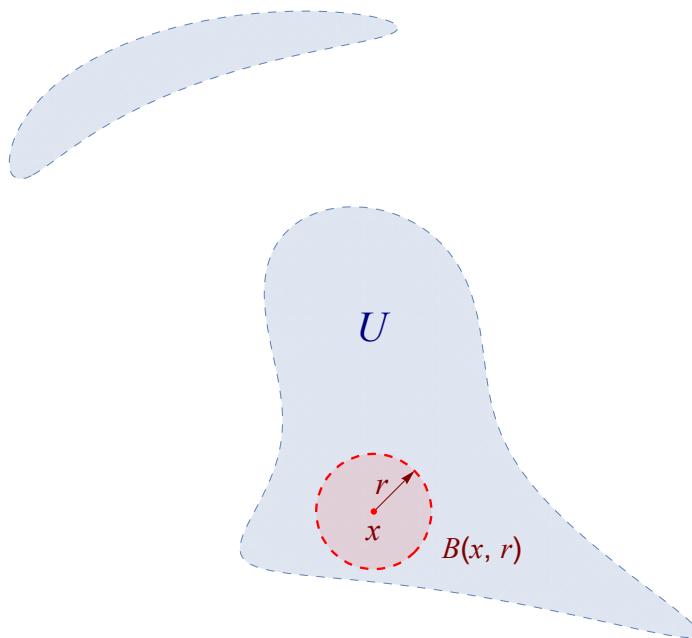


Figura 1.6: Um ponto x em U cabe com folga.

Intuitivamente, um aberto em X é um conjunto no qual todo ponto cabe com certa folga dentro de U . De fato, se $x \in U$, existe $B(x, r) \subset U$ com $r > 0$. Isso quer dizer que se $y \notin U$, então $d_X(x, y) \geq r$.

Exercício 1.36. Sejam d_1, d_2 métricas Lipschitz equivalentes em X . Mostre que U é aberto com respeito a d_1 se, e só se, for aberto com respeito a d_2 .

Exercício 1.37. Considere \mathbb{R} com as métricas $d_1(x, y) = |x - y|$ e $d_2(x, y) = \min(1, |x - y|)$. Mostre que

- d_2 é de fato uma métrica;
- as métricas d_1, d_2 **não** são Lipschitz equivalentes;
- as métricas d_1, d_2 produzem os mesmos abertos em \mathbb{R} .

Exercício 1.38. Mostre que $[0, 1)$ é aberto de $\mathbb{R}_{\geq 0}$, mas não é aberto de \mathbb{R} .

Proposição 1.39. A bola aberta $B(x, r)$ em (X, d_X) é aberta.

Demonstração. Considere $x' \in B(x, r)$ e $r' = r - d_X(x', x)$, que é positivo. Se $y \in B(x', r')$, então

$$d_X(x, y) \leq d_X(x, x') + d_X(x', y) < r' + d_X(x', y) = r,$$

ou seja, $B(x', r') \subset B(x, r)$. ■

O conjunto vazio \emptyset é aberto de X . Se não fosse aberto, haveria um ponto x em \emptyset para o qual $B(x, r) \not\subset \emptyset$ para todo r , o que é impossível porque \emptyset não possui elementos. Assim, \emptyset é aberto. O conjunto X é aberto de X .

Além disso, se U_1, U_2 são abertos de X , então $U_1 \cap U_2$ é aberto de X . De fato, dado $x \in U_1 \cap U_2$, existem $r_1, r_2 > 0$ tais que $B(x, r_1) \subset U_1$ e $B(x, r_2) \subset U_2$. Tomando $r = \min(r_1, r_2)$, temos

$$B(x, r) \subset U_1 \cap U_2.$$

Por fim, se $\{U_\lambda : \lambda \in \Lambda\}$ é uma família de abertos de X , então $\cup_{\lambda \in \Lambda} U_\lambda$ é aberto de X , onde essa união pode ser finita ou infinita (não necessariamente enumerável). De fato, se $x \in \cup_{\lambda \in \Lambda} U_\lambda$, então existe $\lambda_0 \in \Lambda$ tal que $x \in U_{\lambda_0}$. Como U_{λ_0} é aberto, existe $r > 0$ para o qual

$$B(x, r) \subset U_{\lambda_0} \subset \bigcup_{\lambda \in \Lambda} U_\lambda.$$

Em resumo, obtemos o seguinte resultado:

Proposição 1.40. Seja (X, d_X) espaço métrico.

- \emptyset, X são abertos de X ;
- se U_1, U_2 são abertos de X , então $U_1 \cap U_2$ é aberto de X ;
- se $\{U_\lambda : \lambda \in \Lambda\}$ é uma família de abertos de X , então $\bigcup_{\lambda \in \Lambda} U_\lambda$ é aberto de X .

Exercício 1.41. Mostre que todo aberto U de um espaço métrico X é união das bolas abertas contidas em U .

Exercício 1.42. Mostre que

- todo intervalo aberto de \mathbb{R} contém um número racional;
- todo subconjunto de \mathbb{R} é união de intervalos.

Dica: $A \subset \mathbb{R}$ é intervalo se para $a_1, a_2 \in A$, com $a_1 \leq a_2$, temos $[a_1, a_2] \subset A$;

- um aberto de \mathbb{R} é união enumerável de intervalos abertos disjuntos.

1.3 Topologia de espaço métrico

Definição 1.43. Considere um conjunto X com uma família τ de subconjuntos de X . Dizemos que τ é uma topologia se satisfaz as seguintes propriedades:

- $\emptyset, X \in \tau$;
- se $U_1, U_2 \in \tau$, então $U_1 \cap U_2 \in \tau$;
- se $\{U_\lambda : \lambda \in \Lambda\} \subset \tau$, então $\bigcup_{\lambda \in \Lambda} U_\lambda \in \tau$.

Nesse caso, chamamos (X, τ) de espaço topológico e os elementos de τ são chamados de abertos de X . Quando a topologia está clara do contexto, escrevemos $U \subset X$ para denotar que U é aberto de X .

Em resumo, uma topologia é uma família distinguida de subconjuntos chamados abertos, que devem seguir regras básicas. O espaço todo e o vazio devem ser abertos, intersecção de finitos abertos é aberto, e união arbitrária de abertos é aberta.

Exemplo 1.44. Seja (X, d_X) um espaço métrico. Lembre-se que definimos aberto em X de acordo com a Definição 1.35. O conjunto $\tau = \{U \subset X : U \text{ é aberto de } X\}$ define uma topologia em X , por causa da Proposição 1.40. Dizemos que essa topologia é induzida pela métrica d_X .

A partir desse ponto, assumiremos que todo espaço tem uma métrica.

Exemplo 1.45. Exemplos extremos. Seja X um conjunto.

- A maior topologia em X é a discreta, onde todo subconjunto é aberto:

$$\tau_{\text{discreta}} = \{U : U \subset X\}.$$

Essa topologia é induzida pela métrica $d(x_1, x_2)$ que vale 1 se $x_1 \neq x_2$ e 0 se $x_1 = x_2$.

- A menor topologia em X é a caótica, que tem somente \emptyset e X como abertos:

$$\tau_{\text{caótica}} = \{\emptyset, X\}.$$

Essa topologia também vem de uma métrica. De fato, basta tomar $d(x_1, x_2) = 0$ para todo $x_1, x_2 \in X$.

Definição 1.46. Uma propriedade topológica é aquela que pode ser caracterizada a partir da topologia.

Como veremos, continuidade, convergência de sequências, conexidade, conexidade por caminhos e compacidade são propriedades caracterizadas a partir da topologia.

Exemplo 1.47. Considere em reta \mathbb{R} as métricas $d_1(x, y) = |x - y|$ e $d_2(x, y) = \min(|x - y|, 1)$. O conjunto \mathbb{Z} é ilimitado com respeito a métrica d_1 e limitado com respeito a métrica d_2 , isto é, $d_2(x, 0) \leq 1$ para todo $x \in \mathbb{Z}$. No entanto, d_1, d_2 induzem a mesma topologia. Assim, a propriedade “ser limitado” não é uma propriedade topológica, i.e., não pode ser caracterizada a partir da topologia, que é a discreta nesse exemplo.

Em breve daremos uma definição mais apropriada de propriedade topológica, como sendo aquela que é invariante por homeomorfismo.

Exercício 1.48. Considere um espaço métrico X . Mostre que para pontos $x, y \in X$ distintos, existem abertos U e V de X tais que $U \cap V = \emptyset$, $x \in U$ e $y \in V$.

Exercício 1.49. Seja (X, d_X) espaço métrico e Y subconjunto. Considere em Y a métrica d_Y induzida de X , isto é, $d_Y(y_1, y_2) := d_X(y_1, y_2)$. Mostre que $A \subset Y$ é aberto com respeito a métrica d_Y se, e somente se, existe um aberto U de X tal que $A = Y \cap U$.

Dica: Escreva A como uma união de bolas de Y e use que bolas de Y são bolas de X intersectadas com Y . Construa U como união de bolas de X .

Exercício 1.50. Considere $\mathbb{Z}^d \subset \mathbb{R}^d$. Mostre que topologia induzida em \mathbb{Z}^d é a discreta.

Exercício 1.51. Se Y é aberto de X , então U é aberto de Y se, e só se, U é aberto de X e $U \subset Y$.

Em um espaço métrico, sempre usaremos a topologia induzida pela sua métrica. Além disso, lembre-se que métricas Lipschitz equivalentes induzem a mesma topologia (Exercício 1.36), então em espaços normados, por exemplo, podemos trabalhar com qualquer norma quando interessados em propriedades relacionadas a topologia.

Exercício 1.52. Considere os espaços métricos $(X_1, d_{X_1}), \dots, (X_q, d_{X_q})$.

- Mostre que as métricas

$$d_1((x^1, \dots, x^q), (y^1, \dots, y^q)) := \sum_{i=1}^q d_{X_i}(x^i, y^i),$$

$$d_2((x^1, \dots, x^q), (y^1, \dots, y^q)) := \sqrt{\sum_{i=1}^q d_{X_i}(x^i, y^i)^2},$$

$$d_\infty((x^1, \dots, x^q), (y^1, \dots, y^q)) := \max_{1 \leq i \leq q} (d_{X_i}(x^i, y^i))$$

em $X_1 \times \dots \times X_q$ são Lipschitz equivalentes e, consequentemente, definem a mesma topologia.

- Mostre $\Omega \subset X_1 \times \dots \times X_q$ é aberto se, e só se, para cada $x = (x^1, \dots, x^q) \in \Omega$ existem abertos U_1, \dots, U_q com $U_i \subset X_i$ tais que

$$x \in U_1 \times \dots \times U_q \subset \Omega.$$

Dica: Utilize a métrica d_∞ .

Nós chamamos a topologia descrita no exercício acima de **topologia produto**. Obviamente, a topologia de \mathbb{R}^d (oriunda de uma norma) coincide com a topologia produto.

Definição 1.53. Seja X espaço métrico. O subconjunto $F \subset X$ é fechado se $X \setminus F$ é aberto.

Exemplo 1.54. Pontos são fechados. Mais precisamente, se $x \in X$, então $\{x\}$ é fechado. De fato, dado $y \in X \setminus \{x\}$ temos que $B(y, d(y, x)) \subset X \setminus \{x\}$, ou seja, $X \setminus \{x\}$ é aberto.

Exercício 1.55. Mostre que a bola fechada

$$B[x, r] = \{y \in X : d(x, y) \leq r\}$$

é fechada.

Proposição 1.56. Temos que

- X e \emptyset são fechados;
- se F_1, F_2 são fechados, então $F_1 \cup F_2$ é fechado;
- se $\{F_\lambda\}_{\lambda \in \Lambda}$ é uma família de fechados, então $\bigcap_{\lambda \in \Lambda} F_\lambda$ é fechado.

Demonstração. Basta usar as identidades de conjuntos

$$X \setminus \bigcup_i A_i = \bigcap_i (X \setminus A_i),$$

$$X \setminus \bigcap_i A_i = \bigcup_i (X \setminus A_i).$$

■

Definição 1.57. Uma vizinhança de $x \in X$ é um aberto contendo x .

Exemplo 1.58. Em um espaço métrico X , temos que bolas $B(x, r)$ são vizinhanças de x .

Exemplo 1.59. Se X está munido da topologia discreta, então $\{x\}$ é uma vizinhança de x . Se X está munido da topologia caótica, a única vizinhança de um ponto é o espaço X todo.

Definição 1.60. Considere um subconjunto A de X e ponto $x \in X$. Dizemos que x é ponto aderente de A em X se para toda vizinhança V de x temos que $V \cap A \neq \emptyset$. O fecho \bar{A} de A em X é o conjunto dos pontos aderentes de A em X .

Em outras palavras, pontos aderentes são aqueles próximos de A , mas que não necessariamente estão em A . O fecho \bar{A} de A é formado pelos pontos de X próximos de A .

Exercício 1.61. Considere o subconjunto A do espaço métrico X . Mostre que $x \in X$ é ponto aderente de A em X se, e só se, $B(x, r) \cap A \neq \emptyset$ para todo $r > 0$.

Exemplo 1.62. O ponto $\sqrt{2}$ é aderente de \mathbb{Q} em \mathbb{R} . De fato, dado $r > 0$, considere $N \in \mathbb{Z}_{\geq 1}$ tal que $\frac{1}{N} < r$. Seja k o maior inteiro satisfazendo $k \leq N\sqrt{2}$. Temos que $N\sqrt{2} < k + 1$. Portanto,

$$0 \leq \frac{k}{N} - \sqrt{2} < \frac{1}{N} < r.$$

Em outras palavras, $\frac{k}{N} \in (\sqrt{2} - r, \sqrt{2} + r) = B(\sqrt{2}, r)$.

Portanto, $B(\sqrt{2}, r) \cap \mathbb{Q} \neq \emptyset$ para todo $r > 0$ e, consequentemente, $\sqrt{2} \in \overline{\mathbb{Q}}$.

Agora caracterizaremos fechados em termos de pontos aderentes. Um conjunto fechado será aquele que contém seus pontos aderentes.

Proposição 1.63. Seja A subconjunto de um espaço métrico X .

- \bar{A} é fechado.
- A é fechado se, e só se, $A = \bar{A}$.

Demonstração. Provemos que \bar{A} é fechado. Tome $x \in X \setminus \bar{A}$, isto é, x não é aderente de A em X . Existe uma bola $B(x, r)$ de x tal que $B(x, r) \cap A = \emptyset$. Para todo $y \in B(x, r)$ temos que $B(x, r)$ é vizinhança de y disjunta de A , ou seja, nenhum ponto de $B(x, r)$ é aderente a A em X . Portanto, $B(x, r) \subset X \setminus \bar{A}$. Daí segue que $X \setminus \bar{A}$ é aberto em X e \bar{A} é fechado em X .

Agora mostremos que A é fechado se, e só se, $A = \bar{A}$. Como vimos acima, \bar{A} é sempre fechado. Assim sendo, se $A = \bar{A}$, então A é fechado.

Provemos a recíproca. Considere A fechado e mostremos que $A = \bar{A}$. Tome $x \notin A$. Existe uma bola $B(x, r)$ de x tal que $B(x, r) \subset X \setminus A$, ou seja, $B(x, r) \cap A = \emptyset$ e, consequentemente, x não é aderente a A em X . Portanto, $\bar{A} \subset A$. Como $A \subset \bar{A}$ sempre vale, concluímos que $A = \bar{A}$. ■

Exercício 1.64. Mostre que $\overline{\mathbb{Q}}$ em \mathbb{R} é \mathbb{R} .

Quando um subconjunto A de X tem fecho $\bar{A} = X$, dizemos que A é denso em X . Com isso em mente, \mathbb{Q} é denso em \mathbb{R} seguindo o exercício acima. O conjunto dos números irracionais também é denso em \mathbb{R} e \mathbb{Q}^d é denso em \mathbb{R}^d .

Exercício 1.65. Seja X espaço métrico e A um subconjunto de X . Definimos a distância de $x \in X$ a A como sendo

$$d_X(x, A) := \inf_{a \in A} d_X(x, a).$$

Mostre que $x \in \bar{A}$ se, e só se, $d_X(x, A) = 0$.

1.4 Sequência em espaços métricos

Definição 1.66. Seja X espaço métrico. A sequência x_n converge para x se dada uma vizinhança U de x existe $n_0 \in \mathbb{Z}_{\geq 1}$ tal que $x_n \in U$ para $n \geq n_0$. Denotamos essa convergência por $x_n \rightarrow x$ ou $\lim_{n \rightarrow \infty} x_n = x$.

Intuitivamente, isso quer dizer que para qualquer vizinhança de x a sequência x_n entra nela eventualmente e permanece lá. É comum dizer também que para qualquer vizinhança U de x temos $x_n \in U$ para n suficientemente grande.

Do ponto de vista da métrica, temos que $x_n \rightarrow x$ se, e só se, para cada $\varepsilon > 0$ existe $n_0 \in \mathbb{Z}_{\geq 1}$ tal que $x_n \in B(x, \varepsilon)$ para $n \geq n_0$, ou seja, se

$$d_X(x, x_n) < \varepsilon \quad \text{para } n \geq n_0.$$

Em outras palavras, temos que $x_n \rightarrow x$ se, e só se, $d_X(x_n, x) \rightarrow 0$.

Exercício 1.67. Considere um espaço métrico X , a sequência x_n e os pontos x, x' em X . Mostre que se $x_n \rightarrow x$ e $x_n \rightarrow x'$, então $x = x'$.

Observação 1.68. Quando lidando com matrizes reais ou complexas, uma norma útil é a dada por $|A|_\infty := \max_{i,j} |a_{ij}|$. Como todas as normas de \mathbb{R}^{d^2} (ou \mathbb{C}^{d^2}) são todas equivalentes entre si (veja a Proposição 1.31), temos que a norma $|\cdot|_\infty$ define a topologia do espaço de matrizes como qualquer outra. Desta norma fica fácil ver que $A_n \rightarrow A$ se, e só se, as entradas de A_n convergem para as correspondentes entradas de A .

Exemplo 1.69. Nesse exemplo, todas as matrizes são complexas. O resultado abaixo é falso para matrizes reais.

Seja A uma matriz $d \times d$ complexa. A decomposição de Schur nos diz que existem uma matriz triangular superior U e uma matriz unitária Q tais que

$$A = Q U Q^*,$$

onde Q^* é a transposta da matriz conjugada de Q . Por triangular superior queremos dizer que U tem os termos abaixo de sua diagonal nulos e por unitária queremos dizer que $Q Q^* = \text{Id} = Q^* Q$, isto é, Q é uma isometria da forma Hermitiana canônica de \mathbb{C}^d .

$$A = Q \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1d-1} & u_{1d} \\ 0 & u_{22} & u_{23} & \cdots & u_{2d-1} & u_{2d} \\ 0 & 0 & u_{33} & \cdots & u_{3d-1} & u_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{d-1d-1} & u_{d-1d} \\ 0 & 0 & 0 & \cdots & 0 & u_{dd} \end{bmatrix} Q^*.$$

A seguir mostraremos que perturbando a diagonal de U , obtemos matrizes diagonalizáveis suficientemente próximas de A .

Repare que os autovalores de A são $u_{11}, u_{22}, \dots, u_{dd}$, os termos da diagonal de U . Dado $n \in \mathbb{Z}_{\geq 1}$, escolha $\varepsilon_1[n], \varepsilon_2[n], \dots, \varepsilon_d[n] \geq 0$ tais que

$$u_{ii} + \varepsilon_i[n] \neq u_{jj} + \varepsilon_j[n]$$

para $i \neq j$ e $\varepsilon_1[n], \dots, \varepsilon_d[n] < \frac{1}{n}$.

Adicionamos uma perturbação a diagonal de U usando tais ε 's e definimos:

$$A_n := Q \begin{bmatrix} u_{11} + \varepsilon_1[n] & u_{12} & u_{13} & \cdots & u_{1d-1} & u_{1d} \\ 0 & u_{22} + \varepsilon_2[n] & u_{23} & \cdots & u_{2d-1} & u_{2d} \\ 0 & 0 & u_{33} + \varepsilon_3[n] & \cdots & u_{3d-1} & u_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{d-1d-1} + \varepsilon_{d-1}[n] & u_{d-1d} \\ 0 & 0 & 0 & \cdots & 0 & u_{dd} + \varepsilon_d[n] \end{bmatrix} Q^*.$$

Assim temos que cada A_n é diagonalizável porque tem os autovalores dois a dois distintos. Além disso, $A_n \rightarrow A$, pois as entradas de A_n convergem para as entradas de A , já que $\varepsilon_j[n] < \frac{1}{n}$. Logo, para toda matriz A , existe uma sequência A_n de matrizes diagonalizáveis que converge para A .

Como vimos ao fim da Secção 1.3. Um subconjunto D de X é denso se $\overline{D} = X$. No caso de matrizes complexas $M_d(\mathbb{C})$, temos que o espaço D das matrizes $d \times d$ diagonalizáveis é denso em $M_d(\mathbb{C})$. Mais precisamente, o argumento acima prova que $\overline{D} = M_d(\mathbb{C})$. Deixemos isso mais claro.

Exercício 1.70. Prove o teorema de Cayley-Hamilton. Considere uma matriz complexa $A \in M_d(\mathbb{C})$. Seja $p_A(x) = \det(x\text{Id} - A)$ o polinômio característico de A . Podemos escrever

$$p_A(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1x + c_0,$$

onde c_i depende polinomialmente das entradas de A . Por exemplo, $c_{d-1} = -\text{tr}(A)$ e $c_0 = (-1)^d \det(A)$.

- Mostre que se $B = QAQ^{-1}$ para Q invertível, então $p_B = p_A$.
- Mostre que se A é diagonalizável, então

$$A^d + c_{d-1}A^{d-1} + \cdots + c_1A + c_0\text{Id} = 0.$$

- **Teorema de Cayley-Hamilton:** Conclua que

$$A^d + c_{d-1}A^{d-1} + \cdots + c_1A + c_0\text{Id} = 0$$

para toda matriz complexa $A \in M_d(\mathbb{C})$ porque matrizes diagonalizáveis são densas.

Proposição 1.71. Considere o espaço métrico X . Temos que $D \subset X$ é denso em X , isto é, $\overline{D} = X$, se, e só se, para cada $x \in X$ existir uma sequência $x_n \in D$ tal que $x_n \rightarrow x$.

Demoração. Assuma que D é denso em X e considere $x \in X$. Para cada bola $B(x, \frac{1}{n})$, existe $x_n \in D$, pois $B(x, \frac{1}{n}) \cap D \neq \emptyset$. Logo, $d_X(x_n, x) < \frac{1}{n} \rightarrow 0$ e $x_n \in X$.

Por outro lado, assuma que para cada $x \in X$ existe uma sequência $x_n \in D$ convergindo a x e mostremos que $\overline{D} = X$. Fixe $x \in X$. Se $x_n \in D$ converge para $x \in X$, então para toda vizinhança U de x há n_0 suficientemente grande tal que $x_{n_0} \in U$. Portanto, a intersecção $U \cap D$ é não vazia e, consequentemente, $x \in \overline{D}$. ■

Da mesma forma, podemos provar que \mathbb{Q} é denso em \mathbb{R} aproximando cada $x \in \mathbb{R}$ por uma sequência de \mathbb{Q} . Isso é fácil de fazer usando expansão decimal por exemplo. A menos de

trocar x por $-x$, podemos assumir que $x \geq 0$. Na base 10 o elemento x se escreve como

$$x = a_0.a_1a_2a_3\cdots = a_0 + \sum_{i=1}^{\infty} \frac{a_i}{10^i},$$

com $a_0 \in \mathbb{Z}_{\geq 0}$ e $a_1, a_2, \dots \in \{0, 1, \dots, 9\}$. Truncando x obtemos $x_n := a_0.a_1a_2a_3\cdots a_n \in \mathbb{Q}$, que converge para x , pois

$$|x - x_n| \leq \sum_{i=n+1}^{\infty} \frac{9}{10^i} = \frac{1}{10^n} \sum_{i=1}^{\infty} \frac{9}{10^i} = \frac{1}{10^n} 0.9999\cdots = \frac{1}{10^n} \rightarrow 0.$$

Portanto, \mathbb{Q} é denso em \mathbb{R} .

Observação 1.72. Definindo $y_n := x_n(1 + \sqrt{2}/n)$ temos uma sequência de números irracionais convergindo a x . Logo, $(\overline{\mathbb{R} \setminus \mathbb{Q}}) = \mathbb{R}$.

Exercício 1.73. Considere o espaço métrico X e $A \subset X$. Mostre que $x \in \overline{A}$ se, e só se, existe uma sequência $x_n \in A$ tal que $x_n \rightarrow x$.

Exercício 1.74. Considere o espaço métrico X e $A \subset X$. Mostre que A é fechado se, e só se, para toda sequência convergente $x_n \in A$ com limite $x \in X$ temos que $x \in A$.

Assim, o conceito de fechado em espaços métricos pode ser caracterizado via sequências: fechado é aquele cujas sequências não escapam para fora do conjunto.

Observação 1.75. Se temos o produto de espaços métricos $X_1 \times \cdots \times X_q$ e uma sequência $x_n = (x_n^1, \dots, x_n^q)$, então temos que tal sequência converge para um ponto x se, e só se, $x_n^i \rightarrow x^i$, pois a topologia do produto pode, por exemplo, ser dada por

$$d_1(x, y) = \sum_{i=1}^q d_{X_i}(x^i, y^i)$$

e temos que $d_1(x_n, x) \rightarrow 0$ se, e só se, $d_{X_i}(x_n^i, x^i) \rightarrow 0$ para todo $i \in \{1, \dots, q\}$.

Em particular, uma sequência x_n em \mathbb{R}^d converge se, e só se, suas coordenadas x_n^i convergem.

Definição 1.76. Considere $A \subset X$, uma função $f : A \rightarrow Y$ e um ponto $x_0 \in \overline{A}$. Dizemos que $\lim_{x \rightarrow x_0} f(x) = L$ se $f(x_n) \rightarrow L$ para toda sequência $x_n \in A$ convergindo a x .

Depois veremos que pode-se considerar $\mathbb{R}^d \cup \{\infty\}$ como espaço métrico (veja Observação 1.106). Para tal métrica temos que $x_n \rightarrow \infty$ se, e só se, $\lim |x_n| = \infty$.

Proposição 1.77. Mostre que $\lim_{x \rightarrow x_0} f(x) = L$ se, e só se, para cada $\epsilon > 0$ existe $\delta > 0$ tal que $d_Y(f(x), L) < \epsilon$ quando $x \in A$ e $d_X(x, x_0) < \delta$.

Demonstração. Assuma que

$$\lim_{x \rightarrow x_0} f(x) = L,$$

e mostremos que para cada $\epsilon > 0$ existe $\delta > 0$ tal que $d_Y(f(x), L) < \epsilon$ quando $x \in A$ e $d_X(x, x_0) < \delta$.

Suponha que tal afirmação é falsa. Portanto, existe $\epsilon > 0$ tal que para todo $\delta > 0$ temos $d_Y(f(x), L) \geq \epsilon$ quando $x \in A$ e $d_X(x, x_0) < \delta$. Para cada n existe $x_n \in A$ tal que $d_X(x_n, x_0) < \frac{1}{n}$ e $d_Y(f(x_n), L) \geq \epsilon$. Logo, $x_n \rightarrow x_0$ e $f(x_n) \not\rightarrow L$, uma contradição.

Por outro lado, mostremos que se vale a condição ϵ e δ então temos que para toda sequência $x_n \in A$ convergindo para x_0 vale $f(x_n) \rightarrow L$. De fato, dado $\epsilon > 0$ existe $\delta > 0$ tal que $d(f(x), L) < \epsilon$ quando $x \in A$ e $d(x, x_0) < \delta$. Existe n_0 tal que $x_n \in B_X(x_0, \delta)$ para $n \geq n_0$. Assim, $d_Y(f(x_n), L) < \epsilon$ para $n \geq n_0$. Logo, $f(x_n) \rightarrow L$. ■

1.5 Continuidade

Definição 1.78. Sejam X, Y espaços métricos. Uma função $f : X \rightarrow Y$ é contínua em x se dada uma vizinhança U de $f(x)$, existe uma vizinhança V de x tal que $f(V) \subset U$. Se f é contínua em todo $x \in X$, dizemos que f é contínua.

Note que continuidade é algo que é puramente topológico, não depende de métrica ou de ε e δ .

Exercício 1.79. Mostre que se $f : X \rightarrow Y$ é constante, então f é contínua.

Exercício 1.80. Mostre que a função identidade é contínua.

Proposição 1.81. Sejam X e Y espaços métricos. A função $f : X \rightarrow Y$ é contínua em X se, e só se, dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$f(B_X(x, \delta)) \subset B_Y(f(x), \varepsilon),$$

isto é,

$$d_Y(f(x), f(y)) < \varepsilon \quad \text{quando} \quad d_X(x, y) < \delta.$$

Observação: As bolas no domínio são com respeito a X e as no contradomínio, com respeito a Y .

Demonstração. Se f é contínua em x , então para a vizinhança $B_Y(f(x), \varepsilon)$ de $f(x)$ existe uma vizinhança V de x tal que $f(V) \subset B_Y(f(x), \varepsilon)$. Como $x \in V$ e V é aberto de X , existe uma bola $B_X(x, \delta) \subset V$. Portanto,

$$f(B_X(x, \delta)) \subset B_Y(f(x), \varepsilon).$$

Agora suporemos que dado $\varepsilon > 0$, existe $\delta > 0$ tal que $f(B_X(x, \delta)) \subset B_Y(f(x), \varepsilon)$ e mostraremos que f é contínua.

Dada uma vizinhança U de $f(x)$, há ε tal que $B_Y(f(x), \varepsilon) \subset U$. Por hipótese, há $\delta > 0$ tal que $f(B_X(x, \delta)) \subset B_Y(f(x), \varepsilon)$. Tomando $V = B_X(x, \delta)$ temos $f(V) \subset B_Y(f(x), \varepsilon) \subset U$. ■

Observação 1.82. Repare que a proposição acima, quando aplicada a $X = Y = \mathbb{R}$, se torna a definição de continuidade via ε e δ do cálculo.

Exercício 1.83. Sejam X e Y espaços métricos. Uma função $f : X \rightarrow Y$ é Lipschitz se existe uma constante $C > 0$ tal que $d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2)$ para quaisquer $x_1, x_2 \in X$.

- Mostre que se f é Lipschitz, então f é contínua;
- Dê um exemplo de função contínua não-Lipschitz.

Proposição 1.84. Uma função $f : X \rightarrow Y$ é contínua em x se, e só se, para toda sequência x_n convergindo para x , temos $f(x_n) \rightarrow f(x)$.

Demonstração. Mostremos que se f é contínua em x e $x_n \rightarrow x$, então $f(x_n) \rightarrow f(x)$. Dada a vizinhança U de $f(x)$, existe uma vizinhança V de x tal que $f(V) \subset U$. Como $x_n \rightarrow x$, existe n_0 tal que $x_n \in V$ para $n \geq n_0$. Portanto, $f(x_n) \in U$ para $n \geq n_0$.

Mostremos agora que se f mapeia sequências convergentes para x em sequências convergentes para $f(x)$, então ela deve ser contínua. De fato, suponha que ela não seja contínua. Existe um $\varepsilon > 0$ tal que para todo $\delta > 0$ temos que $f(B_X(x, \delta)) \not\subset B_Y(f(x), \varepsilon)$. Em particular, para cada n , existe $x_n \in B_X(x, \frac{1}{n})$ tal que $f(x_n) \notin B_Y(f(x), \varepsilon)$. Temos que $x_n \rightarrow x$ porque $d_X(x_n, x) < \frac{1}{n}$ e $f(x_n) \not\rightarrow f(x)$ porque $d_Y(f(x_n), f(x)) \geq \varepsilon$, uma contradição. ■

Exercício 1.85.

- Seja X espaço métrico. Considere $f : X \rightarrow \mathbb{R}$ e $g : X \rightarrow \mathbb{R}$ contínuas em x . Mostre que $f + g$ e fg são contínuas em x . Supondo que g não se anula em x , mostre que f/g é contínua em x .
- Mostre que $\pi_i : \mathbb{R}^d \rightarrow \mathbb{R}$, dada por $\pi_i(x^1, \dots, x^d) = x^i$, é contínua (π_i manda x em sua i -ésima coordenada x^i).
- Mostre que toda função polinomial $p : \mathbb{R}^d \rightarrow \mathbb{R}$ é contínua. Por polinomial queremos dizer que p é da forma

$$p(x) = \sum_{i_1=0}^{n_1} \cdots \sum_{i_d=0}^{n_d} a_{i_1 \dots i_d} (x^1)^{i_1} \cdots (x^d)^{i_d}.$$

- Mostre que determinante e traço são contínuas como funções de $M_d(\mathbb{R})$ em \mathbb{R} , onde $M_d(\mathbb{R})$ é o espaço de matrizes quadradas $d \times d$ reais, que pode ser identificado com \mathbb{R}^{d^2} .

Proposição 1.86. Mostre que se $f : X \rightarrow Y$ e $g : Y \rightarrow Z$ são contínuas em x e $f(x)$, então $g \circ f : X \rightarrow Z$ é contínua em x .

Demonstração. Tome $x_n \rightarrow x$. Temos $f(x_n) \rightarrow f(x)$ e, consequentemente, $g(f(x_n)) \rightarrow g(f(x))$. ■

Logo, composta de funções contínuas é contínua.

Exemplo 1.87. Considere a função $f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ dada por $f(x^1, x^2) = \frac{\sin(x^1 + x^2)}{\sqrt{(x^1)^2 + (x^2)^2}}$. Essa função é contínua, pois pode ser construída compondo funções contínuas. As funções abaixo são contínuas porque são compostas de funções contínuas:

$$(x^1, x^2) \mapsto x^1 + x^2 \mapsto \sin(x^1 + x^2),$$

$$(x^1, x^2) \mapsto (x^1)^2 + (x^2)^2 \mapsto \sqrt{(x^1)^2 + (x^2)^2}.$$

Note que

$$(x^1, x^2) \mapsto x^1 + x^2, \quad (x^1, x^2) \mapsto (x^1)^2 + (x^2)^2$$

são contínuas por serem polinomiais e

$$y \mapsto \sin(y), \quad y \mapsto \sqrt{y}$$

são funções contínuas de uma variável real a valores reais.

Dividindo

$$(x^1, x^2) \mapsto \sin(x^1 + x^2)$$

por

$$(x^1, x^2) \mapsto \sqrt{(x^1)^2 + (x^2)^2}.$$

obtemos a função f .

Exercício 1.88. Seja X espaço métrico e Y um subconjunto.

- Mostre que o mapa de inclusão $i : Y \rightarrow X$, dado por $i(y) = y$, é contínuo.
- Se $f : X \rightarrow Z$ é uma função contínua entre dois espaços métricos e $Y \subset X$, então $f|_Y : Y \rightarrow Z$ é contínua.
- Se $f : X \rightarrow Z$ é uma função contínua entre dois espaços métricos e $f(X) \subset W$, então $f : X \rightarrow W$ é contínua.

Exemplo 1.89. Considere uma função contínua $f : X \rightarrow Y$. A função

$$\lambda : \{(x, f(x)) \in X \times Y : x \in X\} \rightarrow X,$$

dada por $(x, f(x)) \mapsto x$ é contínua, por ser restrição de uma função contínua.

Exercício 1.90. Considere os espaços métricos X, Y_1, Y_2, \dots, Y_d . Mostre que:

- a projeção $\pi_i : Y_1 \times \dots \times Y_d \rightarrow Y_i$, dada por $\pi_i(y) = y^i$, é contínua.
- uma função $f : X \rightarrow Y_1 \times \dots \times Y_d$ dada por $f = (f_1, \dots, f_d)$ é contínua se, e só se, cada f_i é contínua.

Exemplo 1.91. Considere o espaço M_d de matrizes quadradas $d \times d$. Temos que as operações básicas são contínuas. As funções Soma, Mult : $M_d \times M_d \rightarrow M_d$, dadas por $\text{Soma}(A, B) = A + B$ e $\text{Mult}(A, B) = AB$, tem como entradas polinômios e, portanto, são contínuas. Por exemplo, AB tem entrada i, j igual a

$$\sum_k a_{ik} b_{kj},$$

que é polinomial nas entradas de A e B .

Se considerarmos o grupo multiplicativo $\text{GL}(d, \mathbb{R}) := \det^{-1}(\mathbb{R}_{\neq 0})$, que é aberto porque pré-imagem de aberto é aberto, então podemos definir a função

$$\text{Inv} : \text{GL}(d, \mathbb{R}) \rightarrow \text{GL}(d, \mathbb{R})$$

pela fórmula $\text{Inv}(A) = A^{-1}$.

Mostremos que a função Inv é contínua. Considere o polinômio característico de $A \in \text{GL}(d, \mathbb{R})$. Temos

$$p(t) := \det(t - A) = t^d - c_{d-1}t^{d-1} + \dots + tc_1 + c_0,$$

onde $c_0 = (-1)^d \det(A)$ porque $p(0) = c_0$. Note que cada c_i é um polinômio nas entradas de A . Pelo teorema de Cayley-Hamilton (veja Exercício 1.70), temos que

$$A^d + c_{d-1}A^{d-1} + \dots + Ac_1 + c_0 = 0,$$

ou seja,

$$A(A^{d-1} + c_{d-1}A^{d-2} + \dots + c_1) = -c_0,$$

$$A^{-1} = \frac{(-1)^{d+1}}{\det(A)} (A^{d-1} + c_{d-1}A^{d-2} + \dots + c_1),$$

cujas entradas são quocientes de polinômios nas entradas de A .

Exemplo 1.92. Considere a esfera \mathbb{S}^d e a matriz A com tamanho $(d+1) \times (d+1)$. Enxergaremos os vetores de \mathbb{R}^{d+1} como vetores coluna e o produto interno como $\langle x, y \rangle = y^T x$. Definimos o mapa $E : \mathbb{S}^d \rightarrow \mathbb{R}$ como sendo

$$E(x) = \langle Ax, x \rangle = x^T Ax.$$

Essa função é contínua porque é restrição de uma função polinomial $\langle Ax, x \rangle$, que está definida em \mathbb{R}^{d+1} , na esfera.

O conceito de aberto nos permite caracterizar quando a função é contínua em todos os pontos do seu domínio.

Proposição 1.93. Sejam X e Y espaços métricos. Uma função é contínua se, e só se, pré-imagem de aberto é aberto, isto é, $f^{-1}(U) \subset \circ X$ quando $U \subset \circ Y$.

Demonstração. Seja f contínua e mostremos que $f^{-1}(U) \subset \circ X$ para $U \subset \circ Y$. De fato, tome $x \in f^{-1}(U)$. Temos que $f(x) \in U$. Como f é contínua em x , existe uma vizinhança V_x de x

tal que $f(V_x) \subset U$, ou seja, $V_x \subset f^{-1}(U)$. Como $f^{-1}(U) = \cup_{x \in f^{-1}(U)}(V_x)$, que é uma união de abertos de X , temos que $f^{-1}(U)$ é aberto.

Agora suponha que $f^{-1}(U) \subset X$ quando $U \subset Y$ e mostremos que f é contínua. Tome $x \in X$ e uma vizinhança U de $f(x)$. Como $V := f^{-1}(U)$ é aberto e $x \in V$, temos que V é vizinhança de x . Além disso, $f(V) \subset U$. Portanto, f é contínua em x e como x é arbitrário, f é contínua. ■

Mostremos que composta de funções contínuas é contínua novamente.

Proposição 1.94. Se $f : X \rightarrow Y$ e $g : Y \rightarrow Z$ são contínuas, então $g \circ f$ é contínua.

Demonstração. Se $U \subset Z$, então $g^{-1}(U) \subset Y$ e, consequentemente, $f^{-1}(g^{-1}(U)) \subset X$. Como $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$ é aberto em X , temos que $g \circ f$ é contínua. ■

Proposição 1.95. Sejam X, Y espaços métricos. Temos que $f : X \rightarrow Y$ é contínua se, e só se, $f^{-1}(F)$ é fechado em X para todo fechado F em Y .

Demonstração. Basta usar a identidade de conjunto

$$f^{-1}(Y \setminus F) = X \setminus f^{-1}(F)$$

e que pré-imagem de aberto é aberto se, e só se, a função for contínua. ■

Exemplo 1.96. Seja $f : X \rightarrow \mathbb{R}$ contínua e a um número real. Temos que

$$\{x \in X : f(x) \geq a\}, \quad \{x \in X : f(x) \leq a\} \quad \text{e} \quad f^{-1}(a)$$

são fechados, pois conjuntos $(-\infty, a]$, $[a, \infty)$ e $\{a\}$ são fechados em \mathbb{R} .

Exemplo 1.97. Seja (X, d) espaço métrico. Dado $a \in X$, a função $x \mapsto d(x, a)$ é contínua, porque

$$|d(x, a) - d(x', a)| \leq d(x, x')$$

e assim a função é Lipschitz (verifique isso!). Portanto, a bola fechada

$$B[a, r] := \{x \in X : d(x, a) \leq r\}$$

é fechada em X .

Exemplo 1.98. O grupo $\mathrm{SL}(d, \mathbb{R})$ das matrizes $d \times d$ com determinante 1 é fechado em $\mathrm{GL}(d, \mathbb{R})$ e em M_d porque $\mathrm{SL}(d, \mathbb{R}) = \det^{-1}(1)$.

Exercício 1.99. Sejam (X, d_X) espaço métrico e A subconjunto de X . Mostre que:

- $x \mapsto d_X(x, A) := \inf_{a \in A} d_X(x, a)$ é contínua;
- se A é fechado, então $f^{-1}(0) = A$;
- se A_1, A_2 são fechados disjuntos, então

$$\varphi(x) = \frac{d_X(x, A_2)}{d_X(x, A_1) + d_X(x, A_2)}$$

é uma função contínua que vale

$$\varphi|_{A_1} = 0, \quad \varphi|_{A_2} = 1, \quad 0 < \varphi(x) < 1 \quad \text{para} \quad x \in X \setminus (A_1 \cup A_2).$$

- os conjuntos

$$U_1 = \varphi^{-1}\left(\left(-\infty, \frac{1}{2}\right)\right) \quad \text{e} \quad U_2 = \varphi^{-1}\left(\left(\frac{1}{2}, \infty\right)\right)$$

são abertos disjuntos de X satisfazendo

$$A_1 \subset U_1, \quad A_2 \subset U_2 \quad \text{e} \quad U_1 \cap U_2 = \emptyset.$$

1.6 Homeomorfismos

Definição 1.100. Um homeomorfismo entre X e Y é uma bijeção $f : X \rightarrow Y$ contínua cuja inversa $f^{-1} : Y \rightarrow X$ é contínua. Se um homeomorfismo entre X e Y existe, então dizemos que X e Y são homeomorfos. Note que homeomorfismo define uma relação de equivalência.

Exemplo 1.101. A função identidade $\text{Id} : X \rightarrow X$, com domínio e contra domínio com mesma topologia, é homeomorfismo.

Exemplo 1.102. Considere $f : \mathbb{R} \rightarrow (0, 1)$ dada por $f(x) = \frac{e^x}{1+e^x}$ é homeomorfismo. De fato, computemos sua inversa e mostremos que essa é contínua.

Se $y = f(x)$, então

$$\begin{aligned} e^x y + y &= e^x, \\ e^x &= \frac{y}{1-y}, \\ x &= \log\left(\frac{y}{1-y}\right). \end{aligned}$$

Portanto,

$$f^{-1}(y) = \log\left(\frac{y}{1-y}\right)$$

é contínua.

Exercício 1.103. Mostre que todo intervalo (a, b) é homeomorfo a $(0, 1)$ e que o intervalo $[0, 1]$ é homeomorfo a $(0, 1]$.

Exemplo 1.104. Considere o grupo das matrizes $d \times d$ invertíveis $\text{GL}(d, \mathbb{R})$. Considere

$$Z = \{(A, t) \in M_d \times \mathbb{R} \times \mathbb{R} : \det(A)t = 1\}.$$

Temos o mapa $f : \text{GL}(d, \mathbb{R}) \rightarrow Z$ dado por

$$f(A) = \left(A, \frac{1}{\det(A)}\right),$$

que é contínuo. Sua inversa é $f^{-1}(A, t) = A$, que é contínua também.

Logo, f é um homeomorfismo entre $\text{GL}(d, \mathbb{R})$ e Z .

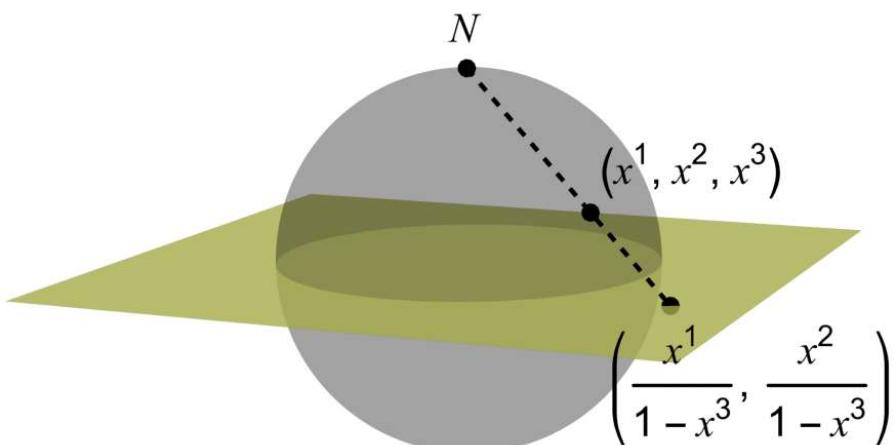


Figura 1.7: Projeção estereográfica.

Exemplo 1.105. (Projeção estereográfica) Considere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ e seja N o polo norte da esfera, i.e., $N = (0, 0, \dots, 0, 1)$. A projeção estereográfica

$$f : \mathbb{S}^d \setminus \{N\} \rightarrow \mathbb{R}^d$$

é dada por

$$f(x^1, x^2, \dots, x^d, x^{d+1}) = \left(\frac{x^1}{1 - x^{d+1}}, \frac{x^2}{1 - x^{d+1}}, \dots, \frac{x^d}{1 - x^{d+1}} \right),$$

que é contínua e tem inversa contínua

$$f^{-1} : \mathbb{R}^d \rightarrow \mathbb{S}^d \setminus \{N\}$$

dada por

$$f^{-1}(y^1, y^2, \dots, y^d) = \left(\frac{2y^1}{1 + \sum_{i=1}^d (y^i)^2}, \frac{2y^2}{1 + \sum_{i=1}^d (y^i)^2}, \dots, \frac{2y^d}{1 + \sum_{i=1}^d (y^i)^2}, \frac{-1 + \sum_{i=1}^d (y^i)^2}{1 + \sum_{i=1}^d (y^i)^2} \right).$$

Logo, $\mathbb{S}^d \setminus \{N\}$ é homeomorfo a \mathbb{R}^d .

Observação 1.106. Se considerarmos o símbolo ∞ e adiciona-lo a \mathbb{R}^d , obtemos o conjunto $\mathbb{R}^d \cup \{\infty\}$. Considere a função $f : \mathbb{S}^d \rightarrow \mathbb{R}^d \cup \{\infty\}$ que mapeia pontos $x \neq N$ para um ponto de \mathbb{R}^d pela projeção estereográfica e $N \mapsto \infty$.

Podemos definir uma topologia em $\mathbb{R}^d \cup \{\infty\}$ a partir de \mathbb{S}^d . Dizemos que $U \subset \mathbb{R}^d \cup \{\infty\}$ se $f^{-1}(U)$ é aberto na esfera. Isso define uma topologia em $\mathbb{R}^d \cup \{\infty\}$ de modo que \mathbb{S}^d e $\mathbb{R}^d \cup \{\infty\}$ são homeomorfos. Uma vizinhança de ∞ é um subconjunto U de $\mathbb{R}^d \cup \{\infty\}$ contendo ∞ tal que $U \setminus \{\infty\}$ é aberto de \mathbb{R}^d e existe $R > 0$ tal que

$$\{x \in \mathbb{R}^d : |x| > R\} \subset U.$$

Chamamos procedimentos como o descrito de compactificação.

No plano complexo \mathbb{C} , ao adicionar o ∞ , obtemos $\mathbb{C} \cup \{\infty\}$, que é uma esfera a menos de homeomorfismo. Quando se estuda funções da forma $f(z) = \frac{az+b}{cz+d}$, com $ad - bc \neq 0$, podemos estender f para uma função contínua de $\mathbb{C} \cup \{\infty\}$ em $\mathbb{C} \cup \{\infty\}$. Se faz isso da seguinte forma. Se $c = 0$ definimos

$$f(\infty) := \lim_{z \rightarrow \infty} \frac{az+b}{d} = \infty$$

e se $c \neq 0$ definimos

$$f(-d/c) := \lim_{z \rightarrow -d/c} \frac{az+b}{cz+d} = \infty,$$

$$f(\infty) := \lim_{z \rightarrow \infty} \frac{az+b}{cz+d} = \lim_{z \rightarrow \infty} \frac{a + \frac{b}{z}}{c + \frac{d}{z}} = \frac{a}{c}.$$

Além disso, essa função da esfera na esfera é homeomorfismo, porque é contínua (argumente via sequências) e tem inversa

$$f^{-1}(w) = \frac{dw - b}{-cw + a}.$$

Funções racionais como a descrita são chamadas de funções de Möbius. Por exemplo, se temos $f(z) = \frac{z}{3z-2}$, podemos definir $f(2/3) = \infty$ e $f(\infty) = 1/3$, assim obtendo uma função contínua da esfera na esfera.

Note que se temos

$$w = \frac{z}{3z-2},$$

então

$$3zw - 2w = z,$$

$$z = \frac{2w}{3w - 1}.$$

Logo, a inversa de f é dada por

$$f^{-1}(w) = \frac{2w}{3w - 1}.$$

Funções de Möbius são importantes porque preservam ângulo e orientação (mas não preservam distâncias).

Outros exemplos de homeomorfismos, que não provaremos aqui, são:

Exemplo 1.107. O toro, que em \mathbb{R}^3 pode ser dado por

$$(\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2,$$

com $0 < r < R$, é homeomorfo a $\mathbb{S}^1 \times \mathbb{S}^1$.

Exemplo 1.108. O cubo $Q = [-1, 1]^d \setminus (-1, 1)^d$ é homeomorfo a esfera \mathbb{S}^d .

Definição 1.109. Uma propriedade é invariante topológico, se for invariante por homeomorfismo.

A seguir estudaremos nosso primeiro invariante topológico, a conexidade.

1.7 Conexidade

Definição 1.110. Um espaço métrico X é desconexo se existem U_1 e U_2 abertos em X e não vazios tais que

$$U_1 \cap U_2 = \emptyset \quad \text{e} \quad X = U_1 \cup U_2.$$

Se X não é desconexo, então dizemos que X é conexo.

Repare que como $X \setminus U_1 = U_2$ e $X \setminus U_2 = U_1$, tanto U_1 quanto U_2 são fechados. Assim, poderíamos definir X como desconexo se esse pode ser escrito como a união de dois fechados não triviais disjuntos.

Um ponto $\{x\}$ é conexo, pois os únicos abertos no espaço $\{x\}$ são \emptyset e $\{x\}$. Já um espaço formado por dois pontos $\{x, y\}$ com a topologia discreta é desconexo porque $\{x, y\} = \{x\} \cup \{y\}$ é união de dois abertos (ou fechados) não triviais e disjuntos.

Proposição 1.111. Se X é desconexo, então há uma função contínua tal que $\text{cor} : X \rightarrow \{1, 2\}$ sobrejetora, onde o contra-domínio $\{1, 2\}$ tem a topologia induzida de \mathbb{R} , que é discreta. Reciprocamente, se tal função existe, X é desconexo.

Demonstração. Se X é desconexo, então temos abertos não triviais U_1, U_2 em X disjuntos tais que $X = U_1 \cup U_2$. Defina $\text{cor}(x) = 1$ se $x \in U_1$ e $\text{cor}(x) = 2$ se $x \in U_2$. Como pré-imagem de aberto é aberto, temos que f é contínua. De fato,

$$\text{cor}^{-1}(\{1\}) = U_1, \quad \text{cor}^{-1}(\{2\}) = U_2, \quad \text{cor}^{-1}(\emptyset) = \emptyset, \quad \text{cor}^{-1}(\{1, 2\}) = X.$$

Por outro lado, se tal função existe, então temos os abertos disjuntos e não vazios

$$U_1 := \text{cor}^{-1}(\{1\}) \quad \text{e} \quad U_2 := \text{cor}^{-1}(\{2\})$$

tais que $X = U_1 \cup U_2$. ■

A ideia da função cor é a seguinte, se X é desconexo, então conseguimos separá-lo com duas cores, aqui denotadas por 1 e 2, de modo contínuo.

Os conjuntos conexos de \mathbb{R} são os intervalos.

Observação 1.112. Um subconjunto I de \mathbb{R} é dito ser intervalo se para cada $a, b \in I$ com $a < b$ temos que $[a, b] \subset I$.

Exercício 1.113. Mostre que um intervalo é de um dos seguintes tipos:

$$\mathbb{R}, \quad (-\infty, b), \quad (-\infty, b], \quad (a, b), \quad (a, b], \quad [a, b], \quad [a, \infty), \quad (a, \infty).$$

Proposição 1.114. Um subconjunto $I \subset \mathbb{R}$ é conexo se, e só se, for I é intervalo.

Demonstração. Considere I conexo. Mostremos que I é um intervalo. Suponha que esse não é o caso. Existem $a, b \in I$ com $[a, b] \not\subset I$. Assim, há $c \notin I$ tal que $a < c < b$. Logo,

$$I = (I \cap (-\infty, c)) \cup (I \cap (c, \infty)),$$

ou seja, I é a união de dois abertos de I disjuntos e não triviais (observe que esses abertos são abertos de I , mas não necessariamente de \mathbb{R}). Logo, concluímos que I é desconexo, uma contradição.

Agora mostremos que se I é intervalo, então é conexo. Suponha que I é desconexo. Pela Proposição 1.111, existe $\text{cor} : I \rightarrow \{1, 2\}$ sobrejetora e contínua. Tome $a, b \in I$ tal que $\text{cor}(a) = 1$ e $\text{cor}(b) = 2$. A menos de trocar a e b de símbolos, podemos supor que $a < b$. Pelo teorema do valor intermediário aplicado a função $\text{cor}|_{[a,b]} : [a, b] \rightarrow \{1, 2\}$, existe c em (a, b) tal que $\text{cor}(c) = 1.5$, uma contradição, pois a imagem de cor é $\{1, 2\}$. ■

Agora provemos que imagem de conexo é conexo.

Proposição 1.115. Se $f : X \rightarrow Y$ é contínua e X conexo, então $f(X)$ é conexo.

Demonstração. Suponha que $f(X)$ é desconexo. Há uma função contínua sobrejetora $\text{cor} : f(X) \rightarrow \{1, 2\}$ pela Proposição 1.111. Assim, temos a função contínua e sobrejetora $\text{cor} \circ f$ de X em $\{1, 2\}$. Provando que X é desconexo. Contradizendo a hipótese do teorema. ■

Uma consequência imediata desse fato é que para funções contínuas de \mathbb{R} em \mathbb{R} , imagem de intervalo é sempre intervalo.

Observação 1.116. Conexidade é invariante topológico por causa da proposição acima. Se X e Y são homeomorfos e X é conexo, então Y é conexo também.

Podemos usar isso para provar a inexistência de homeomorfismos. Por exemplo, mostremos que \mathbb{R} não é homeomorfo a \mathbb{S}^1 . Suponha, por absurdo, que eles são: existe um homeomorfismo $f : \mathbb{S}^1 \rightarrow \mathbb{R}$. Removendo $N = (0, 1)$ do domínio e $f(N)$ do contra-domínio, obtemos o homeomorfismo $\mathbb{S}^1 \setminus \{N\} \rightarrow \mathbb{R} \setminus \{f(N)\}$. Note que temos os homeomorfismos

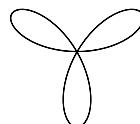
$$\mathbb{S}^1 \setminus \{N\} \simeq \mathbb{R} \quad \text{e} \quad \mathbb{R} \setminus \{f(N)\} \simeq (-\infty, 0) \cup (0, \infty),$$

mas o primeiro é conexo e o segundo não é. Assim, temos uma contradição.

Exercício 1.117. Mostre que a figura



não é homeomorfa a \mathbb{S}^1 . Mostre que a primeira figura não é homeomorfa a



Definição 1.118. Uma curva contínua em X é uma função contínua $\gamma : [0, 1] \rightarrow X$.

Proposição 1.119. Se $\gamma_1 : [0, 1] \rightarrow X$ e $\gamma_2 : [0, 1] \rightarrow X$ são curvas contínuas com $\gamma_1(1) = \gamma_2(0)$, então $\gamma_3 : [0, 1] \rightarrow X$ dada por

$$\gamma_3(t) = \begin{cases} \gamma_1(2t) & \text{se } 0 \leq t \leq 0.5 \\ \gamma_2(2t - 1) & \text{se } 0.5 \leq t \leq 1 \end{cases}$$

é uma curva contínua com $\gamma_3(0) = \gamma_1(0)$ e $\gamma_3(1) = \gamma_2(1)$.

Demonstração. É claro que γ_3 é contínua para $[0, 0.5] \cup (0.5, 1]$. Provemos sua continuidade em $t = 0.5$. Dada uma vizinhança V de $\gamma_1(1) = \gamma_2(0)$, existe $\delta > 0$ (que podemos tomar menor que 0.5) tal que

$$\gamma_1(2t) \in V \quad \text{para } t \in (0.5 - \delta, 0.5],$$

$$\gamma_2(2t - 1) \in V \quad \text{para } t \in [0.5, 0.5 + \delta),$$

pois tanto $\gamma_1(2t)$ quanto $\gamma_2(2t - 1)$ são contínuas em $t = 0.5$. Assim, temos que

$$\gamma_3((0.5 - \delta, 0.5 + \delta)) \subset V.$$

Logo, γ_3 é contínua em $t = 0.5$. ■

Em outras palavras, dadas duas curvas γ_1 e γ_2 tais que o final de γ_1 é o início de γ_2 , podemos junta-las em uma nova curva γ_3 . É comum denotar essa concatenação de curvas γ_3 por $\gamma_1\gamma_2$.

Definição 1.120. O espaço X é conexo por caminhos se para quaisquer dois pontos x e y de X existe uma curva conectando um ao outro, i.e., existe $\gamma : [0, 1] \rightarrow X$ tal que $\gamma(0) = x$ e $\gamma(1) = y$.

Note que se γ é um caminho que vai de x a y então $t \mapsto \gamma(1-t)$ é um caminho que vai de y a x .

Exercício 1.121. Mostre que $I \subset \mathbb{R}$ é conexo se, e só se, for conexo por caminhos.

Conexidade por caminhos é outro invariante topológico e costuma ser relacionado a conexidade. Sempre se tem que conexidade por caminhos implica em conexidade, mas a recíproca não é verdadeira (exemplo desse fato pode ser encontrado em livros de topologia geral). No entanto, para uma classe grande de casos, conexidade por caminhos é equivalente a conexidade. Mais precisamente, se um espaço é conexo e todo ponto tem uma vizinhança conexa por caminhos, então tal espaço é conexo por caminhos, como veremos adiante.

A vantagem da conexidade por caminhos está na facilidade de verificar sua validade.

O espaço \mathbb{R}^d é conexo por caminhos porque dois pontos podem ser ligados por um segmento de reta. A esfera \mathbb{S}^d também é conexa por caminhos.

A fim de ver isso, mostremos primeiro que se removermos um ponto z de \mathbb{S}^d obtemos que

$$\mathbb{S}^d \setminus \{z\} \simeq \mathbb{S}^d \setminus \{(0, \dots, 0, 1)\}.$$

De fato, fixe uma base ortonormal v_1, v_2, \dots, v_d com $v_d = z$. A isometria que leva e_i em v_i leva a esfera $\mathbb{S}^d \setminus \{z\}$ em $\mathbb{S}^d \setminus \{(0, \dots, 0, 1)\}$. Como esse mapa é a restrição de um isomorfismo linear, se trata de um homeomorfismo.

Agora mostremos que \mathbb{S}^d é conexo por caminhos. Dado dois pontos $x, y \in \mathbb{S}^d$, tomamos um ponto $z \in \mathbb{S}^d$ diferente de x, y . Repare que $\mathbb{S}^d \setminus \{z\}$ é homeomorfo a \mathbb{R}^d porque

$$\mathbb{S}^d \setminus \{z\} \simeq \mathbb{S}^d \setminus \{(0, \dots, 0, 1)\} \simeq \mathbb{R}^d.$$

Como esse é conexo por caminho, podemos ligar x a y em $\mathbb{S}^d \setminus \{z\}$ e assim estamos conectando x a y por um caminho em \mathbb{S}^d .

Exercício 1.122. Mostre que se $f : X \rightarrow Y$ é contínua e X é conexo por caminho, então $f(X)$ é conexo por caminhos. Em particular, se f for homeomorfismo, então Y tem de ser conexo por caminhos (e assim conexidade por caminhos é invariante topológico).

Proposição 1.123. Mostre que se X é conexo por caminhos, então X é conexo.

Demonstração. Suponha que X é desconexo. Existe $\text{col} : X \rightarrow \{1, 2\}$ contínua e sobrejetora. Tome $x, y \in X$ tal que $\text{col}(x) = 1$ e $\text{col}(y) = 2$. Seja $\gamma : [0, 1] \rightarrow X$ uma curva ligando x a y . Temos que $\text{col} \circ \gamma$ é um mapa contínuo e sobrejetor de $[0, 1]$ em $\{1, 2\}$, contradizendo que $[0, 1]$ é conexo. ■

Um espaço X é localmente conexo por caminhos se para todo x e toda vizinhança V de x existe uma vizinhança V' de x conexa por caminhos e contida em V . Todo aberto U de \mathbb{R}^d é localmente conexo por caminhos porque bolas em \mathbb{R}^d são conexas por caminhos e dado $x \in U$ existe $r > 0$ tal que $B(x, r) \subset U$.

Verifiquemos que bolas de \mathbb{R}^n são conexas por caminho para qualquer norma. Mais que isso, são convexas, i.e., dois pontos podem ser ligados por um segmento de reta. Seja \mathbf{p} uma norma em \mathbb{R}^d e considere a bola

$$\{z \in \mathbb{R}^d : \mathbf{p}(x - z) < r\}$$

de raio r centrada em x . Tome x_1, x_2 nessa bola. Temos que o caminho $\gamma : [0, 1] \rightarrow X$ dado por $\gamma(t) = (1-t)x + ty$ está inteiramente contido em tal bola. De fato,

$$\mathbf{p}(\gamma(t)) \leq (1-t)\mathbf{p}(x) + t\mathbf{p}(y) < (1-t)r + tr = r.$$

Proposição 1.124. Mostre que um aberto U em \mathbb{R}^d é conexo por caminhos quando for conexo.

Demonstração. Suponha que U não é conexo por caminhos. Fixe um ponto $x_0 \in U$. Considere

$$V_1 := \{x \in U : \text{existe um caminho conectando } x_0 \text{ a } x\}$$

Seja $V_2 := \mathbb{R}^d \setminus V_1$.

Se $x \in V_1$, então existe uma bola aberta B_x centrada em x contida em U , porque U é aberto. Como a bola é conexa por caminhos, todo ponto de B pode ser conectado a x , que por sua vez pode ser conectado a x_0 . Desta forma, $B_x \subset V_1$. Assim, V_1 é aberto, porque

$$V_1 = \bigcup_{x \in U_1} B_x,$$

e não vazio, porque $x_0 \in V_1$.

Por outro lado, V_2 é aberto também. Se $x \in V_2$, então não podemos conectar x a x_0 . Mas existe uma bola B_x centrada em x e contida em U . Todo ponto de B_x está conectado a x , que não pode ser conectado a x_0 . Assim, nenhum ponto da bola B_x pode ser conectado a x_0 . Portanto, $B_x \subset V_2$. Logo, V_2 é aberto e não vazio (pois estamos supondo que U não é conexo por caminhos). Portanto, $U = V_1 \cup V_2$, com V_1, V_2 abertos em U , disjuntos e não triviais. Logo, U é desconexo, uma contradição. ■

Pelo mesmo argumento, temos que, em \mathbb{R}^d , um aberto é conexo se, e só se, dois pontos podem ser conectados por um caminho poligonal, i.e., uma curva feita de segmentos de reta concatenados.

Exemplo 1.125. Uma álgebra real é um espaço vetorial real A munido de multiplicação satisfazendo:

- $(\alpha x)y = x(\alpha y) = \alpha(xy)$ para $\alpha \in \mathbb{R}, x, y \in A$;
- $x(y+z) = xy + xz$ e $(y+z)x = yx + zx$ para $x, y, z \in A$;

Dizemos que a álgebra A é unital se ela tem unidade: existe e tal que $ex = xe = x$ para $x \in A$. E uma álgebra unital A é de divisão se dados $a \neq 0$ e b em V , existem $x, x' \in V$ tais que $ax = b$ e $x'a = b$, ou seja, se é uma álgebra onde é possível resolver equações.

Exemplos de álgebras são vários. O produto vetorial em \mathbb{R}^3 é um exemplo de álgebra, mas ela não é unital. O espaço $C(X)$ de funções contínuas de X em \mathbb{R} é uma álgebra unital assim como o espaço de matrizes quadradas M_d .

Nenhuma dessas álgebras é de divisão. Exemplos de álgebras de divisão são \mathbb{R} e \mathbb{C} . Note que $\dim_{\mathbb{R}} \mathbb{R} = 1$ e $\dim_{\mathbb{R}} \mathbb{C} = 2$. Além dessas, existem somente outras duas de dimensão finita, a dos quaternions e a dos octonions, que tem dimensão 4 e 8.

Mostraremos que espaços tridimensionais não podem ter estrutura de álgebra de divisão. Suponha que \mathbb{R}^3 tem estrutura de álgebra de divisão real. Para cada $a \in \mathbb{R}^3$, existe a função $f_a : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $f_a(x) = ax$, que é linear. Note que $a \mapsto f_a$ é uma função de \mathbb{R}^3 no espaço $\text{Lin}(\mathbb{R}^3, \mathbb{R}^3)$, das transformações lineares de \mathbb{R}^3 em \mathbb{R}^3 (que pode ser identificado com M_3). Além disso, essa função é contínua.

De fato, na base canônica e_1, e_2, e_3 há constantes $c_{ij}^k \in \mathbb{R}$ tais que

$$e_i e_j = \sum_{1 \leq i, j, k \leq 3} c_{ij}^k e_k.$$

e escrevendo $a = \sum_i a^i e_i$ e $x = \sum_j x^j e_j$, temos $f_a(x) = \sum_{i,j,k} a^i x^j c_{ij}^k e_k$ e, consequentemente, identificamos f_a com a matriz

$$M(a) = \begin{bmatrix} \sum_j a^j c_{1j}^1 & \sum_j a^j c_{2j}^1 & \sum_j a^j c_{3j}^1 \\ \sum_j a^j c_{1j}^2 & \sum_j a^j c_{2j}^2 & \sum_j a^j c_{3j}^2 \\ \sum_j a^j c_{1j}^3 & \sum_j a^j c_{2j}^3 & \sum_j a^j c_{3j}^3 \end{bmatrix}.$$

A aplicação $a \mapsto f_a$ de \mathbb{R}^3 em $\text{Lin}(\mathbb{R}^3, \mathbb{R}^3)$ é contínua, pois $M(a)$ é contínua. Repare que para $a \in \mathbb{R}^3 \setminus \{0\}$, temos que f_a é sobrejetora. De fato, dado y , existe x tal que $ax = y$. Assim, f_a é isomorfismo linear e consequentemente

$$\det(M(a)) \neq 0 \quad \text{para } a \in \mathbb{R}^3 \setminus \{0\}.$$

Seja e a unidade de \mathbb{R}^3 , temos que

$$f_e x = ex = x \quad \text{e} \quad f_{-e} = (-e)x = -(ex) = -x.$$

Assim, $f_e = \text{Id}$ e $f_{-e} = -\text{Id}$ e, consequentemente,

$$M(e) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad M(-e) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Agora note que $\mathbb{R}^3 \setminus \{0\}$ é conexo por caminhos. Existe uma curva contínua $\gamma : [0, 1] \rightarrow \mathbb{R}^3 \setminus \{0\}$ tal que $\gamma(0) = \text{Id}$ e $\gamma(1) = -\text{Id}$. Como a função $a \mapsto \det(M(a))$ é contínua, temos que $t \mapsto \det M(\gamma(t))$ é função contínua.

Logo, temos uma curva contínua em \mathbb{R} ligando $\det M(e) = 1$ a $\det M(-e) = -1$, ou seja, há $t' \in [0, 1]$ tal que $\det M(\gamma(t')) = 0$, contradizendo que $\gamma(t') \in \mathbb{R}^3 \setminus \{0\}$. Portanto, \mathbb{R}^3 não pode admitir estrutura de álgebra de divisão real.

Antes de terminar essa seção gostaria de adicionar que muitas vezes se usa a expressão “componente conexa”. Se X é espaço métrico, então podemos definir uma relação entre pontos que é $x \sim y$ se existe $C \subset X$ conexo tal que $x, y \in C$. Essa relação é de equivalência e assim temos que X é união disjunta de suas classes de equivalência. Cada classe de equivalência é conexa. Assim, quebramos X como união de conjuntos conexos.

Similarmente, podemos quebrar X em componentes conexas por caminho.

Exercício 1.126. Verifique que a relação descrita acima é de equivalência.

Dica: Use que C é conexo se não existe uma função $\text{cor} : C \rightarrow \{1, 2\}$ sobrejetora e contínua.

Além disso, mostre que as classes laterais são conexas.

Exercício 1.127. Se X é localmente conexo (como é todo aberto de \mathbb{R}^d), então as componentes conexas são abertas.

Exercício 1.128. Mostre que um aberto de \mathbb{R}^d só pode ter uma quantidade enumerável de componentes.

1.8 Compacidade

Definição 1.129. Um espaço métrico X é compacto se toda sequência x_n em X admite subsequência convergente.

Observação 1.130. Em topologia geral, um espaço no qual toda sequência admite uma subsequência convergente é dito ser sequencialmente compacto. Um espaço é compacto em topologia geral se toda cobertura de X por abertos admite uma subcobertura finita. Como veremos posteriormente na Proposição 1.145, em espaços métricos, compacidade e compacidade sequencial são conceitos equivalentes.

Exemplo 1.131. Todo intervalo fechado e limitado $[a, b]$ é compacto. De fato, se x_n é sequência em $[a, b]$, ela admite subsequência convergente x_{n_k} por ser limitada (veja Teorema 1.11). Seja x o limite de x_{n_k} . O ponto x pertence a $[a, b]$ porque esse é fechado (um subconjunto A de um espaço métrico X é fechado se, e só se, toda sequência de A convergente em X converge em A , isto é, o limite da sequência pertence a A).

Proposição 1.132. Se X_1 e X_2 são espaços métricos compactos, então $X_1 \times X_2$ é compacto.

Demonstração. Se $x_n = (x_n^1, x_n^2)$ é uma sequência de $X_1 \times X_2$, então temos que x_n^1 tem subsequência convergente $x_{n_k}^1$ em X_1 . A sequência $x_{n_k}^2$ em X_2 , que é a sequência na segunda entrada com os índices de $n_1 < n_2 < \dots$, tem uma subsequência convergente x_{m_j} . Note que os índices $m_1 < m_2 < m_3 < \dots$ são tomados entre $n_1 < n_2 < n_3 < \dots$. Assim, $x_{m_j} = (x_{m_j}^1, x_{m_j}^2)$ converge em $X_1 \times X_2$. ■

Por indução, temos que se X_1, \dots, X_k são compactos, então $X_1 \times \dots \times X_k$ é compacto.

Exemplo 1.133. Uma caixa $[a_1, b_1] \times \dots \times [a_d, b_d]$ em \mathbb{R}^d , com $a_i < b_i$, é compacta por ser produto de compactos.

Proposição 1.134. Um subconjunto $A \subset \mathbb{R}^d$ é compacto se, e só se, A é fechado e limitado (com respeito a norma canônica).

Demonstração. Suponha A compacto. Considere uma sequência convergente $x_n \in A$ com limite x em \mathbb{R}^d . Como A é compacto, existe uma subsequência que converge em A e seu limite tem de ser x . Logo, $x \in A$. Agora mostremos que A é limitado. Suponha, por absurdo, que A é ilimitado. Para cada n , existe $x_n \in A$ tal que $|x_n| > n$. Essa sequência não tem subsequência convergente, uma contradição. Logo, se A é compacto, então A é fechado e limitado em \mathbb{R}^d .

Reciprocamente, suponha que A é fechado e limitado. Existe uma caixa compacta $Q = [a_1, b_1] \times \dots \times [a_d, b_d]$ contendo A . Se x_n é uma sequência de A , ela é sequência de Q e admite subsequência convergente em Q . Como A é fechado e os pontos da subsequência estão em A , o limite dessa subsequência pertence a A . Logo, A é compacto. ■

Exemplo 1.135. Como consequência, a esfera \mathbb{S}^d é compacta, porque é limitada e fechada. Ela é fechada porque a função $f(x) = \sum_{i=1}^{d+1} (x^i)^2$ de \mathbb{R}^{d+1} em \mathbb{R} é contínua e, consequentemente, $\mathbb{S}^d = f^{-1}(1)$ é fechado por ser pré-imagem de um fechado por uma função contínua.

Exercício 1.136. Mostre que se X é compacto e $A \subset X$ é fechado, então A é compacto.

Proposição 1.137. Se $f : X \rightarrow Y$ é contínua e X é compacto, então $f(X)$ é compacto.

Demonstração. Se y_n é uma sequência em $f(X)$, então para cada n existe $x_n \in X$ tal que $y_n = f(x_n)$. Como X é compacto, x_n admite uma subsequência x_{n_k} que converge para $x \in X$. Logo, $y_{n_k} = f(x_{n_k})$ converge para $f(x)$, ou seja, $f(X)$ é compacto. ■

Em particular, se $f : X \rightarrow Y$ é sobrejetora e contínua, Y é compacto. Assim, compacidade é invariante por homeomorfismo. Por exemplo, \mathbb{R}^d não é compacto (porque não é limitado em \mathbb{R}^d com a norma usual). Assim, \mathbb{S}^d não pode ser homeomorfo a \mathbb{R}^d .

Proposição 1.138. Se $f : X \rightarrow \mathbb{R}$ é contínua e X é compacto, então f atinge máximo e mínimo.

Demonstração. Sabemos que $f(X)$ é compacto. Assim $f(X)$ é fechado e limitado em \mathbb{R} . Seja $M = \sup_{x \in X} f(x)$. Existe uma sequência x_n tal que $f(x_n) \rightarrow M$. Como $f(X)$ é fechado, $M \in f(X)$. Logo, existe x tal que $f(x) = M$. O argumento para existência de mínimo é análogo. ■

Exemplo 1.139. Mostremos que normas em \mathbb{R}^d são Lipschitz equivalentes novamente. Seja \mathbf{p} uma norma. Pela desigualdade de Cauchy-Schwartz 1.17, temos que

$$\mathbf{p}(x) = \mathbf{p}\left(\sum_{i=1}^d x^i e_i\right) \leq \sum_{i=1}^d |x^i| \mathbf{p}(e_i) \leq \left(\sum_{i=1}^d \mathbf{p}(e_i)^2\right)^{\frac{1}{2}} |x|.$$

Assim, $\mathbf{p} : \mathbb{R}^d \rightarrow \mathbb{R}$ é contínua. Sejam $M > 0$ e $m > 0$ máximo e mínimo de \mathbf{p} na esfera \mathbb{S}^{d-1} . Daí segue que

$$m|x| \leq \mathbf{p}(x) \leq M|x|$$

para $x \in \mathbb{R}^d$, pois para $x = 0$ a desigualdade acima é válida e para $x \neq 0$, temos

$$m \leq \mathbf{p}\left(\frac{x}{|x|}\right) \leq M$$

porque $\frac{x}{|x|} \in \mathbb{S}^{d-1}$.

Proposição 1.140 (Lema do número de Lebesgue). Seja X compacto. Considere uma cobertura $\{U_\lambda\}_{\lambda \in \Lambda}$ de X por abertos, isto é,

$$X = \bigcup_{\lambda \in \Lambda} U_\lambda.$$

Existe $r > 0$, chamado número de Lebesgue, com a seguinte propriedade: dado $x \in X$ existe $\lambda \in \Lambda$ tal que $B(x, r) \subset U_\lambda$.

Demonstração. Suponha que o resultado seja falso: dado $r > 0$, existe x para o qual $B(x, r) \not\subset U_\lambda$ para todo λ . Assim, dado $n \in \mathbb{Z}_{\geq 1}$ existe x_n tal que $B(x_n, \frac{1}{n}) \not\subset U_\lambda$ para todo λ . Como X é compacto, x_n admite uma subsequência convergente x_{n_k} . Seja x seu limite e U_{λ_0} um aberto da cobertura contendo x . Existe $s > 0$ tal que $B(x, s) \subset U_{\lambda_0}$. Temos que existe k_0 tal que $x_{n_k} \in B(x, \frac{s}{2})$ e $\frac{1}{n_k} < \frac{s}{2}$ para $k \geq k_0$. Logo, para $y \in B(x_{n_{k_0}}, \frac{1}{n_{k_0}})$ temos

$$d_X(x, y) \leq d_X(x, x_{n_{k_0}}) + d_X(x_{n_{k_0}}, y) < \frac{s}{2} + \frac{s}{2} = s.$$

Portanto,

$$B\left(x_{n_{k_0}}, \frac{1}{n_{k_0}}\right) \subset B(x, s) \subset U_{\lambda_0},$$

uma contradição. ■

Corolário 1.141. Seja $f : X \rightarrow \mathbb{R}$ contínua com X compacto. A função f é uniformemente contínua, isto é, dado $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x_1) - f(x_2)| < \varepsilon$ para $x_1, x_2 \in X$ satisfazendo $d_X(x_1, x_2) < \delta$.

Demonstração. Dado x considere o aberto $U_x = f^{-1}(B(f(x), \frac{\varepsilon}{2}))$. Temos que $X = \cup_x U_x$. Assim, pela Proposição 1.140, existe $\delta > 0$ tal que $B(x, \delta)$ está contido em algum $U_{x'}$. Tome x_1, x_2 tal que $d_X(x_1, x_2) < \delta$. Existe $U_{x'}$ tal que $B(x_1, \delta) \subset U_{x'}$. Logo,

$$|f(x_1) - f(x_2)| \leq |f(x_1) - f(x')| + |f(x') - f(x_2)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

porque $x_1, x_2 \in B(x_1, \delta) \subset U_{x'}$. ■

Exemplo 1.142. Seja $f : X \times [a, b] \rightarrow \mathbb{R}$ contínua. A função $g(x) = \int_a^b f(x, t)dt$ é contínua. De fato, dado $\varepsilon > 0$, existe δ tal que $|f(x, t) - f(x', t')| < \varepsilon$ quando $d_X(x, x') < \delta$ e $|t - t'| < \delta$. Assim, $|g(x) - g(x')| \leq \int_a^b |f(x, t) - f(x', t)|dt \leq \varepsilon(b-a)$.

Exercício 1.143. Seja $f : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$ contínua. Suponha que existe $\varphi : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ contínua tal que para cada $x \in \mathbb{R}^d$ temos $|f(x, t)| \leq \varphi(t)$ para $t \in \mathbb{R}$. Além disso, suponha que $\int_{-\infty}^{\infty} \varphi(t)dt$ é finita. Mostre que $g(x) = \int_{-\infty}^{\infty} f(x, t)dt$ está bem definida e é contínua.

Lema 1.144. Se X é compacto e $r > 0$, então X pode ser coberto por finitas bolas de raio r .

Demonstração. Suponha que não. Tome $x_1 \in X$. Tome $x_2 \in X \setminus B(x_1, r)$, que existe porque X não pode ser coberto por uma bola. Similarmente, existe $x_3 \in X \setminus (B(x_1, r) \cup B(x_2, r))$. Indutivamente, existem x_1, \dots, x_n, \dots tais que x_k não pertence a $B(x_i, r)$ para $i < k$. Logo, para $i < k$, temos que $d(x_i, x_k) \geq r$. Portanto, x_n não admite uma subsequência convergente. ■

Proposição 1.145. O espaço X é compacto se, e só se, para cada cobertura $\{U_\lambda\}_\Lambda$ de X por abertos, isto é, $X = \bigcup_{\lambda \in \Lambda} U_\lambda$, existem $\lambda_1, \dots, \lambda_k$ tais que $X = \bigcup_{i=1}^k U_{\lambda_i}$.

Demonstração. Se X é compacto, então toda cobertura admite uma subcobertura finita de abertos. De fato, basta combinar Lema 1.140 com Lema 1.144.

Reciprocamente, assuma que toda cobertura de X por abertos admite cobertura finita e mostremos que X é compacto. De fato, suponha que X não é compacto: há uma sequência x_n que não admite subsequência convergente. Para cada $z \in X$ existe uma bola $B(z, r_z)$ tal que $\{n \in \mathbb{Z}_{\geq 1} : x_n \in B(z, r_z)\}$ é finito. Como $X = \bigcup_z B(z, r_z)$, há z_1, \dots, z_k tal que $X = \bigcup_i B(z_i, r_{z_i})$ e assim

$$\{n \in \mathbb{Z}_{\geq 1} : x_n \in X\} \subset \bigcup_{i=1}^k \{n \in \mathbb{Z}_{\geq 1} : x_n \in B(z_i, r_{z_i})\}$$

é finito, uma contradição. ■

1.9 Sequências de Cauchy

Definição 1.146. Uma sequência x_n em X é de Cauchy se para todo $\varepsilon > 0$ existe $N \in \mathbb{Z}_{\geq 1}$ tal que

$$d_X(x_n, x_m) < \varepsilon$$

para $n, m \geq N$.

Em outras palavras, x_n é de Cauchy se $\lim_{m,n \rightarrow \infty} d_X(x_n, x_m) = 0$.

Exemplo 1.147. É claro que toda sequência convergente é de Cauchy (verifique isso). Em \mathbb{R}^d a recíproca vale, toda sequência de Cauchy converge. De fato, se x_n de Cauchy, então existe N tal que $|x_n - x_N| < 1$ para $n \geq N$, ou seja, a sequência é limitada porque $x_n \in B(x_N, 1)$ exceto

finitos pontos. Assim, x_n é uma sequência contida em uma caixa compacta suficientemente grande. Em particular, x_n tem uma subsequência x_{n_k} convergente. Denotemos seu limite por x . Dado $\epsilon > 0$, existe N' tal que

$$|x_n - x_{n_k}| < \epsilon$$

para $n, n_k \geq N'$. Fazendo $k \rightarrow \infty$, temos $|x_n - x| \leq \epsilon$ para $n \geq N'$. Portanto, $x_n \rightarrow x$. Logo, em \mathbb{R}^d , toda sequência de Cauchy converge.

Definição 1.148. Um espaço métrico (X, d_X) é completo se toda sequência de Cauchy converge.

Portanto, \mathbb{R}^d é completo.

Exercício 1.149. Mostre que todo espaço métrico compacto é completo.

Observação 1.150. Um espaço completo pode ser homeomorfo a um espaço não completo, ou seja, ser completo não é invariante topológico. Por exemplo, considere $X = \{\frac{1}{n} : n \in \mathbb{Z}_{\geq 1}\}$ e $Y = \mathbb{Z}_{\geq 1}$ com as métricas induzidas de \mathbb{R} . O espaço Y é completo porque toda sequência de Cauchy é eventualmente constante, isto é, se y_n é de Cauchy em Y , então existe N tal que $y_n = y_N$ para $n \geq N$. Por outro lado, X possui a sequência de Cauchy $x_n = \frac{1}{n}$ que não converge em X . Logo, X não é completo. Os espaços X e Y são homeomorfos pelo mapa $X \rightarrow Y, x \mapsto x^{-1}$.

Por outro lado, se dois espaços X e Y são bi-Lipschitz isomorfos, isto é, existe uma bijeção $f : X \rightarrow Y$ tal que f e f^{-1} são Lipschitz, então X é completo se, e só se, Y é completo. Em particular, métricas Lipschitz equivalentes induzem as mesmas sequências de Cauchy.

Exercício 1.151. Considere a matriz quadrada A . Mostre que a série $\sum_{k=0}^{\infty} \frac{A^k}{k!}$ converge. Denotamos essa série por $\exp(A)$.

Dica: utilize a norma

$$|A| := \sup_{|x|=1} |Ax|$$

e que $|AB| \leq |A||B|$.

Exemplo 1.152. Mostremos que $\exp : M_d \rightarrow M_d$ é contínua.

Considere $\epsilon > 0$. Tome $N \in \mathbb{Z}_{\geq 1}$ tal que

$$\sum_{k>N} \frac{(|A|+1)^k}{k!} < \frac{\epsilon}{3}$$

Para uma matriz H tal que $|H| < 1$ temos

$$|\exp(A+H) - \exp(A)| \leq \sum_{k=0}^N \frac{|(A+H)^k - A^k|}{k!} + \frac{2\epsilon}{3}.$$

Tome $\delta \in (0, 1)$ tal que para $|H| < \delta$ temos

$$\frac{|(A+H)^k - A^k|}{k!} \leq \frac{\epsilon}{3N} \quad \text{para } k = 0, 1, \dots, N.$$

Assim, para $|H| < \delta$ temos

$$|\exp(A+H) - \exp(A)| < \epsilon,$$

ou seja, \exp é contínua em A .

Exercício 1.153. Use que matrizes diagonalizáveis complexas são densas no espaço de matrizes quadradas complexas para demonstrar que

$$\exp(\text{tr}(A)) = \det(\exp(A)).$$

Dica: Mostre que se $B = QAQ^{-1}$ para uma matriz invertível Q , então $\exp(\text{tr}(B)) = \exp(\text{tr}(A))$ e $\det(\exp(B)) = \det(\exp(A))$. Assim, podemos verificar que a fórmula acima vale para matrizes diagonais a fim de provar para matrizes diagonalizáveis. Provado para matrizes diagonais, utilize a densidade para concluir que a fórmula vale para qualquer matriz $A \in M_d(\mathbb{C})$, pois \exp é contínua.

Observação 1.154. O Exercício 1.153 é bastante importante porque nos permite computar álgebras de Lie de vários grupos de Lie matriciais, por exemplo. Aparece também ao resolver problemas variacionais envolvendo volume.

Exemplo 1.155. Seja X espaço métrico e $\text{BC}(X)$ o espaço das funções contínuas limitadas de X em \mathbb{R} , o símbolo BC significa “bounded continuous”. Claramente, $\text{BC}(X)$ é espaço vetorial e consideraremos nesse espaço a norma

$$\|f\|_{\infty} = \sup_{x \in X} |f(x)|.$$

O espaço $\text{BC}(X)$ é completo. Espaços vetoriais normados cujas métricas são completas são chamados de espaços de Banach. Assim $\text{BC}(X)$ é espaço de Banach.

Além disso, se X é compacto, então $\text{C}(X) = \text{BC}(X)$, isto é, toda função contínua é limitada. O espaço $\text{BC}(X)$ é completo (veja Teorema 4.1).

Observação 1.156. De forma geral, $\text{C}(X)$ é o espaço das funções contínuas de X em \mathbb{R} . As vezes se escreve $\text{C}(X, \mathbb{R})$ ou $\text{C}(X \rightarrow \mathbb{R})$. Mesmo quando X não é compacto $\text{C}(X)$ possui topologia natural e no caso em que X é aberto de \mathbb{R}^d , essa vem de uma métrica. No entanto, não adentrarei em tais tópicos aqui.

Capítulo 2

Diferenciabilidade

2.1 Curvas diferenciáveis

Definição 2.1. Uma curva (ou caminho) em \mathbb{R}^d é uma função contínua $\gamma : I \rightarrow \mathbb{R}^d$, onde I é um intervalo de \mathbb{R} . Escrevemos $\gamma(t) = (\gamma^1(t), \dots, \gamma^d(t))$. Dizemos que essa curva é diferenciável se γ^i é diferenciável para todo i . A derivada de γ é definida como

$$\frac{d\gamma}{dt} := \left(\frac{d\gamma^1}{dt}, \dots, \frac{d\gamma^d}{dt} \right).$$

As vezes denotamos a primeira derivada por $\dot{\gamma}$, a segunda derivada por $\ddot{\gamma}$, e assim por diante. Dizemos que γ é de classe C^r se cada γ^i é de classe C^r , isto é, as derivadas

$$\frac{d\gamma^i}{dt}, \frac{d^2\gamma^i}{dt^2}, \dots, \frac{d^r\gamma^i}{dt^r}$$

existem e são contínuas.

Exercício 2.2.

- Mostre que se α, β são curvas de classe C^1 , então

$$\frac{d}{dt} \langle \alpha, \beta \rangle = \langle \dot{\alpha}, \beta \rangle + \langle \alpha, \dot{\beta} \rangle.$$

- Mostre que

$$\frac{d|\alpha|^2}{dt} = 2 \langle \alpha, \dot{\alpha} \rangle.$$

- Mostre que se α não passa pela origem no instante t , então

$$\frac{d|\alpha|}{dt} = \left\langle \frac{\alpha}{|\alpha|}, \dot{\alpha} \right\rangle.$$

- Para $d = 3$, mostre que

$$\frac{d\alpha \times \beta}{dt} = \dot{\alpha} \times \beta + \alpha \times \dot{\beta}.$$

A velocidade de uma curva γ é a primeira derivada $\dot{\gamma}$. A segunda derivada $\ddot{\gamma}$ é a aceleração.

Exemplo 2.3 (Força central). Curvas no plano. Em vez de trabalhar com \mathbb{R}^2 , trabalharemos com os números complexos $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$.

Considere uma partícula com massa m e com trajetória γ que não passa pela origem. Temos que em coordenadas polares:

$$\gamma(t) = r(t) \exp(i\theta(t)),$$

onde $r > 0$ e $\exp(i\alpha) := \cos(\alpha) + i \sin(\alpha)$. A velocidade e a aceleração são:

$$\dot{\gamma} = \dot{r} \exp(i\theta(t)) + r \dot{\theta} i \exp(i\theta(t)),$$

$$\ddot{\gamma} = (\ddot{r} - r^2 \dot{\theta}^2) \exp(i\theta(t)) + (2\dot{r}\dot{\theta} + r\ddot{\theta}) i \exp(i\theta(t)).$$

O vetor $\exp(i\theta(t))$ é unitário e radial, isto é, $\exp(i\theta(t)) = \frac{\gamma}{|\gamma|}$. O vetor $i \exp(i\theta(t))$ é obtido ao rotacionar $\exp(i\theta(t))$ no sentido anti-horário por 90° .

Uma força central $f : \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}^d$ é um campo vetorial (uma função que associa a cada x um vetor $f(x)$) da forma $f(x) = \varphi(|x|) \frac{x}{|x|}$ onde φ é uma função de $\mathbb{R}_{>0}$ em \mathbb{R} . Em outras palavras, a força é central se $f(x)$ é sempre paralelo a x . Se a força f age sobre um corpo de massa m percorrendo a trajetória γ , então, pela segunda lei de Newton,

$$m\ddot{\gamma} = f(\gamma).$$

Assim, se γ descreve o movimento de uma partícula em \mathbb{C} sob a ação de uma força central, o que quer dizer que $\ddot{\gamma}$ aponta na direção radial, então

$$2\dot{r}\dot{\theta} + r\ddot{\theta} = 0$$

ou seja,

$$\frac{d}{dt} r^2 \dot{\theta} = 0,$$

$r^2 \dot{\theta}$ é constante.

A grandeza $L = mr^2\dot{\theta}$ se chama momento angular e, como vimos, é constante para forças centrais. Forças centrais são extremamente importantes. Dois exemplos distinguidos são:

$$f(x) = -kx \quad (\text{Lei de Hook})$$

$$f(x) = \frac{k}{|x|^2} \frac{x}{|x|} \quad (\text{Lei do inverso do quadrado}).$$

Integração de curvas é feita coordenada a coordenada

$$\int_a^b \gamma(t) dt := \left(\int_a^b \gamma^1(t) dt, \dots, \int_a^b \gamma^d(t) dt \right).$$

É óbvio que o teorema fundamental do cálculo vale:

$$\int_a^b \dot{\gamma}(t) dt = \gamma(b) - \gamma(a),$$

$$\frac{d}{ds} \int_a^s \gamma(t) dt = \gamma(s),$$

porque a integral de curvas é definida coordenada a coordenada.

Proposição 2.4. Considere a curva γ definida em $[a, b]$. Temos que

$$\left| \int_a^b \gamma(t) dt \right| \leq \int_a^b |\gamma(t)| dt.$$

Demonstração. Se $\int_a^b \gamma(t) dt = 0$, então a identidade acima é trivialmente verdadeira. Podemos assumir que tal vetor é não nulo.

Para qualquer vetor v , temos

$$\left\langle \int_a^b \gamma(t) dt, v \right\rangle = \sum_i v^i \int_a^b \gamma^i(t) dt$$

$$\begin{aligned}
&= \sum_i \int_a^b \gamma^i(t) v^i dt \\
&= \int_a^b \langle \gamma(t), v \rangle dt \\
&\leq \int_a^b |\gamma(t)| |v| dt \\
&= \int_a^b |\gamma(t)| dt |v|
\end{aligned}$$

por causa da desigualdade de Cauchy-Schwartz 1.17.

Tomando

$$v = \int_a^b \gamma(t) dt$$

obtemos

$$\begin{aligned}
\left| \int_a^b \gamma(t) dt \right|^2 &\leq \int_a^b |\gamma(t)| dt \left| \int_a^b \gamma(t) dt \right|, \\
\left| \int_a^b \gamma(t) dt \right| &\leq \int_a^b |\gamma(t)| dt.
\end{aligned}$$

■

Proposição 2.5 (Teorema do valor médio). Considere a curva $\gamma : [a, b] \rightarrow \mathbb{R}^d$ de classe C^1 . Temos que

$$|\gamma(b) - \gamma(a)| \leq \max_{a \leq t \leq b} |\dot{\gamma}(t)|(b - a).$$

Demonstração. Pelo teorema fundamental do cálculo:

$$\gamma(b) - \gamma(a) = \int_a^b \dot{\gamma}(t) dt.$$

Logo, pela Proposição 2.4, temos

$$|\gamma(b) - \gamma(a)| \leq \int_a^b |\dot{\gamma}(t)| dt \leq \max_{a \leq t \leq b} |\dot{\gamma}(t)|(b - a).$$

■

O comprimento de uma curva $\gamma : [a, b] \rightarrow \mathbb{R}^d$ de classe C^1 é

$$L(\gamma, [a, b]) = \int_a^b |\dot{\gamma}(t)| dt.$$

Pelo teorema fundamental do cálculo e a Proposição 2.4

$$|\gamma(b) - \gamma(a)| \leq \int_a^b |\dot{\gamma}| dt = L(\gamma, [a, b]),$$

ou seja, o comprimento de uma curva ligando $\gamma(a)$ a $\gamma(b)$ é maior que ou igual a $|\gamma(b) - \gamma(a)|$.

Uma reparametrização é uma função $\varphi : [a', b'] \rightarrow [a, b]$ bijetora de classe C^1 . Note que

$$L(\gamma, [a, b]) = L(\gamma \circ \varphi, [a', b'])$$

porque

$$\int_a^b |\dot{\gamma}| dt = \int_{a'}^{b'} |\dot{\gamma}(\varphi(s))| \left| \frac{d\varphi}{ds} \right| ds = \int_{a'}^{b'} \left| \frac{d}{ds} \gamma(\varphi(s)) \right| ds.$$

Definição 2.6. Uma curva γ é regular se $\dot{\gamma}(t) \neq 0$ para todo t .

Curvas regulares são interessantes porque podemos reparametrizá-las de modo que a nova curva tenha velocidade constante e igual a 1. De fato, suponha que a curva γ está

definida no intervalo $[a, b]$. Seja L o comprimento de γ e considere $s(t) = \int_a^t |\dot{\gamma}(u)| du$. Temos que $\frac{ds}{dt} = |\dot{\gamma}(t)| > 0$, ou seja, s é estritamente crescente. Além disso, $s(a) = 0$ e $s(b) = L$. A inversa $\tau : [0, L] \rightarrow [a, b]$ de s tem derivada

$$\frac{d\tau}{ds} = \frac{1}{|\dot{\gamma}(\tau(s))|}.$$

E assim, a curva $\dot{\gamma}(s) = \gamma(\tau(s))$ tem velocidade com módulo 1.

$$|\dot{\gamma}(s)| = |\dot{\gamma}(\tau(s))| \frac{d\tau}{ds} = 1.$$

Uma curva regular com velocidade 1 em valor absoluto é chamada de curva parametrizada por comprimento de arco.

Exemplo 2.7. A curva $\gamma(s) = \exp(is)$ é parametrizada por comprimento de arco porque $|\dot{\gamma}| = |i \exp(is)| = 1$.

Exemplo 2.8. Considere uma curva $\gamma : \mathbb{R} \rightarrow \mathbb{C}$ parametrizada por comprimento de arco. Em particular, a aceleração é perpendicular a velocidade:

$$|\dot{\gamma}(s)|^2 = 1,$$

$$\langle \dot{\gamma}(s), \ddot{\gamma}(s) \rangle = 0.$$

Assim, não há aceleração tangencial. Como $\ddot{\gamma}$ é paralelo a $i\dot{\gamma}$, existe uma função $k(s)$ tal que

$$\ddot{\gamma}(s) = k(s)i\dot{\gamma}(s).$$

A função k se chama curvatura. O valor absoluto de $k(s)$ é dado por

$$|k(s)| = |\ddot{\gamma}(s)|.$$

A curvatura em s_0 mede o raio do círculo que melhor aproxima a curva γ em s_0 . Esse círculo leva o nome de círculo osculante. O raio desse círculo é $r_0 = \frac{1}{|k(s_0)|}$, onde assumimos $k(s_0) \neq 0$, e o seu centro é

$$c_0 := \gamma(s_0) + r_0 i\dot{\gamma}(s_0).$$

O círculo osculante parametrizado por comprimento de arco é

$$\lambda(s) := c(s_0) - r_0 i\dot{\gamma}(s_0) \exp\left(i \frac{s - s_0}{r_0}\right).$$

Note que

$$\lambda(s_0) = \gamma(s_0),$$

$$\dot{\lambda}(s_0) := \dot{\gamma}(s_0),$$

$$\ddot{\lambda}(s_0) := \frac{i\dot{\gamma}(s_0)}{r_0} = \ddot{\gamma}(s_0),$$

ou seja, o círculo osculante aproxima-se de γ no sentido que

$$|\gamma(s) - \lambda(s)| = o(|s - s_0|^2).$$

Observação 2.9. Seja $f : U \rightarrow \mathbb{R}^k$ uma função contínua definida em uma vizinhança aberta U de x_0 em um espaço métrico. Assuma que f não se anula para $x \neq x_0$. Ao escrevermos uma função $g : U \rightarrow \mathbb{R}^l$ satisfaz

$$g(x) = o(f(x))$$

queremos dizer que

$$\lim_{x \rightarrow x_0} \frac{g(x)}{|f(x)|} = 0.$$

Assim, ao escrevermos $|\gamma(s) - \lambda(s)| = o(|s - s_0|^2)$ acima, queremos dizer que

$$\frac{|\gamma(s) - \lambda(s)|}{|s - s_0|^2} \rightarrow 0$$

quando $|s - s_0| \rightarrow 0$.

Se $k(s_0) = 0$, então aproximamos γ pela reta $\lambda(s) = \gamma(s_0) + \dot{\gamma}(s_0)(s - s_0)$ e temos

$$|\gamma(s) - \lambda(s)| = o(|s - s_0|^2).$$

Podemos interpretar esse cenário como sendo um círculo osculante de raio infinito. De fato, para $r(s_0)$ muito grande temos

$$\lambda(s) := c(s_0) - r_0 i \dot{\gamma}(s_0) \exp\left(i \frac{s - s_0}{r(s_0)}\right) \simeq c(s_0) - r_0 i \dot{\gamma}(s_0) \left(1 + i \frac{s - s_0}{r(s_0)}\right) = \gamma(s_0) + \dot{\gamma}(s_0)(s - s_0).$$

Como discutimos na secção sobre conexidade, temos que um aberto é conexo se, e só se, é conexo por caminhos. Além disso, sempre podemos tomar caminhos poligonais. Podemos também tomar caminhos regulares simples (simples quer dizer que o caminho não se auto-interseca), mas provar a existência de tais caminhos é não trivial: segue de dois fatos, todo aberto admite uma métrica riemanniana completa e dois pontos podem ser ligados por uma geodésica minimizante de velocidade 1 na presença de uma métrica riemanniana completa.

Uma curva contínua $\gamma : [a, b] \rightarrow \mathbb{R}^d$ é C^1 por partes se há $a = t_1 < t_2 < \dots < t_k = b$ tais que γ é de classe C^1 em (t_i, t_{i+1}) e os limites laterais

$$\lim_{t \rightarrow t_i^+} \dot{\gamma}(t), \quad \lim_{t \rightarrow t_i^-} \dot{\gamma}(t)$$

existem. Curvas poligonais são exemplos de curvas C^1 por partes.

Note que comprimento de curva está definido para curvas C^1 por partes.

$$L(\gamma, [a, b]) := \sum_i \int_{t_i}^{t_{i+1}} |\dot{\gamma}(t)| dt.$$

De forma geral, se V é uma função de $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ em \mathbb{R}^k , por exemplo, podemos definir

$$\int_a^b V(t, \gamma(t), \dot{\gamma}) dt := \sum_i \int_{t_i}^{t_{i+1}} V(t, \gamma(t), \dot{\gamma}(t)) dt.$$

Podemos considerar V 's que dependam de $\ddot{\gamma}$ ou derivadas superiores. Mas em geral, usamos apenas até a primeira derivada.

Exemplo 2.10. O trabalho de uma força $f : U \rightarrow \mathbb{R}^d$, com U aberto em \mathbb{R}^d , ao longo de um caminho γ em U , que assumimos ser C^1 por partes, é definido como sendo

$$W(f, \gamma) := \int_a^b \langle f(\gamma(t)), \dot{\gamma} \rangle dt.$$

Se a força é central, isto é, $U = \mathbb{R}^d \setminus \{0\}$ e $f(x) = \varphi(|x|) \frac{x}{|x|}$ para uma função $\varphi : \mathbb{R}_{>0} \rightarrow \mathbb{R}$, então $W(f, \gamma) = 0$ para todo caminho γ fechado. Mostremos que esse é o caso. Considere $V(r) = \int_1^r \varphi(s) ds$. Temos que

$$\frac{d}{dt} V(|\gamma(t)|) = \varphi(|\gamma(t)|) \left\langle \frac{\gamma(t)}{|\gamma(t)|}, \dot{\gamma}(t) \right\rangle$$

Assim, como $\gamma(a) = \gamma(b)$, temos

$$W(f, \gamma) = \int_a^b \varphi(|\gamma(t)|) \left\langle \frac{\gamma(t)}{|\gamma(t)|}, \dot{\gamma}(t) \right\rangle dt = \int \frac{d}{dt} V(|\gamma(t)|) dt = V(|\gamma(b)|) - V(|\gamma(a)|) = 0.$$

Exercício 2.11. Mostre que se o trabalho para curvas fechadas é sempre zero, então o trabalho para ir de x_0 a x_1 independe da curva conectando esses pontos.

Exercício 2.12. Mostre que se γ é C^1 por partes, então

$$\int_a^b \dot{\gamma}(t) dt = \gamma(b) - \gamma(a).$$

2.2 Funções diferenciáveis

Definição 2.13. Considere U aberto em \mathbb{R}^p e a função $f : U \rightarrow \mathbb{R}^q$. A função f é diferenciável em $x_0 \in U$ se existe uma transformação \mathbb{R} -linear $\lambda : \mathbb{R}^p \rightarrow \mathbb{R}^q$ tal que

$$f(x_0 + h) = f(x_0) + \lambda(h) + o(h)$$

para $x_0 + h \in U$. Em outras palavras,

$$f(x_0 + h) - f(x_0) - \lambda(h) = o(h),$$

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \lambda(h)|}{|h|} = 0.$$

Veja Observação 2.9 sobre a notação o .

A transformação linear λ é única (veja Proposição 2.14), se chama **diferencial de f em x_0** (ou derivada de Fréchet em x_0) e a denotamos por df_{x_0} . Existem várias outras notações para o diferencial, como $df(x_0)$, Df_{x_0} , $Df(x_0)$, $f'(x_0)$ etc.

Se f é diferenciável para todo $x_0 \in U$, então dizemos que f é diferenciável em U .

Proposição 2.14. A transformação linear λ descrita acima é única.

Demonstração. Suponha que existem λ_1, λ_2 satisfazendo

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \lambda_i(h)|}{|h|} = 0.$$

Temos

$$\lim_{h \rightarrow 0} \frac{|\lambda_1(h) - \lambda_2(h)|}{|h|} = 0$$

porque

$$|\lambda_1(h) - \lambda_2(h)| \leq |f(x_0 + h) - f(x_0) - \lambda_1(h)| + |f(x_0 + h) - f(x_0) - \lambda_2(h)|.$$

Considere $v \in \mathbb{R}^p \setminus \{0\}$ e $h_n = \frac{1}{n}v$. Temos

$$\frac{|\lambda_1(v) - \lambda_2(v)|}{|v|} = \lim_{n \rightarrow \infty} \frac{|\lambda_1(h_n) - \lambda_2(h_n)|}{|h_n|} = 0,$$

ou seja, $\lambda_1(v) = \lambda_2(v)$ para $v \neq 0$.

Como $\lambda_1 0 = \lambda_2 0 = 0$, obtemos $\lambda_1 = \lambda_2$. ■

Geometricamente, o diferencial fornece a melhor aproximação afim de f em $f(x_0)$. Mais precisamente, temos que a função

$$x \mapsto f(x)$$

é bem aproximada pela aplicação afim

$$x \mapsto f(x_0) + df_{x_0}(x - x_0)$$

onde bem aproximado quer dizer que

$$f(x) = f(x_0) + df_{x_0}(x - x_0) + o(x - x_0).$$

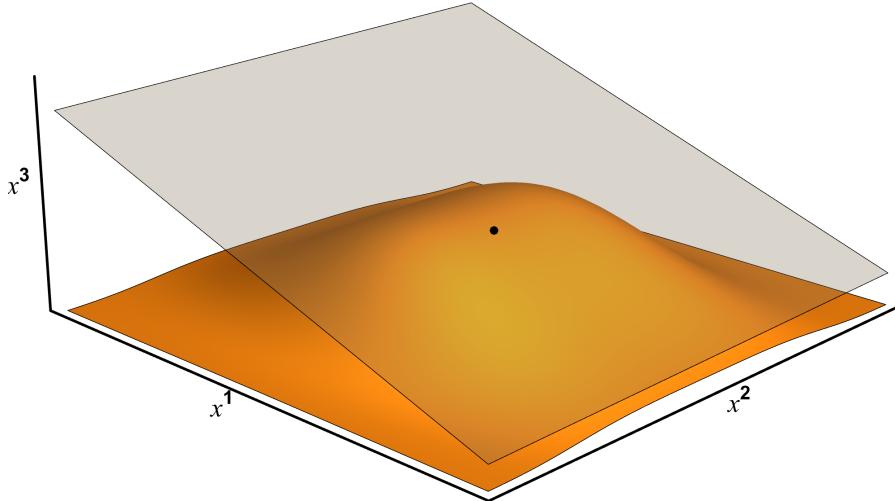


Figura 2.1: A função $f(x)$ e sua aproximação afim $f(x_0) + df_{x_0}(x - x_0)$ no ponto x_0 .

Exemplo 2.15. Considere $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ dada por $f(x) = Ax$, onde A é matriz $q \times p$. Temos que $df_x = A$ porque

$$A(x + h) = A(x) + A(h).$$

Exemplo 2.16. Considere $f : \mathbb{R}^p \rightarrow \mathbb{R}$ dada por $f(x) = \langle Ax, x \rangle$, onde A é transformação linear de $p \times p$. Temos que

$$f(x + h) - f(x) = \langle A(x + h), x + h \rangle - \langle A(x), x \rangle = \langle Ax, h \rangle + \langle Ah, x \rangle + \langle Ah, h \rangle.$$

Como $\langle Ah, h \rangle = o(h)$, temos

$$df_x(h) = \langle Ax, h \rangle + \langle Ah, x \rangle = \langle Ax + A^T x, h \rangle.$$

Quando A é simétrica, temos $df_x(h) = 2\langle Ax, h \rangle$.

Exemplo 2.17. Considere $f(x^1, x^2) = (x^1)^2 + x^1 x^2$.

$$\begin{aligned} f(x + h) - f(x) &= (x^1 + h^1)^2 + (x^1 + h^1)(x^2 + h^2) - (x^1)^2 - x^1 x^2 \\ &= 2x^1 h^1 + (h^1)^2 + x^1 h^2 + h^1 x^2 + h^1 h^2 = (2x^1 + x^2)h^1 + x^1 h^2 + o(h). \end{aligned}$$

Portanto,

$$df_x(h) = \begin{bmatrix} 2x^1 + x^2 & x^1 \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \end{bmatrix}.$$

Exercício 2.18. Identifique \mathbb{C} e \mathbb{R}^2 Compute a derivada da função $f(z) = z^2$ em z . Faça o mesmo para $f(z) = z^n$.

Exemplo 2.19. Se f é diferenciável em x_0 , então f é contínua em x_0 . De fato,

$$\lim_{x \rightarrow x_0} f(x) = \lim_{h \rightarrow 0} f(x_0 + h) = \lim_{h \rightarrow 0} f(x_0) + df_{x_0}(h) + o(h) = f(x_0),$$

pois

$$df_{x_0}(h) = \sum_{i=1}^p h^i df_{x_0}(e_i) \rightarrow 0 \quad \text{quando } h \rightarrow 0$$

e

$$o(h) = |h| \frac{o(h)}{|h|} \rightarrow 0 \quad \text{quando } h \rightarrow 0.$$

Definição 2.20. Considere a função $f : U \rightarrow \mathbb{R}^q$, onde U é aberto de \mathbb{R}^p . A derivada parcial $\frac{\partial f}{\partial x^i}(x_0)$ de f com respeito a variável x^i no ponto x_0 é o limite

$$\lim_{t \rightarrow 0} \frac{f(x_0^1, \dots, x_0^{i-1}, x_0^i + t, x_0^{i+1}, \dots, x_0^p)}{t},$$

se esse existe. É costumeiro denotar $\frac{\partial f}{\partial x^i}(x_0)$ por $\partial_i f(x_0)$ quando não houver ambiguidade sobre qual variável estamos usando no domínio.

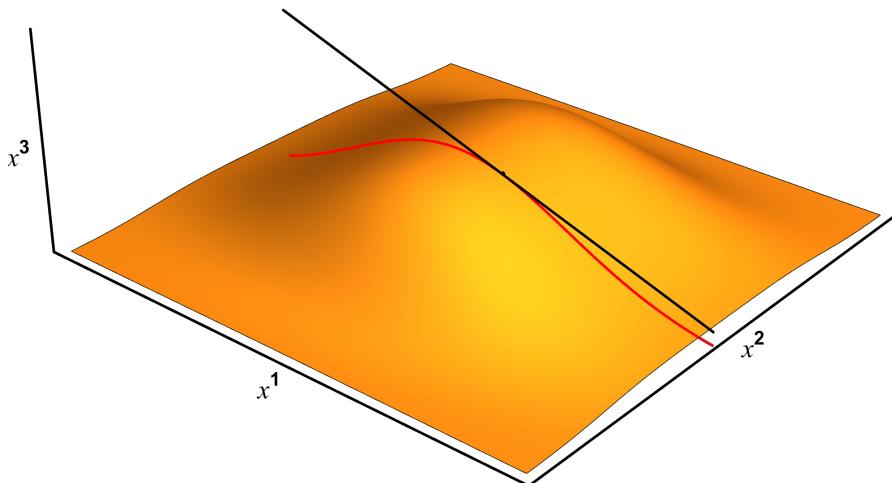


Figura 2.2: $\partial_1 f(x_0)$ é o coeficiente angular da curva $x^1 \mapsto f(x^1, x_0^2)$ em vermelho no instante $x^1 = x_0^1$. A curva em vermelho é a intersecção de $x^3 = f(x^1, x^2)$ com o plano $x^2 = x_0^2$.

As derivadas parciais de uma função diferenciável $f : U \rightarrow \mathbb{R}^q$ sempre existem. De fato, para a base canônica e_i de \mathbb{R}^p , temos

$$f(x_0 + te_i) = f(x_0) + d_{x_0} f(te_i) + o(te_i),$$

$$f(x_0 + te_i) = f(x_0) + td_{x_0} f(e_i) + o(t),$$

$$\frac{f(x_0 + te_i) - f(x_0)}{t} = d_{x_0} f(e_i) + \frac{o(t)}{t} \rightarrow d_{x_0} f(e_i).$$

Portanto $\frac{\partial f}{\partial x^i}(x_0) = df_{x_0}(e_i)$.

Assim,

$$df_{x_0}(h) = \sum_{j=1}^p \frac{\partial f}{\partial x^j}(x_0) h^j.$$

Se considerarmos a função coordenada $x^j : \mathbb{R}^p \rightarrow \mathbb{R}$, dada por $x^j(a) = a^i$, temos

$$dx_{x_0}^j(h) = h^j,$$

que não depende de x_0 . Assim, denotamos $dx_{x_0}^j$ por dx^j apenas.

Desta forma, temos

$$df_{x_0} = \sum_{j=1}^p \frac{\partial f}{\partial x^j}(x_0) dx^j.$$

Exemplo 2.21. Para $f(x^1, x^2) = (x^1)^2 + x^1 x^2$, temos

$$df_{(x^1, x^2)} = (2x^1 + x^2) dx^1 + x^1 dx^2.$$

Observação 2.22. Considere $g_1, \dots, g_p : U \rightarrow \mathbb{R}$ de classe C^∞ . Para cada x temos o funcional linear

$$\omega_x = \sum_{i=1}^p g_i(x) dx^i.$$

Chamamos essa função $\omega : U \rightarrow \mathbb{R}$ de 1-forma. Para uma função $f : U \rightarrow \mathbb{R}$ de classe C^∞ , temos que

$$df = \sum_{i=1}^p \frac{\partial f}{\partial x^i} dx^i$$

é uma 1-forma.

Definição 2.23. A matriz Jacobiana de $f : U \rightarrow \mathbb{R}^l$ diferenciável em x_0 é a matriz

$$\begin{bmatrix} \frac{\partial f^1}{\partial x^1}(x_0) & \frac{\partial f^1}{\partial x^2}(x_0) & \cdots & \frac{\partial f^1}{\partial x^p}(x_0) \\ \frac{\partial f^2}{\partial x^1}(x_0) & \frac{\partial f^2}{\partial x^2}(x_0) & \cdots & \frac{\partial f^2}{\partial x^p}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^q}{\partial x^1}(x_0) & \frac{\partial f^q}{\partial x^2}(x_0) & \cdots & \frac{\partial f^q}{\partial x^p}(x_0) \end{bmatrix}$$

associada a df_{x_0} , isto é,

$$df_{x_0}(h) = \begin{bmatrix} \frac{\partial f^1}{\partial x^1}(x_0) & \frac{\partial f^1}{\partial x^2}(x_0) & \cdots & \frac{\partial f^1}{\partial x^p}(x_0) \\ \frac{\partial f^2}{\partial x^1}(x_0) & \frac{\partial f^2}{\partial x^2}(x_0) & \cdots & \frac{\partial f^2}{\partial x^p}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^q}{\partial x^1}(x_0) & \frac{\partial f^q}{\partial x^2}(x_0) & \cdots & \frac{\partial f^q}{\partial x^p}(x_0) \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ \vdots \\ h^k \end{bmatrix}.$$

Observação 2.24. Denotamos uma matriz

$$\begin{bmatrix} a_1^1 & a_2^1 & \cdots & a_p^1 \\ a_1^2 & a_2^2 & \cdots & a_p^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^q & a_2^q & \cdots & a_p^q \end{bmatrix}$$

por (a_j^i) , onde i é o índice das linhas e j das colunas. Se temos um produto de matrizes $(c_j^i) = (a_j^i)(b_j^i)$, então

$$c_j^i = \sum_k a_k^i b_j^k.$$

No caso da matriz jacobiana, escrevemos

$$df = \left(\frac{\partial f^i}{\partial x^j} \right).$$

Teorema 2.25. Regra da cadeia: Sejam $f : U \rightarrow V$ e $g : V \rightarrow \mathbb{R}^m$ funções diferenciáveis em x_0 e $y_0 = f(x_0)$, onde $U \subset \mathbb{R}^p$ e $V \subset \mathbb{R}^q$ são abertos. Temos que

$$d(g \circ f)_{x_0} = dg_{y_0} \circ df_{x_0}.$$

Demonstração. Defina

$$r(k) = \lim_{k \rightarrow 0} \frac{g(y_0 + k) - g(y_0) - dg_{y_0}(k)}{|k|}$$

para $k \neq 0$ e $r(0) = 0$. Como g é diferenciável em y_0 , $r(k) \rightarrow 0$ quando $k \rightarrow 0$, ou seja, r é contínua em 0. Portanto, temos

$$g(y_0 + k) = g(y_0) + dg_{y_0}(k) + |k|r(k).$$

Usando a diferenciabilidade de f em x_0 junto à formula acima, obtemos

$$\begin{aligned} g(f(x_0 + h)) &= g(f(x_0) + df_{x_0}(h) + o(h)) \\ &= g(y_0) + dg_{y_0}(df_{x_0}(h) + o(h)) + |df_{x_0}(h) + o(h)| r(df_{x_0}(h) + o(h)), \\ &= g(y_0) + dg_{y_0}(df_{x_0}(h)) + dg_{y_0}(o(h)) + |df_{x_0}(h) + o(h)| r(df_{x_0}(h) + o(h)). \end{aligned}$$

Note que

$$\frac{dg_{y_0}(o(h)) + |df_{x_0}(h) + o(h)| r(df_{x_0}(h) + o(h))}{|h|} \rightarrow 0,$$

pois,

$$\begin{aligned} \frac{dg_{y_0}(o(h))}{|h|} &= dg_{y_0}\left(\frac{o(h)}{|h|}\right) \rightarrow 0, \\ \frac{|df_{x_0}(h) + o(h)|}{|h|} &\leq \left|df_{x_0}\left(\frac{h}{|h|}\right)\right| + \frac{|o(h)|}{|h|} \leq \max_{|v|=1} |df_{x_0}(v)| + 1 \end{aligned}$$

para h suficientemente pequeno, e

$$r(df_{x_0}(h) + o(h)) \rightarrow 0.$$

Logo,

$$g(f(x_0 + h)) = g(y_0) + dg_{y_0}(df_{x_0}(h)) + o(h).$$

■

Usando x e y para denotar as coordenadas de U e V , temos que

$$d(f \circ g)_{x_0} = dg_{y_0} \circ df_{x_0}$$

se torna

$$\left(\frac{\partial(g^i \circ f)}{\partial x^j}(x_0) \right) = \left(\frac{\partial g^i}{\partial y^j}(y_0) \right) \left(\frac{\partial f^i}{\partial x^j}(x_0) \right),$$

ou seja,

$$\begin{aligned} \frac{\partial(g^i \circ f)}{\partial x^j}(x_0) &= \sum_{k=1}^q \frac{\partial g^i}{\partial y^k}(y_0) \frac{\partial f^k}{\partial x^j}(x_0), \\ \frac{\partial(g \circ f)}{\partial x^j}(x_0) &= \sum_{k=1}^q \frac{\partial g}{\partial y^k}(y_0) \frac{\partial f^k}{\partial x^j}(x_0). \end{aligned}$$

Escrevendo $z = g \circ f(y)$ e $y = f(x)$, temos

$$\frac{\partial z^i}{\partial x^j}(x_0) = \sum_{k=1}^q \frac{\partial z^i}{\partial y^k}(y_0) \frac{\partial y^k}{\partial x^j}(x_0).$$

Se f e g são diferenciáveis em U e V , então é comum omitir o ponto:

$$\frac{\partial z^i}{\partial x^j} = \sum_{k=1}^q \frac{\partial z^i}{\partial y^k} \frac{\partial y^k}{\partial x^j}.$$

Exemplo 2.26. Sejam $U \subset \mathbb{R}^p$ aberto e $f : U \rightarrow \mathbb{R}^q$ diferenciável. Considere uma curva $\gamma : (-\epsilon, \epsilon) \rightarrow U$ de classe C^1 tal que $\gamma(0) = x_0$ e $\dot{\gamma}(0) = h$. Pela regra da cadeia, temos

$$\frac{d(f \circ \gamma)}{dt}(0) = \sum_{k=1}^p \frac{\partial f}{\partial x^k}(\gamma(0)) \frac{d\gamma^k}{dt}(0),$$

ou seja,

$$\frac{d(f \circ \gamma)}{dt}(0) = \sum_{k=1}^p \frac{\partial f}{\partial x^k}(x_0) h^k = df_{x_0}(h).$$

Observação 2.27. Considere $f : U \rightarrow \mathbb{R}$ diferenciável em x_0 . A derivada direcional de f em x_0 na direção $v \in \mathbb{R}^p$ é definida como

$$\frac{\partial f}{\partial v}(x_0) := \left. \frac{d}{dt} \right|_{t=0} f(x_0 + tv),$$

ou seja, pela regra da cadeia temos

$$\frac{\partial f}{\partial v}(x_0) = df_{x_0}(v).$$

O gradiente de f em x_0 é o vetor

$$\nabla f(x_0) := \sum_{k=1}^p \frac{\partial f}{\partial x^k}(x_0) e_k,$$

onde e_1, \dots, e_p é a base canônica de \mathbb{R}^p , ou seja,

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x^1}(x_0), \frac{\partial f}{\partial x^2}(x_0), \dots, \frac{\partial f}{\partial x^p}(x_0) \right).$$

Portanto,

$$\frac{\partial f}{\partial v}(x_0) = df_{x_0}(v) = \langle \nabla f(x_0), v \rangle.$$

Definição 2.28. Considere a função $f : U \rightarrow \mathbb{R}^q$, onde U é aberto de \mathbb{R}^p . A função f é de classe C^1 se for contínua e as derivadas parciais $\frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \dots, \frac{\partial f}{\partial x^p}$ existem em são funções contínuas. Se f é contínua é todas as derivadas parciais até ordem k existem e são contínuas, então dizemos que f é de classe C^k . Se todas as derivadas parciais existem e são contínuas, então dizemos que f é de classe C^∞ ou suave. De forma geral, denotamos o espaço das funções de classe C^k de U em \mathbb{R}^q por $C^k(U, \mathbb{R}^q)$.

Exemplo 2.29. Existem funções que são diferenciáveis e não são de classe C^1 . Por exemplo, $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

A derivada de f em 0 vale 0:

$$f'(0) = \lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} h \sin\left(\frac{1}{h}\right) = 0.$$

Para $x \neq 0$,

$$f'(x) = 2x \sin\left(\frac{1}{x}\right) + x^2 \left(\frac{-1}{x^2}\right) \cos\left(\frac{1}{x}\right) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

Note que f' é descontínua e portanto f não é de classe C^1 .

Para ver que f é diferenciável em $x_0 \in \mathbb{R}$, note que definindo

$$r(h) = f(x_0 + h) - f(x_0) - f'(x_0)h$$

temos $r(h)/|h| \rightarrow 0$, ou seja,

$$f(x_0 + h) = f(x_0) + f'(x_0)h + o(h).$$

Portanto, f é diferenciável com diferencial $d_{x_0}f(h) = f'(x_0)h$, mas não é de classe C^1 .

Para fazer um exemplo em duas variáveis, podemos tomar $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $g(x, y) = f(x)$.

Temos as derivadas parciais $\frac{\partial g}{\partial x}(x_0, y_0) = f'(x_0)$ e $\frac{\partial g}{\partial y}(x_0, y_0) = 0$, que é descontínua.

Além disso,

$$g(x_0 + r, y_0 + s) = f(x_0 + r) = f(x_0) + f'(x_0)r + o(|r|) = g(x_0, y_0) + f'(x_0)r + o((r, s)),$$

ou seja, possui derivadas parciais, g é diferenciável e não é de classe C^1 .

Exemplo 2.30. Um exemplo de função contínua que possui derivadas parciais mas não é diferenciável é o seguinte. Considere a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

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

Essa função é contínua em \mathbb{R}^2 , porque $\frac{x^2}{x^2 + y^2}, \frac{y^2}{x^2 + y^2} \leq 1$. Suas derivadas parciais são

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$$

e para $(x, y) \neq 0$

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

Note que as derivadas parciais são descontínuas em $(0, 0)$. Por exemplo,

$$\lim_{x \rightarrow 0} \frac{\partial f}{\partial y}(x, 0) = \lim_{x \rightarrow 0} \frac{x^2(x^2 + 2x0 - 0^2)}{(x^2 + 0^2)^2} = 1 \neq \frac{\partial f}{\partial y}(0, 0).$$

Além disso, f não é diferenciável em $(0, 0)$. De fato, suponha que por contradição que f é diferenciável em $(0, 0)$. Temos $df_{(0,0)}(r, s) = \frac{\partial f}{\partial x}(0, 0)r + \frac{\partial f}{\partial y}(0, 0)s = 0$ e, portanto,

$$f(r, s) = o((r, s)),$$

ou seja,

$$\lim_{(r,s) \rightarrow (0,0)} \frac{f(r, s)}{\sqrt{r^2 + s^2}} = 0.$$

Isso não é verdade, por exemplo,

$$\lim_{t \rightarrow 0^+} \frac{f(t, t)}{\sqrt{t^2 + t^2}} = \lim_{t \rightarrow 0^+} \frac{2t^3}{(2t^2)^{3/2}} = \frac{1}{\sqrt{2}} \neq 0.$$

Assim, há funções que possuem derivadas parciais, mas não são diferenciáveis.

Teorema 2.31. Se $f : U \rightarrow \mathbb{R}^q$ é de classe C^1 , então f é diferenciável.

Demonstração. Pelo teorema fundamental do cálculo

$$\begin{aligned} f(x + h) - f(x) &= \int_0^1 \frac{d}{dt} (f(x + th)) dt \\ &= \sum_{k=1}^p h^k \int_0^1 \frac{\partial f}{\partial x^k}(x + th) dt. \end{aligned}$$

Logo,

$$\begin{aligned} f(x+h) - f(x) - df_x(h) &= \int_0^1 \frac{d}{dt} (f(x+th)) dt \\ &= \sum_{k=1}^p h^k \int_0^1 \left(\frac{\partial f}{\partial x^k}(x+th) - \frac{\partial f}{\partial x^k}(x) \right) dt. \end{aligned}$$

Provando que

$$\lim_{h \rightarrow 0} \int_0^1 \left(\frac{\partial f}{\partial x^k}(x+th) - \frac{\partial f}{\partial x^k}(x) \right) dt = 0,$$

obtemos que f é diferenciável

$$f(x+h) = f(x) + df_x(h) + o(h).$$

Para computar o limite acima, considere uma bola compacta B centrada em x contida em U . Por causa da compacidade de B , temos que $\frac{\partial f}{\partial x^k} : B \rightarrow \mathbb{R}^q$ é uniformemente contínua, isto é, dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\left| \frac{\partial f}{\partial x^k}(x_2) - \frac{\partial f}{\partial x^k}(x_1) \right| < \varepsilon$$

para $|x_1 - x_2| < \delta$, $x_1, x_2 \in B$. Assim, para $|h| < \delta$, temos

$$\left| \int_0^1 \left(\frac{\partial f}{\partial x^k}(x+th) - \frac{\partial f}{\partial x^k}(x) \right) dt \right| \leq \int_0^1 \left| \frac{\partial f}{\partial x^k}(x+th) - \frac{\partial f}{\partial x^k}(x) \right| dt < \varepsilon.$$

■

Observação 2.32. Considere uma transformação linear $A : \mathbb{R}^p \rightarrow \mathbb{R}^q$. A norma de operador é definida pela fórmula

$$\|A\| := \max\{|A(v)| \in \mathbb{R} : |v| = 1\}.$$

Para $x \in \mathbb{R}^p$, temos

$$|A(x)| \leq \|A\| |x|.$$

De fato, a identidade vale para $x = 0$ trivialmente e para $x \neq 0$ temos

$$\left| A\left(\frac{x}{|x|}\right) \right| \leq \|A\|,$$

ou seja,

$$|A(x)| \leq \|A\| |x|.$$

Se temos as transformações lineares $A : \mathbb{R}^q \rightarrow \mathbb{R}^m$ e $B : \mathbb{R}^p \rightarrow \mathbb{R}^q$

$$|AB(x)| \leq \|A\| |B(x)| \leq \|A\| \|B\| |x|.$$

Portanto,

$$\|AB\| \leq \|A\| \|B\|$$

pois

$$\|AB\| = \max\{|AB(x)| \in \mathbb{R} : |x| = 1\} \leq \|A\| \|B\|.$$

Teorema 2.33. (Teorema do valor médio) Sejam $f : U \rightarrow \mathbb{R}^q$ de classe C^1 e $C \subset U$ compacto e convexo. Para

$$M = \max\{\|df_x\| \in \mathbb{R} : x \in C\},$$

que existe porque é o máximo da função contínua $x \mapsto \|df_x\|$ no compacto C , temos

$$|f(y) - f(x)| \leq M |y - x|.$$

Observação 2.34. A função $x \mapsto \|df_x\|$ é contínua. Considere a norma L^1

$$\|A\|_1 = \sum_{i,j} |a_j^i|.$$

Como normas são equivalentes em um espaço vetorial de dimensão finita, temos que existe $L > 0$ tal que $\|A\| \leq L\|A\|_1$, ou seja,

$$\|df_x - df_{x_0}\| \leq L \sum_{i,j} \left| \frac{\partial f^i}{\partial x^j}(x) - \frac{\partial f^i}{\partial x^j}(x_0) \right| \rightarrow 0$$

quando $x \rightarrow x_0$.

Demonstração. Como

$$f(y) - f(x) = \int_0^1 \frac{d}{dt} (f(x + t(y-x))) dt = \int_0^1 df_{x+t(y-x)}(y-x) dt$$

obtemos

$$|f(y) - f(x)| \leq \int_0^1 |df_{x+t(y-x)}(y-x)| dt \leq \int_0^1 \|df_{x+t(y-x)}\| |y-x| dt \leq M |y-x|. \quad \blacksquare$$

Em particular, como podemos considerar bolas compactas em U , temos que se f é de classe C^1 , então f é localmente Lipschitz.

Exemplo 2.35. Se $f : U \rightarrow \mathbb{R}^q$ é função de classe C^1 , $df \equiv 0$, e $U \subset \mathbb{R}^p$ é aberto conexo, então f é constante. De fato, pelo teorema do valor médio, em bolas compactas temos $M = 0$ e f é constante, isto é, dado $x \in U$ temos que f é constante em uma vizinhança aberta de x .

Considere $x_0 \in U$ e

$$V := \{x \in U : f(x) = f(x_0)\}.$$

Note que $x_0 \in V$ e, portanto, $V \neq \emptyset$. Além disso, V é fechado em U porque $V = f^{-1}(x_0)$ é aberto em U porque f é localmente constante. Assim, V é não vazio, aberto e fechado em U , ou seja, $V = U$ pois U é conexo.

Lema 2.36. Seja $g : [a, b] \times [c, d] \rightarrow \mathbb{R}$ contínua. Suponha que $\frac{\partial g}{\partial y}$ existe e é contínua. Então temos que

$$G(x) := \int_c^d g(x, y) dy, \quad x \in [a, b]$$

é de classe C^1 com

$$G'(x) = \int_c^d \frac{\partial g}{\partial x}(x, y) dy.$$

Demonstração. A prova segue da continuidade uniforme de g e $\frac{\partial g}{\partial x}$. Dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\left| \frac{\partial g}{\partial x}(x_2, y_2) - \frac{\partial g}{\partial x}(x_1, y_1) \right| < \varepsilon$$

para $|x_2 - x_1| < \delta$ e $|y_1 - y_2| < \delta$. Para $|h| < \delta$ temos

$$\left| \frac{G(x+h) - G(x) - h \int_c^d \frac{\partial g}{\partial x}(x, y) dy}{|h|} \right| \leq \int_c^d \left| \frac{g(x+h, y) - g(x, y)}{h} - \frac{\partial g}{\partial x}(x, y) \right| dy \leq (d-c)\varepsilon$$

porque a desigualdade segue do teorema do valor médio na reta. De fato, existe $\theta \in (0, 1)$ tal que

$$\left| \frac{g(x+h, y) - g(x, y)}{h} - \frac{\partial g}{\partial x}(x, y) \right| = \left| \frac{\partial g}{\partial x}(x + \theta h, y) - \frac{\partial g}{\partial x}(x, y) \right| < \varepsilon.$$

Logo, G é diferenciável e

$$G'(x) = \int_c^d \frac{\partial g}{\partial x}(x, y) dy.$$

Para ver que G' é contínua, note que

$$|G'(x+h) - G'(x)| \leq \int_c^d \left| \frac{\partial g}{\partial x}(x+h, y) - \frac{\partial g}{\partial x}(x, y) \right| dy < (d-c)\varepsilon$$

■

Em particular,

$$\frac{\partial}{\partial x} \int_c^d g(x, y) dy = \int_c^d \frac{\partial g}{\partial x}(x, y) dy$$

Teorema 2.37. (Teorema de Schwarz/Teorema de Clairaut/Teorema de Young)

Se $f : U \rightarrow \mathbb{R}^q$ de classe C^2 , onde $U \subset \mathbb{R}^p$ é aberto, então

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}.$$

Demonstração. Podemos supor que $k = 2$ e $l = 1$. Usemos coordenadas (x, y) em \mathbb{R}^2 . Tome $(x_0, y_0) \in U$ e uma caixa $B = [x_0 - r, x_0 + r] \times [y_0 - r, y_0 + r] \subset U$. Fixe $y \in [y_0 - r, y_0 + r]$. Pelo teorema fundamental do cálculo, temos

$$f(x, y) - f(x, y_0) = \int_{y_0}^y \frac{\partial f}{\partial y}(x, v) dv.$$

Pelo Lemma 2.36 temos

$$\frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x, y_0) = \int_{y_0}^y \frac{\partial^2 f}{\partial x \partial y}(x, v) dv$$

e, portanto, derivando com respeito y , obtemos

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

■

2.3 Fórmula de Taylor

Lema 2.38. (Teorema do valor médio para integrais) Seja $f : [a, b] \rightarrow \mathbb{R}$ de classe C^1 e $p : [a, b] \rightarrow \mathbb{R}$ contínua. Suponha que $p(x) > 0$ exceto um número finito de pontos. Existe $c \in (a, b)$ tal que

$$\int_a^b f(x)p(x) dx = f(c) \int_a^b p(x) dx.$$

Demonstração. Sejam M e m o máximo e mínimo de f . Temos

$$m \leq \frac{\int_a^b f(x)p(x) dx}{\int_a^b p(x) dx} \leq M.$$

Se f é constante, então $f(x) = M = m$ e qualquer $c \in (a, b)$ funciona. Podemos supor que f é não constante.

Existem $c, d \in [a, b]$, com $c < d$, tal que $f(x) < M$ para $c \leq x \leq d$. Temos

$$\int_a^b (M - f(x))p(x) dx \geq \int_c^d (M - f(x))p(x) dx > 0,$$

ou seja,

$$\int_a^b f(x)p(x)dx < M \int_a^b p(x)dx.$$

Similarmente,

$$m \int_a^b p(x)dx < \int_a^b f(x)p(x)dx.$$

Assim,

$$m < \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx} < M.$$

Seja x_1 um ponto de máximo e x_2 um ponto de mínimo. Assim, pelo teorema do valor intermediário, há c entre x_1 e x_2 com $c \neq x_1$ e $c \neq x_2$ tal que

$$f(c) = \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx}.$$

Em particular, $c \in (a, b)$. ■

Corolário 2.39. (Teorema do valor médio em \mathbb{R}) Para $f : [a, b] \rightarrow \mathbb{R}$ de classe C^1 , existe $c \in (a, b)$ tal que

$$f(b) - f(a) = f'(c)(b - a).$$

Demonstração. Pelo teorema fundamental do cálculo,

$$f(b) - f(a) = \int_a^b f'(x)dx.$$

Aplicando Lema 2.38 com $p \equiv 1$, existe $c \in (a, b)$ tal que

$$f(b) - f(a) = f'(c)(b - a). ■$$

Lema 2.40. Considere $\varphi : [0, 1] \rightarrow \mathbb{R}$ de classe C^{k+1} . Temos que

$$\varphi(1) = \sum_{i=0}^k \frac{1}{i!} \frac{d^i \varphi}{dt^i}(0) + \int_0^1 \frac{(1-t)^k}{k!} \frac{d^{k+1} \varphi}{dt^{k+1}}(t) dt.$$

Além disso, há $\theta \in (0, 1)$ tal que

$$\int_0^1 \frac{(1-t)^k}{k!} \frac{d^{k+1} \varphi}{dt^{k+1}}(t) dt = \frac{1}{(k+1)!} \frac{d^{k+1} \varphi}{dt^{k+1}}(\theta).$$

Demonstração. Aplicando Lemma 2.38 para $p(t) = \frac{(1-t)^k}{k!}$, existe $\theta \in (0, 1)$ tal que

$$\int_0^1 \frac{(1-t)^k}{k!} \frac{d^{k+1} \varphi}{dt^{k+1}}(t) dt = \frac{d^{k+1} \varphi}{dt^{k+1}}(\theta) \int_0^1 \frac{(1-t)^k}{k!} dt = \frac{1}{(k+1)!} \frac{d^{k+1} \varphi}{dt^{k+1}}(\theta),$$

mostrando a identidade no final do enunciado.

A identidade

$$\varphi(1) = \sum_{i=0}^k \frac{1}{i!} \frac{d^i \varphi}{dt^i}(0) + \int_0^1 \frac{(1-t)^k}{k!} \frac{d^{k+1} \varphi}{dt^{k+1}}(t) dt$$

segue do teorema fundamental do cálculo:

$$\varphi(1) = \varphi(0) + \int_0^1 \frac{d\varphi}{dt}(t) dt.$$

Note que

$$\frac{d}{dt} \left((1-t) \frac{d\varphi}{dt}(t) \right) = -\frac{d\varphi}{dt}(t) + (1-t) \frac{d^2 \varphi}{dt^2}(t),$$

e, integrando, obtemos

$$\begin{aligned}-\frac{d\varphi}{dt}(0) &= - \int_0^1 \frac{d\varphi}{dt}(t) dt + \int_0^1 (1-t) \frac{d^2\varphi}{dt^2}(t) dt, \\ \int_0^1 \frac{d\varphi}{dt}(t) dt &= \frac{d\varphi}{dt}(0) + \int_0^1 (1-t) \frac{d^2\varphi}{dt^2}(t) dt.\end{aligned}$$

Portanto,

$$\varphi(1) = \varphi(0) + \frac{d\varphi}{dt}(0) + \int_0^1 (1-t) \frac{d^2\varphi}{dt^2}(t) dt.$$

Suponhamos que o resultado para $n < k$ é provado para $n+1$, isto é, estamos supondo que

$$\varphi(1) = \sum_{i=0}^n \frac{1}{i!} \frac{d^i\varphi}{dt^i}(0) + \int_0^1 \frac{(1-t)^n}{n!} \frac{d^{n+1}\varphi}{dt^{n+1}}(t) dt.$$

Como

$$\frac{d}{dt} \left(\frac{(1-t)^{n+1}}{(n+1)!} \frac{d^{n+1}\varphi}{dt^{n+1}}(t) \right) = -\frac{(1-t)^n}{n!} \frac{d^{n+1}\varphi}{dt^{n+1}}(t) + \frac{(1-t)^{n+1}}{(n+1)!} \frac{d^{n+2}\varphi}{dt^{n+2}}(t),$$

obtemos, integrando, a seguinte identidade

$$\begin{aligned}-\frac{1}{(n+1)!} \frac{d^{n+1}\varphi}{dt^{n+1}}(0) &= - \int_0^1 \frac{(1-t)^n}{n!} \frac{d^{n+1}\varphi}{dt^{n+1}}(t) dt + \int_0^1 \frac{(1-t)^{n+1}}{(n+1)!} \frac{d^{n+2}\varphi}{dt^{n+2}}(t) dt, \\ \int_0^1 \frac{(1-t)^n}{n!} \frac{d^{n+1}\varphi}{dt^{n+1}}(t) dt &= \frac{1}{(n+1)!} \frac{d^{n+1}\varphi}{dt^{n+1}}(0) + \int_0^1 \frac{(1-t)^{n+1}}{(n+1)!} \frac{d^{n+2}\varphi}{dt^{n+2}}(t) dt.\end{aligned}$$

Portanto,

$$\begin{aligned}\varphi(1) &= \sum_{i=0}^n \frac{1}{i!} \frac{d^i\varphi}{dt^i}(0) + \frac{1}{(n+1)!} \frac{d^{n+1}\varphi}{dt^{n+1}}(0) + \int_0^1 \frac{(1-t)^{n+1}}{(n+1)!} \frac{d^{n+2}\varphi}{dt^{n+2}}(t) dt \\ &= \sum_{i=0}^{n+1} \frac{1}{i!} \frac{d^i\varphi}{dt^i}(0) + \int_0^1 \frac{(1-t)^{n+1}}{(n+1)!} \frac{d^{n+2}\varphi}{dt^{n+2}}(t) dt.\end{aligned}$$

Logo, o resultado é verdadeiro. ■

Teorema 2.41. (Fórmula de Taylor em \mathbb{R}) Seja I um intervalo, $x_0 \in I$ e $x_0 + h \in I$. Considere $f : I \rightarrow \mathbb{R}$ de classe C^{k+1} . Temos

$$f(x_0 + h) = \sum_{i=0}^k \frac{1}{i!} \frac{d^i f}{dx^i}(x_0) h^i + h^{k+1} \int_0^1 \frac{(1-t)^k}{k!} \frac{d^{k+1} f}{dx^{k+1}}(x_0 + th) dt.$$

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

$$h^{k+1} \int_0^1 \frac{(1-t)^k}{k!} \frac{d^{k+1} f}{dx^{k+1}}(x_0 + th) dt = \frac{1}{(k+1)!} \frac{d^{k+1} f}{dx^{k+1}}(x_0 + \theta h) h^{k+1}.$$

Demonstração. Considere $\varphi(t) = f(x_0 + th)$. O resultado segue do Lema 2.40. Basta notar que

$$\frac{d^i \varphi}{dt^i}(t) = \frac{d^i f}{dx^i}(x_0 + th) h^i.$$

■

Em particular,

$$f(x_0 + h) = \sum_{i=0}^k \frac{1}{i!} \frac{d^i f}{dx^i}(x_0) h^i + o(|h|^k).$$

Agora, provemos a fórmula de Taylor para várias variáveis.

Considere $f : U \rightarrow \mathbb{R}$ de classe C^{k+1} , com $U \subset \mathbb{R}^p$ aberto, temos que

$$\frac{d}{dt} f(x + th) = \left(\sum_{i=1}^p h^i \partial_i \right) f(x + th).$$

Assim, para $n = 0, 1, \dots, k, k+1$, obtemos

$$\frac{d^n}{dt^n} f(x + th) = \left(\sum_{i=1}^p h^i \partial_i \right)^n f(x + th).$$

Teorema 2.42. (Fórmula de Taylor) Considere uma função $f : U \rightarrow \mathbb{R}$ de classe C^{k+1} , onde $U \subset \mathbb{R}^p$ é aberto convexo. Se $x_0, x_0 + h \in U$, então temos

$$f(x_0 + h) = \sum_{n=0}^k \frac{1}{n!} \left(\sum_{i=1}^p h^i \partial_i \right)^n f(x_0) + \int_0^1 \frac{(1-t)^k}{k!} \left(\sum_{i=1}^p h^i \partial_i \right)^{k+1} f(x_0 + th) dt.$$

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

$$\int_0^1 \frac{(1-t)^k}{k!} \left(\sum_{i=1}^p h^i \partial_i \right)^{k+1} f(x_0 + th) dt = \frac{1}{(k+1)!} \left(\sum_{i=1}^p h^i \partial_i \right)^{k+1} f(x_0 + \theta h) h^\alpha.$$

Demonstração. Considere $\varphi(t) = f(x_0 + th)$. Basta notar que

$$\frac{d^n \varphi}{dt^n}(t) = \left(\sum_{i=1}^p h^i \partial_i \right)^n f(x_0 + th)$$

e usar o Lema 2.40. ■

Se f é de classe C^{k+1} , então para cada $x_0, x_0 + h \in U$ existe $\theta \in (0, 1)$ tal que

$$f(x_0 + h) = \sum_{n=0}^k \frac{1}{n!} \left(\sum_{i=1}^p h^i \partial_i \right)^n f(x_0) + \frac{1}{(k+1)!} \left(\sum_{i=1}^p h^i \partial_i \right)^{k+1} f(x_0 + \theta h).$$

Como

$$\lim_{h \rightarrow 0} \frac{\left(\sum_{i=1}^p h^i \partial_i \right)^{k+1} f(x_0 + \theta h) - \left(\sum_{i=1}^p h^i \partial_i \right)^{k+1} f(x_0)}{|h|^{k+1}} = 0,$$

temos

$$f(x_0 + h) = \sum_{n=0}^{k+1} \frac{1}{n!} \left(\sum_{i=1}^p h^i \partial_i \right)^n f(x_0) + o(|h|^{k+1}).$$

Definição 2.43. O Hessiano de $f : U \rightarrow \mathbb{R}$ de classe C^2 é a forma quadrática

$$\mathbf{Hess} f_x(h) = \left(\sum_{i=1}^p h^i \partial_i \right)^2 f(x) = \sum_{1 \leq i, j \leq p} h^i h^j \partial_i \partial_j f(x).$$

A matriz Hessiana de f é a matriz simétrica

$$H(f)_x = [\partial_i \partial_j f(x)],$$

onde i, j são os índices de linhas e colunas respectivamente. Temos

$$\begin{aligned} \mathbf{Hess} f_x(h) &= \begin{bmatrix} h^1 & h^2 & \cdots & h^p \end{bmatrix} \begin{bmatrix} \partial_1 \partial_1 f(x) & \partial_1 \partial_2 f(x) & \cdots & \partial_1 \partial_p f(x) \\ \partial_2 \partial_1 f(x) & \partial_2 \partial_2 f(x) & \cdots & \partial_2 \partial_p f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_p \partial_1 f(x) & \partial_p \partial_2 f(x) & \cdots & \partial_p \partial_p f(x) \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ \vdots \\ h^p \end{bmatrix} \\ &= h^T H(f)_x h. \end{aligned}$$

Logo,

$$f(x_0 + h) = f(x_0) + df_{x_0}(h) + \frac{1}{2} \mathbf{Hess} f_{x_0}(h) + o(|h|^2).$$

2.4 Convexidade e pontos críticos

Definição 2.44. Seja $C \subset \mathbb{R}^p$ um conjunto convexo e considere a função $f : C \rightarrow \mathbb{R}$. Dizemos que f é convexa se

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

para $x, y \in C$ e $t \in [0, 1]$.

Se para quaisquer $x, y \in C$ distintos temos

$$f((1-t)x + ty) < (1-t)f(x) + tf(y)$$

para $0 < t < 1$, então f é estritamente convexa. A função f é (estritamente) côncava se $-f$ é (estritamente) convexa.

Teorema 2.45. Sejam $U \subset \mathbb{R}^p$ aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função de classe C^2 . São equivalentes:

1. f é convexa;
2. Para $x, x+h \in U$ temos

$$f(x+h) \geq f(x) + df_x(h);$$

3. $\text{Hess } f_x(h) \geq 0$ para $x \in U$ e $h \in \mathbb{R}^p$.

Observação: A equivalência entre (1) e (2) não precisa que f seja C^2 , basta C^1 .

Demonstração. Considere

$$\varphi(t) = f(x+th),$$

$$\dot{\varphi}(t) = df_{x+th}(h),$$

$$\ddot{\varphi}(t) = \text{Hess } f_{x+th}(h).$$

(3) \implies (2): Se $\ddot{\varphi}(t) \geq 0$, então $\dot{\varphi}(t)$ é não decrescente.

$$\varphi(1) - \varphi(0) = \int_0^1 \dot{\varphi}(t) dt \geq \dot{\varphi}(0),$$

$$f(x+h) - f(x) \geq df_x(h).$$

(2) \implies (3): Temos que

$$\varphi(t) - \varphi(0) \geq t\dot{\varphi}(0)$$

porque

$$\varphi(t) - \varphi(0) = f(x+th) - f(x) \geq df_x(th) = t\dot{\varphi}(0),$$

Para cada $n \in \mathbb{Z}_{\geq 1}$ existe $\theta_n \in (0, 1)$ tal que

$$\varphi\left(\frac{1}{n}\right) - \varphi(0) = \frac{1}{n}\dot{\varphi}\left(\frac{\theta_n}{n}\right).$$

Assim, temos

$$\frac{1}{n}\dot{\varphi}\left(\frac{\theta_n}{n}\right) = \varphi\left(\frac{1}{n}\right) - \varphi(0) \geq \frac{1}{n}\dot{\varphi}(0),$$

$$\text{Hess } f_x(h) = \ddot{\varphi}(0) = \lim_{n \rightarrow \infty} \frac{\dot{\varphi}\left(\frac{\theta_n}{n}\right) - \dot{\varphi}(0)}{\frac{\theta_n}{n}} \geq 0.$$

(1) \implies (2): Note que para $t \in (0, 1]$

$$f((1-t)x + t(x+h)) \leq (1-t)f(x) + tf(x+h),$$

$$\begin{aligned} f(x + th) &\leq (1-t)f(x) + tf(x+h), \\ f(x + th) - f(x) &\leq t(f(x+h) - f(x)), \\ f(x+h) &\geq f(x) + \frac{f(x+th) - f(x)}{t}, \end{aligned}$$

e, tomando o limite $t \rightarrow 0^+$, temos

$$f(x+h) \geq f(x) + df_x(h).$$

(2) \implies (1): Considere $J := df_{(1-t)x+ty}$.

$$\begin{aligned} f(y) &\geq f((1-t)x+ty) + J((1-t)(y-x)) = f((1-t)x+ty) + (1-t)J(y-x), \\ f(x) &\geq f((1-t)x+ty) + J(-t(y-x)) = f((1-t)x+ty) - tJ(y-x). \end{aligned}$$

Logo, somando as duas expressões, obtemos

$$tf(y) + (1-t)f(x) \geq f((1-t)x+ty).$$

■

Teorema 2.46. Sejam $U \subset \mathbb{R}^p$ aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função de classe C^1 . São equivalentes:

1. f é estritamente convexa;
2. Para $x, x+h \in U$, $h \neq 0$, temos

$$f(x+h) > f(x) + df_x(h).$$

Demonstração. Temos apenas de provar as desigualdades estritas para o Teorema 2.45.

(1) \implies (2): Considere $x, x+h \in U$ com $h \neq 0$.

$$\begin{aligned} f\left(x + \frac{h}{2^n}\right) &= f\left(\frac{1}{2}x + \frac{1}{2}\left(x + \frac{h}{2^{n-1}}\right)\right) < \frac{1}{2}f(x) + \frac{1}{2}f\left(x + \frac{h}{2^{n-1}}\right), \\ f\left(x + \frac{h}{2^n}\right) - f(x) &< \frac{1}{2}\left(f\left(x + \frac{h}{2^{n-1}}\right) - f(x)\right), \\ \frac{f\left(x + \frac{h}{2^n}\right) - f(x)}{\frac{1}{2^n}} &< \frac{f\left(x + \frac{h}{2^{n-1}}\right) - f(x)}{\frac{1}{2^{n-1}}}. \end{aligned}$$

Recursivamente, para $m < n$

$$\frac{f\left(x + \frac{h}{2^n}\right) - f(x)}{\frac{1}{2^n}} < \frac{f\left(x + \frac{h}{2^m}\right) - f(x)}{\frac{1}{2^m}}.$$

Portanto, para $n > 1$

$$\frac{f\left(x + \frac{h}{2^n}\right) - f(x)}{\frac{1}{2^n}} < \frac{f\left(x + \frac{h}{2}\right) - f(x)}{\frac{1}{2}} < f(x+h) - f(x).$$

Fazendo $n \rightarrow \infty$, obtemos

$$df_x(h) \leq \frac{f\left(x + \frac{h}{2}\right) - f(x)}{\frac{1}{2}} < f(x+h) - f(x).$$

(2) \implies (1): Para $t \in (0, 1)$ e $x, y \in U$ pontos distintos considere $J := df_{(1-t)x+ty}$.

$$\begin{aligned} f(y) &> f((1-t)x+ty) + J((1-t)(y-x)) = f((1-t)x+ty) + (1-t)J(y-x), \\ f(x) &> f((1-t)x+ty) + J(-t(y-x)) = f((1-t)x+ty) - tJ(y-x). \end{aligned}$$

Logo, somando as duas expressões, obtemos

$$tf(y) + (1-t)f(x) > f((1-t)x+ty).$$

■

Teorema 2.47. Sejam $U \subset \mathbb{R}^p$ aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função de classe C^2 . Se

$$\mathbf{Hess}f_x(h) > 0$$

para todo $x \in U$ e $h \in \mathbb{R}^d \setminus \{0\}$, então f é estritamente convexa.

Demonstração. Tome $x, x + h \in U$ com $h \neq 0$. Pela fórmula de Taylor,

$$f(x + h) = f(x) + df_x(h) + \int_0^1 (1-t)\mathbf{Hess}f_{x+th}(h)dt > f(x) + df_x(h).$$

■

Observação 2.48. A recíproca do teorema acima não vale. Considere $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = x^4$. Temos que $f''(0) = 0$ e como $f''(x) = 12x^2 > 0$ para $x \neq 0$, então para $h \neq 0$

$$f(x + h) - f(x) = f(x) + f'(x)h + \int_0^1 (1-t)f''(x + th)dt > f(x) + f'(x)h.$$

Logo, o Hessiano se anula em $x = 0$, mas a função é estritamente convexa.

Definição 2.49. Sejam $U \subset \mathbb{R}^p$ aberto e f de classe C^1 . O ponto crítico de uma função $f : U \rightarrow \mathbb{R}$ é um ponto x_0 tal que $df_{x_0} = 0$, ou seja, $\nabla f(x_0) = 0$.

Teorema 2.50. Sejam $U \subset \mathbb{R}^p$ aberto e f de classe C^1 . Se x_0 é ponto de máximo ou mínimo local de f , então $\nabla f(x_0) = 0$.

Demonstração. Para todo $v \in \mathbb{R}^p$ temos

$$\frac{d}{dt} \Big|_{t=0} f(x_0 + tv) = df_{x_0}(v) = 0.$$

Portanto, $\nabla f(x_0) = 0$.

■

Teorema 2.51. Seja $f : U \rightarrow \mathbb{R}$ de classe C^1 , com $U \subset \mathbb{R}^p$ aberto convexo. Se $x_0 \in U$, $\nabla f(x_0) = 0$ e f é estritamente convexa, então $f(x_0) < f(x)$ para todo $x \neq x_0$.

Observação: Se $H(f)_x$ é matriz positiva definida para todo $x \in U$, temos que f é estritamente convexa.

Demonstração. Basta notar que para $x \neq x_0$

$$f(x) > f(x_0) + df_{x_0}(x - x_0) = f(x_0).$$

■

Exemplo 2.52. Considere $f(x, y) = 3x^2 + 7xy + 5y^2$. Temos

$$\nabla f(x, y) = (6x + 7y, 7x + 10y).$$

Fazendo $\nabla f(x, y) = (0, 0)$ obtemos $(x, y) = (0, 0)$.

Como a matriz Hessiana $H(f)_{(x,y)} = \begin{bmatrix} 6 & 7 \\ 7 & 10 \end{bmatrix}$ tem determinante e traço positivos, temos que $H(f)$ é positiva definida.

Logo, $(0, 0)$ é mínimo local.

Exemplo 2.53. Considere $f(x, y, z) = 3x^2 + xz + 2yz - z^2$. Temos

$$\nabla f(x, y, z) = (6x + z, 2z, x + 2y - 2z).$$

Fazendo $\nabla f(x, y, z) = (0, 0, 0)$ obtemos $(x, y, z) = (0, 0, 0)$.

Como a matriz Hessiana

$$H(f)_{(x,y,z)} = \begin{bmatrix} 6 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & -2 \end{bmatrix}$$

tem autovalores $6.13\dots$, $-3.31\dots$, $1.18\dots$, ou seja, $(0, 0, 0)$ é ponto de cela.

Exemplo 2.54. Considere $f(x, y) = xy \log(x^2 + y^2)$ definida para $(x, y) \neq (0, 0)$.

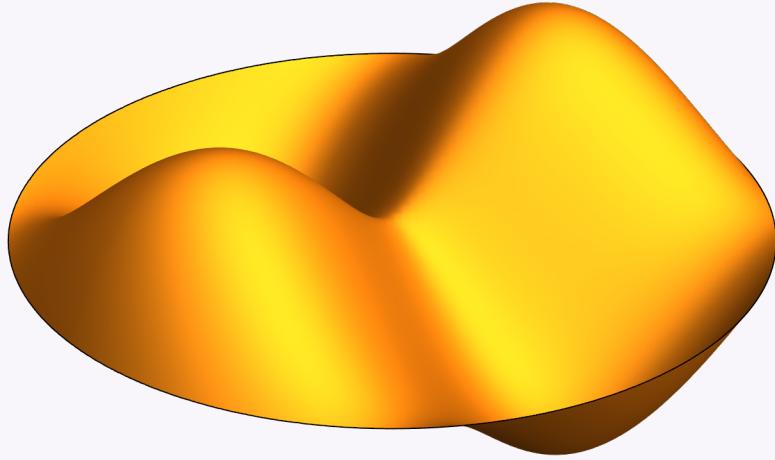


Figura 2.3: Gráfico de $z = xy \log(x^2 + y^2)$.

Temos que

$$\nabla f(x, y) = \left(y \log(x^2 + y^2) + \frac{2x^2y}{x^2 + y^2}, x \log(x^2 + y^2) + \frac{2xy^2}{x^2 + y^2} \right) = (0, 0),$$

ocorre para

$$(\pm 1, 0), (0, \pm 1), \left(\pm \frac{1}{\sqrt{2e}}, \pm \frac{1}{\sqrt{2e}} \right), \left(\pm \frac{1}{\sqrt{2e}}, \mp \frac{1}{\sqrt{2e}} \right).$$

The Hessian of f is

$$\begin{pmatrix} \frac{2xy(x^2+3y^2)}{(x^2+y^2)^2} & \log(x^2+y^2) + \frac{2(x^4+y^4)}{(x^2+y^2)^2} \\ \log(x^2+y^2) + \frac{2(x^4+y^4)}{(x^2+y^2)^2} & \frac{2xy(3x^2+y^2)}{(x^2+y^2)^2} \end{pmatrix}$$

Para os pontos $(\pm 1, 0)$ e $(0, \pm 1)$ temos que o Hessiano é

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$$

e tais pontos não são de máximo ou mínimo local (ponto de sela).

Para os pontos $\left(\pm \frac{1}{\sqrt{2e}}, \pm \frac{1}{\sqrt{2e}} \right)$, temos o Hessiano

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

que é uma matriz positiva (porque tem determinante e traço positivos). Logo, os pontos $\left(\pm \frac{1}{\sqrt{2e}}, \pm \frac{1}{\sqrt{2e}} \right)$ são de mínima.

Similarmente, para $\left(\pm \frac{1}{\sqrt{2e}}, \mp \frac{1}{\sqrt{2e}} \right)$ temos a matriz Hessiana negativa

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Portanto, $\left(\pm \frac{1}{\sqrt{2e}}, \mp \frac{1}{\sqrt{2e}} \right)$ são pontos de máximo.

Exemplo 2.55. (Regressão Linear) Considere o conjunto de dados

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n).$$

Considere a reta $y = ax + b$. Queremos encontrar (a, b) que minimiza

$$E(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

Note que

$$\nabla E(a, b) = \left(\sum_{i=1}^n -x_i(y_i - ax_i - b), \sum_{i=1}^n -2(y_i - ax_i - b) \right) = 0,$$

implica em

$$\begin{cases} \left(\sum_{i=1}^n x_i^2 \right) a + \left(\sum_{i=1}^n x_i \right) b = \sum_{i=1}^n x_i y_i \\ \left(\sum_{i=1}^n x_i \right) a + nb = \sum_{i=1}^n y_i \end{cases}$$

Definindo as médias

$$\bar{x} := \frac{\sum_{i=1}^n x_i}{n} \quad \text{e} \quad \bar{y} := \frac{\sum_{i=1}^n y_i}{n}$$

obtemos

$$\bar{y} = a\bar{x} + b$$

da segunda equação. Resolvendo o sistema, obtemos

$$\begin{aligned} a &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \\ b &= a\bar{x} - \bar{y}. \end{aligned}$$

A matriz Hessiana de E em (a, b) é

$$H(E)_{(a,b)} = \begin{pmatrix} 2\sum_{i=1}^n x_i^2 & 2\sum_{i=1}^n x_i \\ 2\sum_{i=1}^n x_i & 2n \end{pmatrix}.$$

e

$$\det H(E)_{(a,b)} = 4n \sum_{i=1}^n (x_i - \bar{x})^2 > 0 \quad \text{e} \quad \operatorname{tr} H(E)_{(a,b)} > 0.$$

Portanto, os autovalores de $H(E)_{(a,b)}$ são positivos e o ponto crítico é de mínimo.

Definição 2.56. Seja $C \subset \mathbb{R}^p$ um conjunto convexo. Uma combinação convexa de $x_1, x_2, \dots, x_n \in C$ é um vetor da forma

$$\sum_{i=1}^n \lambda_i x_i$$

com $\lambda_1, \dots, \lambda_n \geq 0$ e $\sum_{i=1}^n \lambda_i = 1$.

Teorema 2.57. Seja $C \subset \mathbb{R}^p$ é convexo. Toda combinação convexa de $x_1, \dots, x_n \in C$ pertence a C .

Demonstração. Para $n = 1$ e $n = 2$ o resultado é verdadeiro. Suponha que o resultado vale para n vetores e mostremos que vale para $n + 1$ vetores. Considere

$$\sum_{i=1}^{n+1} \lambda_i x_i$$

com $x_1, \dots, x_{n+1} \in C$ com $\lambda_i \geq 0$ e $\sum_{i=1}^{n+1} \lambda_i = 1$.

Se algum $\lambda_i = 0$, então temos que a soma acima é combinação convexa de n elementos e portanto pertence a C . Podemos supor que $\lambda_i > 0$ para todo i .

De fato,

$$\sum_{i=1}^{n+1} \lambda_i x_i = \sum_{j=1}^n \mu_j y_j \in C,$$

onde

$$y_1 := x_1, \quad y_2 := x_2, \quad \dots, \quad y_{n-1} := x_{n-1}, \quad y_n := \frac{\lambda_n}{\lambda_n + \lambda_{n+1}} x_n + \frac{\lambda_{n+1}}{\lambda_n + \lambda_{n+1}} x_{n+1}$$

são vetores de C e

$$\mu_1 := \lambda_1, \quad \mu_2 := \lambda_2, \quad \dots, \quad \mu_{n-1} := \lambda_{n-1}, \quad \mu_n = \lambda_n + \lambda_{n+1}.$$

■

Exemplo 2.58. Considere $x_1, \dots, x_n \in \mathbb{R}^p$. A envoltória convexa de x_1, \dots, x_n é o conjunto

$$C = \left\{ \sum_{i=1}^n \lambda_i x_i \in \mathbb{R}^p : \lambda_i \geq 0 \text{ e } \sum_{i=1}^n \lambda_i = 1 \right\}.$$

O conjunto C é convexo. De fato, para $t \in [0, 1]$ temos

$$t \sum_{i=1}^n \lambda_i x_i + (1-t) \sum_{i=1}^n \lambda'_i x_i = \sum_{i=1}^n (t\lambda_i + (1-t)\lambda'_i) x_i \in C$$

porque

$$\sum_{i=1}^n t\lambda_i + (1-t)\lambda'_i = t \sum_{i=1}^n \lambda_i + (1-t) \sum_{i=1}^n \lambda'_i = t + (1-t) = 1.$$

Teorema 2.59. (Desigualdade de Jensen) Se $f : C \rightarrow \mathbb{R}$ é uma função convexa e $\sum_{i=1}^n \lambda_i x_i$ é uma combinação convexa de elementos de C , então

$$f \left(\sum_{i=1}^n \lambda_i x_i \right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Demonstração. O caso $n = 2$ segue da definição de convexidade. Suponha que o resultado vale para $n \geq 2$ e mostremos que vale para $n + 1$. Considere uma combinação convexa

$$\sum_{i=1}^{n+1} \lambda_i x_i \in C,$$

onde $x_1, \dots, x_n, x_{n+1} \in C$.

Se $\lambda_i = 0$ para algum i , então podemos excluir um dos termos da combinação e aplicar a hipótese de indução. Podemos portanto supor que $\lambda_i > 0$ para todo i . Temos

$$\begin{aligned} f \left(\sum_{i=1}^{n+1} \lambda_i x_i \right) &= f \left(\sum_{i=1}^{n-1} \lambda_i x_i + (\lambda_n + \lambda_{n+1}) \left(\frac{\lambda_n}{\lambda_n + \lambda_{n+1}} x_n + \frac{\lambda_{n+1}}{\lambda_n + \lambda_{n+1}} x_{n+1} \right) \right) \\ &\leq \sum_{i=1}^{n-1} \lambda_i f(x_i) + (\lambda_n + \lambda_{n+1}) f \left(\frac{\lambda_n}{\lambda_n + \lambda_{n+1}} x_n + \frac{\lambda_{n+1}}{\lambda_n + \lambda_{n+1}} x_{n+1} \right) \\ &\leq \sum_{i=1}^{n-1} \lambda_i f(x_i) + (\lambda_n + \lambda_{n+1}) \left(\frac{\lambda_n}{\lambda_n + \lambda_{n+1}} f(x_n) + \frac{\lambda_{n+1}}{\lambda_n + \lambda_{n+1}} f(x_{n+1}) \right) \\ &= \sum_{i=1}^{n+1} \lambda_i f(x_i). \end{aligned}$$

■

Exemplo 2.60. Seja $f : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ dada por $f(x) = -\log(x)$. Temos $f''(x) = \frac{1}{x^2} > 0$. A função f é convexa por causa do Teorema 2.47. Assim, se $x_1, \dots, x_n > 0$ temos a combinação convexa

$$\frac{1}{n}x_1 + \dots + \frac{1}{n}x_n \in \mathbb{R}_{>0}.$$

Pela desigualdade de Jensen,

$$\begin{aligned} f\left(\frac{1}{n}x_1 + \dots + \frac{1}{n}x_n\right) &\leq \frac{1}{n}f(x_1) + \dots + \frac{1}{n}f(x_n) \\ \log\left(\frac{1}{n}x_1 + \dots + \frac{1}{n}x_n\right) &\geq \frac{1}{n}\log(x_1) + \dots + \frac{1}{n}\log(x_n), \\ \log\left(\frac{1}{n}x_1 + \dots + \frac{1}{n}x_n\right) &\geq \log(\sqrt[n]{x_1 \cdots x_n}), \\ \frac{x_1 + \dots + x_n}{n} &\geq \sqrt[n]{x_1 \cdots x_n} \quad (\text{Desigualdade das médias}). \end{aligned}$$

Teorema 2.61. Sejam $U \subset \mathbb{R}^p$ um aberto convexo e $f : U \rightarrow \mathbb{R}$ convexa. A função f é localmente Lipchitz e, consequentemente, contínua.

Demonstração. Mostremos que f é contínua em $x_0 \in U$. Para $r > 0$ suficientemente pequeno o conjunto convexo

$$B = \left\{ \sum_{i=1}^p \lambda_i (x_0 + re_i) + \mu_i (x_0 - re_i) \in \mathbb{R}^p : \sum_{i=1}^p \lambda_i + \mu_i = 1 \text{ e } \lambda_i, \mu_i \geq 0 \right\}$$

está contido em U . Além disso, x_0 está no interior de V . De fato, B é a bola fechada de raio r com respeito a L^1 , isto é,

$$B = \left\{ x \in \mathbb{R}^p : \sum_{i=1}^p |x^i - x_0^i| \leq r \right\}.$$

Considere

$$x = \sum_{i=1}^p \lambda_i (x_0 + re_i) + \mu_i (x_0 - re_i)$$

em B .

Mostremos que $f : B \rightarrow \mathbb{R}$ é limitada. Note que

$$f(x) \leq \sum_{i=1}^p \lambda_i f(x_0 + re_i) + \mu_i f(x_0 - re_i) \leq c,$$

onde $c := \max_{1 \leq i, j \leq p} \{|f(x_0 + re_i)|, |f(x_0 - re_j)|\}$.

Logo, $f : B \rightarrow \mathbb{R}$ é limitada superiormente. Por outro lado,

$$2x_0 - x = \sum_{i=1}^p \lambda_i (x_0 - re_i) + \mu_i (x_0 + re_i) \in B.$$

Temos

$$\begin{aligned} f(x_0) &= f\left(\frac{1}{2}x + \frac{1}{2}(2x_0 - x)\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(2x_0 - x), \\ f(x) &\geq 2f(x_0) - f(2x_0 - x) \geq 2f(x_0) - c. \end{aligned}$$

Portanto, f é limitada em uma vizinhança de x_0 . Considere uma bola $B(x_0, \varepsilon) \subset U$ onde f é limitada. Existe $M > 0$ tal que $|f(x)| \leq M$ para $x \in B(x_0, \varepsilon)$.

Provemos que f é Lipschitz em uma vizinhança de x_0 . Suponha que f não é Lipschitz para nenhuma vizinhança de x_0 . Existe $x_1, x_2 \in B(x_0, \varepsilon/4)$ tal que

$$|f(x_2) - f(x_1)| > \frac{4M}{\varepsilon} |x_1 - x_2|,$$

trocando x_1 por x_2 se necessário, podemos assumir $f(x_2) - f(x_1) > 0$ e, consequentemente,

$$f(x_2) - f(x_1) > \frac{4M}{\varepsilon} |x_1 - x_2|.$$

Tome x_3 em $B(x_0, \varepsilon)$ tal que x_3 está na reta dada por x_1, x_2 e x_2 está estritamente entre x_1, x_3 . Temos $x_2 = (1-t)x_1 + tx_3$ com $t \in (0, 1)$.

$$f(x_2) \leq (1-t)f(x_1) + tf(x_3),$$

$$f(x_2) - f(x_1) \leq t(f(x_3) - f(x_1)),$$

$$f(x_3) - f(x_1) > \frac{4M}{t\varepsilon} |x_2 - x_1|.$$

Podemos tomar $x_3 \in B(x_0, \varepsilon)$ de modo que $t = \frac{2|x_2-x_1|}{\varepsilon}$, que é menor que 1, porque para

$$x_3 := x_1 + \frac{(x_2 - x_1)}{t}$$

temos

$$|x_3 - x_1| = \frac{|x_2 - x_1|}{t} = \frac{\varepsilon}{2},$$

$$|x_3 - x_0| \leq |x_3 - x_1| + |x_1 - x_0| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} < \varepsilon.$$

ou seja,

$$f(x_3) - f(x_1) > 2M,$$

o que é uma contradição porque $|f(x_3) - f(x_1)| \leq |f(x_3)| + |f(x_1)| \leq 2M$. ■

Observação 2.62. Teorema de Rademacher: Funções Lipschitz são diferenciáveis em quase todo ponto (no sentido de teoria de medida). Como funções convexas são localmente Lipschitz, ela é diferenciável em quase todo ponto.

Capítulo 3

Variedades em \mathbb{R}^p

3.1 Teorema da função inversa

Comecemos provando o teorema do ponto fixo de Banach. É um dos teoremas básicos quando se procura por pontos fixos. Usaremos ele para provar o teorema da função inversa. O teorema de ponto fixo de Banach aparece com frequência em demonstrações de resultados de análise numérica e na prova do teorema de existência e unicidade para equações diferenciais ordinárias.

Teorema 3.1. (Teorema de ponto fixo de Banach) Seja X espaço métrico completo e k uma constante real tal que $0 \leq k < 1$. Se $f : X \rightarrow X$ satisfaz

$$d(f(x_1), f(x_2)) \leq kd(x_1, x_2),$$

então f possui um único ponto fixo x^* e para todo $x \in X$ temos que $\lim_{n \rightarrow \infty} f^n(x) = x^*$.

Demonstração. Como f é contínua porque é Lipschitz. Considere $x \in X$ e defina $x_n = f^n(x)$. Para n inteiro não negativo, temos

$$d(f^{n+1}(x), f^n(x)) \leq k^n d(f(x), x),$$

$$d(x_{n+1}, x_n) \leq k^n d(x_1, x).$$

Para $n > m$, temos, pela desigualdade triangular,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_{m+2}, x_{m+1}) + d(x_{m+1}, x_m) \\ &\leq (k^{n-1} + k^{n-2} + \cdots + k^{m+1} + k^m)d(x_1, x). \end{aligned}$$

Como

$$k^{n-1} + k^{n-2} + \cdots + k^{m+1} + k^m \leq k^m \sum_{i=0}^{\infty} k^i = \frac{k^m}{1-k}$$

temos que x_n é sequência de Cauchy. Defina

$$x^* := \lim_{n \rightarrow \infty} x_n.$$

Como f é contínua,

$$f(x^*) = f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} x_{n+1} = x^*.$$

Portanto, x^* é ponto fixo. Se $x' \in X$ é ponto fixo f , temos

$$d(x^*, x') = d(f(x^*), f(x')) \leq kd(x^*, x'),$$

$$(1 - k)d(x^*, x') \leq 0,$$

$$d(x^*, x') = 0,$$

$$x^* = x'.$$

■

Assim como homeomorfismo é uma função entre espaços que compara suas propriedades topológicas, introduzimos agora difeomorfismo, que faz o mesmo com respeito à diferenciabilidade.

Definição 3.2. Um difeomorfismo de classe C^k é uma função $f : U \rightarrow V$, onde $U, V \subset \mathbb{R}^p$ são abertos, tal que f é bijetora e f, f^{-1} são de classe C^k . Quando dizemos que f é difeomorfismo sem mencionar sua classe de diferenciabilidade estamos assumindo que f é C^∞ .

Observação 3.3. Se f é difeomorfismo com inversa g , então df_x é invertível e pela regra da cadeia

$$dg_y = [df_{g(y)}]^{-1}.$$

Exemplo 3.4. A função $\cos : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow (-1, 1)$ é difeomorfismo com inversa $\arccos : (-1, 1) \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$,

Exemplo 3.5. (Coordenadas polares) Considere \cos como definido no exemplo anterior. A função

$$f : \mathbb{R}_{>0} \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \{(x, y) \in \mathbb{R}^2 : (x, y) \notin \{0\} \times \mathbb{R}_{\leq 0}\}$$

dada por

$$f(r, \theta) = (r \cos(\theta), r \sin(\theta))$$

é difeomorfismo com inversa

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

Teorema 3.6. (Teorema da função inversa) Seja $f : U \rightarrow \mathbb{R}^p$, onde $U \subset \mathbb{R}^p$ é aberto, de classe C^k , $k \geq 1$. Tome $x_0 \in U$ tal que df_{x_0} é invertível. Existe uma vizinhança $V \subset U$ de x_0 tal que

- $f(V)$ é aberto;
- $f : V \rightarrow f(V)$ é difeomorfismo de classe C^k , isto é, f é bijetora e $f^{-1} : f(V) \rightarrow V$ tem classe C^k .

Demonstração. Sem perda de generalidade, podemos supor que $x_0 = f(x_0) = 0$ e $df_{x_0} = -\text{Id}_{\mathbb{R}^p}$, pois podemos trocar f por $x \rightarrow -df_{x_0}^{-1}(f(x+x_0) - f(x_0))$.

Como $df_0 = -\text{Id}_{\mathbb{R}^p}$, existe $r > 0$ tal que

$$\|\text{Id}_{\mathbb{R}^p} + df_x\| < \frac{1}{2} \quad \text{e} \quad \det(df_x) \neq 0$$

para $x \in B(0, r)$, onde a norma $\|\cdot\|$ usada acima é a norma de operador. Pelo teorema fundamental do cálculo, temos que para $|x| < r$

$$f(x) - f(0) = \int_0^1 \frac{d}{dt}(f(tx))dt = \int_0^1 df_{tx}(x)dt,$$

$$f(x) = \int_0^1 df_{tx}(x)dt,$$

$$f(x) + x = \int_0^1 (df_{tx} + \text{Id}_{\mathbb{R}^p})(x)dt,$$

$$|f(x) + x| < \frac{1}{2}|x|.$$

Considere a bola fechada $K = \overline{B(0, r/2)}$. Para cada $s \in B(0, r/4)$ defina $\lambda_s(x) = f(x) + x - s$.

Afirmção: $\lambda_s(K) \subset K$. De fato, se $x \in K$ temos

$$|\lambda_s(x)| \leq |f(x) + x| + |s| < \frac{1}{2}|x| + |s| \leq \frac{r}{4} + \frac{r}{4} = \frac{r}{2},$$

$$\lambda_s(x) \in K.$$

Portanto, $\lambda_s : K \rightarrow K$ está bem definida.

Afirmção: $|\lambda_s(x_1) - \lambda_s(x_2)| < \frac{1}{2}|x_1 - x_2|$. De fato, para $x_1, x_2 \in K$, temos

$$\lambda_s(x_1) - \lambda_s(x_2) = f(x_1) - f(x_2) + x_1 - x_2 = \int_0^1 (df_{(1-t)x_1+tx_2} + \text{Id}_{\mathbb{R}^p})(x_1 - x_2) dt,$$

$$|\lambda_s(x_1) - \lambda_s(x_2)| < \frac{1}{2}|x_1 - x_2|.$$

Logo, pelo teorema do ponto fixo de Banach (K é compacto e portanto completo), existe x_s tal que $\lambda_s(x_s) = x_s$, isto é, $f(x_s) = s$. Logo, $B(0, r/4) \subset f(K)$.

Considere o aberto $V = \{x \in B(0, r) : f(x) \in B(0, r/4)\}$. Temos que $f(V) = B(0, r/4)$, isto é, a função $f : V \rightarrow B(0, r/4)$ é sobrejetora.

Afirmção: $f : V \rightarrow B(0, r/4)$ é injetora. Tome $x_1, x_2 \in V$.

$$\begin{aligned} |x_1 - x_2| &= |(f(x_1) + x_1) - (f(x_2) + x_2) + (f(x_2) - f(x_1))| \\ &\leq |(f(x_1) + x_1) - (f(x_2) + x_2)| + |f(x_2) - f(x_1)| \\ &= \left| \int_0^1 \frac{d}{dt} \left(f((1-t)x_1 + tx_2) + (1-t)x_1 + tx_2 \right) dt \right| + |f(x_2) - f(x_1)| \\ &= \left| \int_0^1 (df_{(1-t)x_1+tx_2} + \text{Id}_{\mathbb{R}^p})(x_1 - x_2) dt \right| + |f(x_2) - f(x_1)| \\ &\leq \frac{1}{2}|x_1 - x_2| + |f(x_2) - f(x_1)|. \end{aligned}$$

Logo,

$$|x_1 - x_2| \leq 2|f(x_1) - f(x_2)|.$$

Portanto, se $f(x_1) = f(x_2)$, temos $x_1 = x_2$.

Logo, $f : V \rightarrow B(0, r/4)$ é bijetora. Seja g sua inversa. Como

$$|g(y_1) - g(y_2)| \leq 2|f(g(y_1)) - f(g(y_2))| = 2|y_1 - y_2|,$$

temos que g é Lipschitz.

Logo, $f : V \rightarrow B(0, r/4)$ é homeomorfismo.

Mostremos que g é diferenciável. Tome $y' \in B(0, r/4)$, $x' = g(y')$, $x = g(y)$ e $A = df_x^{-1}$. Temos

$$g(y') - g(y) - A(y' - y) = x' - x - A(f(x') - f(x)) = x' - x - A(df_x(x' - x)) + o(|x' - x|) = o(|x' - x|)$$

quando $x' \rightarrow x$.

Como $|x' - x| \leq 2|y' - y|$, temos

$$\frac{|g(y') - g(y) - A(y' - y)|}{|y' - y|} = \frac{|g(y') - g(y) - A(y' - y)|}{|x' - x|} \frac{|x' - x|}{|y' - y|} \leq 2 \frac{|g(y') - g(y) - A(y' - y)|}{|x' - x|} \rightarrow 0$$

quando $y' \rightarrow y$.

Portanto,

$$g(y') = g(y) + A(y' - y) + o(|y' - y|).$$

Logo, g é diferenciável e $dg_y = df_{g(y)}^{-1}$ para todo y . Dessa relação, temos que se g é C^i então g é C^{i+1} para $0 \leq i < k$. Logo, g é de classe C^k . ■

Exemplo 3.7. O teorema da função inversa na reta se enuncia da seguinte forma: se I é intervalo aberto e $f : I \rightarrow \mathbb{R}$ é uma função de classe C^k com derivada que não se anula em nenhum ponto, então $J = f(I)$ é um intervalo aberto e $f : I \rightarrow J$ é difeomorfismo. De fato, basta notar que o fato da derivada não se anular implica que a função é estritamente monótona e homeomorfismo com a imagem J aberta. Aplicando o teorema da função inversa mostramos que tal mapa é localmente difeomorfismo e portanto globalmente difeomorfismo por ser homeomorfismo.

Exemplo 3.8. Considere $f : \mathbb{C} \rightarrow \mathbb{C}$ dado por $f(z) = z^n$, $n \geq 2$. Pelo binômio de Newton, temos

$$f(z+h) = (z+h)^n = z^n + nz^{n-1}h + o(|h|),$$

ou seja, $df_z(h) = nz^{n-1}h$. Para $z \neq 0$ temos que df_z é injetor, e portanto, bijetor. Assim, pelo teorema da função inversa, existe uma vizinhança U de z tal que $f(U)$ é aberto e $f : U \rightarrow f(U)$ é difeomorfismo.

Assim, f é localmente difeomorfismo. No entanto, f não é globalmente difeomorfismo porque f não é injetora. De fato, para $z \neq 0$

$$f(z) = f\left(e^{\frac{2\pi i}{n}}z\right).$$

3.2 Variedades mergulhadas em \mathbb{R}^p e espaço tangente

Definição 3.9. Considere os inteiros positivos p e n com $n \leq p$.

Um subconjunto $M \subset \mathbb{R}^p$ é uma n -variedade se para cada $x_0 \in M$ existe um difeomorfismo $\varphi : U \rightarrow V$ onde U é vizinhança de x_0 e V é uma vizinhança de $y_0 := \varphi(x_0)$ tal que

$$\varphi(U \cap M) = \{y \in \mathbb{R}^p : y^{n+1} = y_0^{n+1}, \dots, y^p = y_0^p\} \cap V.$$

Chamamos o difeomorfismo φ de carta.

Observação 3.10. Mais precisamente, o que chamamos acima de n -variedade é conhecido como subvariedade de \mathbb{R}^p . A definição geral de variedade é um pouco mais abstrata, mas não vamos usa-la.

Exemplo 3.11. Seja $f : V \rightarrow W$, onde $U \subset \mathbb{R}^p$ e $W \subset \mathbb{R}^q$ são abertos, uma função suave. Mostremos que o gráfico de f

$$M = \{(x, f(x)) \in \mathbb{R}^p \times \mathbb{R}^q : x \in U\}$$

é uma p -variedade. De fato, para cada $(x_0, f(x_0)) \in M$ temos o o difeomorfismo $\varphi : U \times \mathbb{R}^q \rightarrow U \times \mathbb{R}^q$ dado por $\varphi(x, y) = (x, y - f(x))$, cuja inversa é $\varphi^{-1}(x, v) = (x, v + f(x))$.

Note que $M \cap U \times \mathbb{R}^q = M$ e

$$\varphi(M) = U \times \{0\} = (\mathbb{R}^p \times \{0\}) \cap (U \times \mathbb{R}^q)$$

Exemplo 3.12. Uma p -variedade dentro de \mathbb{R}^p é um aberto.

Proposição 3.13. Seja M uma n -variedade. O espaço tangente $T_x M$ é espaço vetorial de dimensão n para todo $x \in M$.

Demonstração. Tome $x_0 \in M$ e defina $y_0 := \varphi(x_0)$. Existe uma carta $\varphi : U \rightarrow V$ tal que

$$\varphi(M \cap U) = V \cap \mathbb{R}^n \times \{(y_0^{n+1}, \dots, y_0^p)\},$$

onde $y_0 = \varphi(x_0)$.

Provemos que

$$T_{x_0}M = d\varphi_{x_0}^{-1}(\mathbb{R}^n \times \{0\}).$$

Seja $u \in T_{x_0}M$. Existe uma curva $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ tal que $\gamma(0) = x_0$ e $\dot{\gamma}(0) = u$. Temos que $\varphi(\gamma(0)) = y_0$ e $d\varphi_{x_0}(v) \in \mathbb{R}^n \times \{0\}$. Portanto, $T_{x_0}M \subset d\varphi_{x_0}^{-1}(\mathbb{R}^n \times \{0\})$.

Por outro lado, se $u \in \mathbb{R}^n \times \{0\}$, temos $\gamma(t) = y_0 + tu$ em $\mathbb{R}^n \times \{0\}$ e $\varphi^{-1}(\gamma(t))$ é uma curva em M passando por x_0 com velocidade $d(\varphi^{-1})_{y_0}(u) = d\varphi_{x_0}^{-1}(u)$ para $t = 0$.

Logo,

$$d\varphi_{x_0}^{-1}(\mathbb{R}^n \times \{0\}) \subset T_{x_0}M.$$

Portanto, $T_{x_0}M = d\varphi_{x_0}^{-1}(\mathbb{R}^n \times \{0\})$ é espaço vetorial de dimensão n . ■

Exemplo 3.14. No caso do gráfico de acordo com exemplo 3.11, que o espaço tangente $T_{x_0}M$ em x_0 é gerado pelos vetores

$$(e_1, \partial_1 f(x_0)), \quad (e_2, \partial_2 f(x_0)), \quad \dots, \quad (e_p, \partial_p f(x_0)).$$

Exemplo 3.15. (Teorema da função implícita via força bruta) Considere $f : U \rightarrow \mathbb{R}$ suave, onde U é aberto de \mathbb{R}^{p+1} com coordenadas (x, y) , $x \in \mathbb{R}^p$ e $y \in \mathbb{R}$.

Considere $(x_0, y_0) \in f^{-1}(0)$ e suponha que $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Mostremos existe uma vizinhança $V \times W \subset U$ de (x_0, y_0) , com $x_0 \in V \subset \mathbb{R}^p$ e $y_0 \in W \subset \mathbb{R}$, e uma função suave $g : V \rightarrow W$ tais que

$$f^{-1}(0) \cap (V \times W) = \{(x, g(x)) \in \mathbb{R}^p \times \mathbb{R} : x \in V\}.$$

Sem perda de generalidade, suponha que $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Existe $\beta > 0$ tal que

$$\{x_0\} \times [y_0 - \beta, y_0 + \beta] \subset U$$

e

$$\frac{\partial f}{\partial y}(x_0, y) > 0 \quad \text{para } y \in [y_0 - \beta, y_0 + \beta].$$

Temos que $y \rightarrow f(x_0, y)$ é estritamente crescente. Assim,

$$f(x_0, y - \beta) \leq f(x_0, y) \leq f(x_0, y_0 + \beta), \quad y \in [y_0 - \beta, y_0 + \beta]$$

Tomando $y = y_0$, temos

$$f(x_0, y_0 - \beta) < 0 < f(x_0, y_0 + \beta).$$

Existe $\alpha > 0$ suficientemente pequeno garantindo que

$$\overline{B(x_0, \alpha)} \times [y_0 - \beta, y_0 + \beta] \subset U$$

e que para $(x, y) \in \overline{B(x_0, \alpha)} \times [y_0 - \beta, y_0 + \beta]$ valem as desigualdades

$$f(x, y_0 - \beta) < 0, \quad f(x, y_0 + \beta) > 0 \quad \text{e} \quad \frac{\partial f}{\partial y}(x, y) > 0.$$

Em particular, para cada x temos que $f(x, -)$ estritamente crescente.

Portanto,

$$f(x, y_0 - \beta) \leq f(x, y) \leq f(x, y_0 + \beta)$$

quando $(x, y) \in \overline{B(x_0, \alpha)} \times [y_0 - \beta, y_0 + \beta]$.

Assim, para cada $x \in \overline{B(x_0, \alpha)}$ existe $g(x) \in (y_0 - \beta, y_0 + \beta)$ de modo que $f(x, g(x)) = 0$. Além disso, se $f(x, y) = 0$ então $y = g(x)$.

Afirmiação: $g : \overline{B(x_0, \alpha)} \rightarrow [y_0 - \beta, y_0 + \beta]$ é contínua.

Tome $\hat{x} \in \overline{B(x_0, \alpha)}$ e suponha que g é descontínua em \hat{x} . Existem, portanto, $\varepsilon > 0$ e uma sequência $x_n \in \overline{B(x_0, \alpha)}$ tais que $x_n \rightarrow \hat{x}$ e $|g(\hat{x}) - g(x_n)| \geq \varepsilon$.

A menos de passarmos para uma subsequência, podemos supor que $(x_n, g(x_n))$ converge para $(\hat{x}, \hat{y}) \in \overline{B(x_0, \alpha)} \times [y_0 - \beta, y_0 + \beta]$. Assim, $f(x_n, f(x_n)) = 0$ e tomando o limite, obtemos

$f(\hat{x}, \hat{y}) = 0$. Logo, $\hat{y} = g(\hat{x})$ e $g(x_n) \rightarrow g(\hat{x})$ contradizendo que $|g(x_n) - g(\hat{x})| \geq \varepsilon$. Logo, g é contínua.

Afirmacão: $g : B(x_0, \alpha) \rightarrow (y_0 - \beta, y_0 + \beta)$ é suave. Tome $\hat{x} \in B(x_0, \alpha)$, $\hat{x} + te_1 \in B(x_0, \alpha)$ e defina

$$\phi(s) = f(\hat{x} + ste_1, sg(\hat{x} + te_1) + (1-s)g(\hat{x})).$$

Pelo teorema do valor médio, existe $\theta \in (0, 1)$, que depende de t , tal que

$$\phi(1) - \phi(0) = \phi'(\theta).$$

Como

$$\frac{d\phi}{ds}(s) = \frac{\partial f}{\partial x^1}(\hat{x} + ste_1, sg(\hat{x} + te_1)t + (1-s)g(\hat{x}))t + \frac{\partial f}{\partial y}(\hat{x} + ste_1, sg(\hat{x} + te_1) + (1-s)g(\hat{x}))(g(\hat{x} + te_1) - g(\hat{x}))$$

Como $\phi(1) = \phi(0) = 0$, obtemos que $\dot{\phi}(\theta) = 0$ e consequentemente

$$\frac{g(\hat{x} + te_1) - g(\hat{x})}{t} = -\frac{\frac{\partial f}{\partial x^1}(\hat{x} + \theta te_1, \theta g(\hat{x} + te_1)t + (1-\theta)g(\hat{x}))}{\frac{\partial f}{\partial y}(\hat{x} + \theta te_1, \theta g(\hat{x} + te_1)t + (1-\theta)g(\hat{x}))}.$$

Fazendo $t \rightarrow 0$, obtemos

$$\frac{\partial g}{\partial x^1}(\hat{x}) = -\frac{\frac{\partial f}{\partial x^1}(\hat{x}, g(\hat{x}))}{\frac{\partial f}{\partial y}(\hat{x}, g(\hat{x}))}.$$

Em geral,

$$\frac{\partial g}{\partial x^i}(\hat{x}) = -\frac{\frac{\partial f}{\partial x^i}(\hat{x}, g(\hat{x}))}{\frac{\partial f}{\partial y}(\hat{x}, g(\hat{x}))}.$$

e, consequentemente, temos que $\frac{\partial g}{\partial x^i}$ existe e é contínua. Portanto, g é de classe C^1 .

Recursivamente, se f é de classe C^k temos que g é de classe C^k .

Pela construção da função g , temos que definindo $V := B(x_0, \alpha)$ e $W := (y_0 - \beta, y_0 + \beta)$ concluímos que

$$f^{-1}(0) \cap (V \times W) = \{(x, g(x)) \in \mathbb{R}^{p+q} : x \in V\}.$$

Teorema 3.16. (Teorema da função implícita) Seja $f : U \rightarrow \mathbb{R}^q$ suave, onde U é aberto de \mathbb{R}^{p+q} com coordenadas (x, y) , $x \in \mathbb{R}^p$, $y \in \mathbb{R}^q$. Suponha que

$$\det \left(\frac{\partial f^i}{\partial y^j}(x_0, y_0) \right) \neq 0$$

$f(x_0, y_0) = 0$ com $(x_0, y_0) \in \mathbb{R}^p \times \mathbb{R}^q$. Existe uma vizinhança V de x_0 , uma vizinhança W de y_0 , e uma função $g : V \rightarrow W$ suave tais que

- $y_0 = g(x_0)$;
- $f(x, g(x)) = 0$ para $x \in V$;
- se $f(x, y) = 0$ para $(x, y) \in V \times W$, então $y = g(x)$.

Em outras palavras, $f(x, y) = 0$ é gráfico de $y = g(x)$ na vizinhança $V \times W$ de (x_0, y_0) .

Demonstração. Considere

$$h : U \rightarrow \mathbb{R}^{p+q}$$

dada por

$$h(x, y) = (x, f(x, y)).$$

A matriz Jacobiana de f é

$$df_{(x,y)} = \left[\begin{array}{c|c} \text{Id}_{p \times p} & 0_{p \times q} \\ \hline \frac{\partial f^i}{\partial x^j}(x, y) & \frac{\partial f^i}{\partial y^j}(x, y) \end{array} \right]$$

Logo,

$$\det(df_{(x,y)}) = \det\left(\frac{\partial f^i}{\partial y^j}(x,y)\right)$$

e

$$\det(df_{(x_0,y_0)}) \neq 0.$$

Pelo teorema da função inversa, existe uma vizinhança aberta $W' \times W \subset U$ de (x_0, y_0) , com $x_0 \in W'$ e $y_0 \in W$, tal que $h(W' \times W) \subset \mathbb{R}^{p+q}$ é aberto e $h : W' \times W \rightarrow h(W' \times W)$ é difeomorfismo. Tome uma vizinhança $V \times V'$ de $(x_0, 0)$ tal que $V \times V' \subset h(W' \times W)$.

A função $V \ni x \mapsto h^{-1}(x, 0)$ se escreve como $(x, g(x))$ e a função $g : V \rightarrow W$ é suave. De fato, a função g é obtida compondo funções suaves: $g = \pi_W \circ h^{-1} \circ \iota_V$, onde $\pi_W : W' \times W \rightarrow W$ é dada por $(x, z) \mapsto z$ e $\iota_V : V \rightarrow V \times V'$ é dada por $x \mapsto (x, 0)$.

Assim, temos a função suave $g : V \rightarrow W$. Mostremos que

$$f^{-1}(0) \cap (V \times W) = \{(x, g(x)) \in \mathbb{R}^{p+q} : x \in V\}.$$

Tome $(x, y) \in V \times W$ tal que $f(x, y) = 0$. Temos que $h(x, y) = (x, 0)$ e, portanto,

$$h(x, y) = (x, f(x, y)) = (x, 0) = h(x, g(x)),$$

ou seja, $y = g(x)$. Assim,

$$f^{-1}(0) \cap (V \times W) \subset \{(x, g(x)) \in \mathbb{R}^{p+q} : x \in V\}.$$

Por outro lado, se temos $(x, g(x))$ com $x \in V$, então $h(x, g(x)) = (x, 0)$. Como $h(x, y) = (x, f(x, y))$, obtemos $f(x, g(x)) = 0$. Portanto,

$$f^{-1}(0) \cap (V \times W) = \{(x, g(x)) \in \mathbb{R}^{p+q} : x \in V\}.$$

Como $(x_0, y_0) \in f^{-1}(0) \cap V \times W$, obtemos $g(x_0) = y_0$. ■

Observação 3.17. Note que, em particular, o teorema da função implícita nos diz que

$$\{(x, y) \in V \times W : f(x, y) = 0\}$$

é gráfico de uma função suave $g : V \rightarrow W$ e portanto é uma p -variedade.

É comum usar o teorema da função implícita com funções a valores reais $f : U \rightarrow \mathbb{R}$. Se temos $(x^1, \dots, x^p) \mapsto f(x^1, \dots, x^p)$ e $\partial_p f(x_0) \neq 0$ com $f(x_0) = 0$, então podemos $x^p = g(x^1, \dots, x^{p-1})$ e perto de x_0 os pontos dados por $f(x) = 0$ coincidem com o gráfico $\{(x^1, \dots, x^p) : x^p = g(x^1, \dots, x^{p-1})\}$ de g . Similarmente, se $\partial_i f(x_0) \neq 0$, então podemos escrever x^i como função das demais variáveis, $x^i = g(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^p)$, e em uma vizinhança de x_0 temos que os pontos $f(x) = 0$ coincidem com o gráfico $\{(x^1, \dots, x^p) : x^i = g(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^p)\}$.

Exemplo 3.18. Considere $y^2 = x(x^2 - 1)$. Defina $f(x, y) = y^2 - x(x^2 - 1)$ e note que

$$\nabla f(x, y) = (3x^2 - 1, 2y).$$

Para os pontos $(-1, 0), (0, 0), (1, 0)$ temos que $\partial_x f \neq 0$ e $\partial_y f = 0$, ou seja, podemos escrever x como função de y . Para os demais pontos de $f(x, y) = 0$, temos $\partial_y f \neq 0$ e podemos escrever y como função de x :

$$y = \sqrt{x(x^2 - 1)} \quad \text{ou} \quad y = -\sqrt{x(x^2 - 1)}.$$

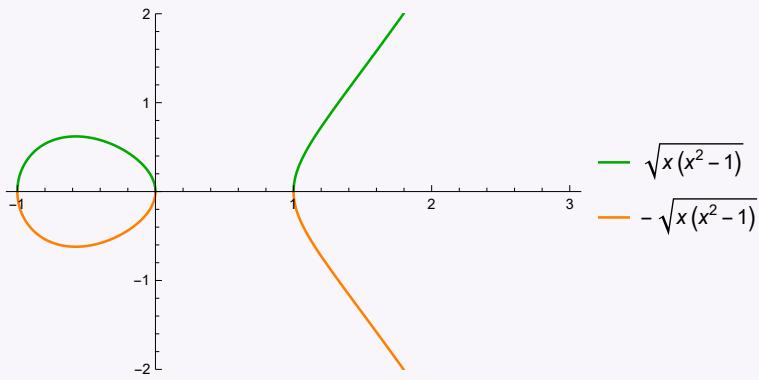


Figura 3.1: Gráfico de $y^2 - x(x^2 - 1) = 0$ com y escrito como funções de x .

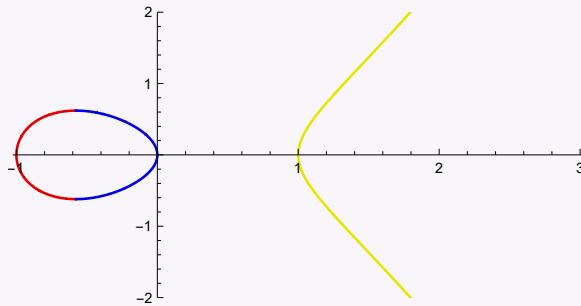


Figura 3.2: Gráfico de $y^2 - x(x^2 - 1) = 0$ com x escrito como funções de y .

Exemplo 3.19. Considere $f(x) = \langle x, x \rangle - 1$, $x \in \mathbb{R}^p$.

$$df_x = [2x^1 \quad 2x^2 \quad \dots \quad 2x^p]$$

Se $x_0 \in f^{-1}(0) = \mathbb{S}^{p-1}$ então uma das colunas de df_{x_0} não se anula. Suponhamos que $2x_0^i \neq 0$. Pelo teorema da função implícita, podemos escrever x^i como função das demais coordenadas e $f^{-1}(0)$ é localmente o gráfico dessa função, isto é, temos uma função $x^i = g(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^p)$ e uma vizinhança U de x_0 tal que $U \cap \mathbb{S}^{p-1}$ é o gráfico de g . No caso da esfera, temos

$$x^i = \sqrt{1 - ((x^1)^2 + \dots + (x^{i-1})^2 + (x^{i+1})^2 + \dots + (x^p)^2)}$$

ou

$$x^i = -\sqrt{1 - ((x^1)^2 + \dots + (x^{i-1})^2 + (x^{i+1})^2 + \dots + (x^p)^2)}$$

dependendo se x_0^i é positivo ou negativo.

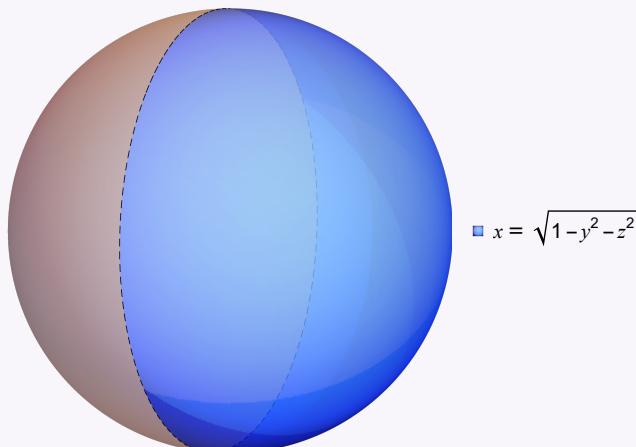


Figura 3.3: $x = \sqrt{1 - y^2 - z^2}$ para $y^2 + z^2 < 1$ em azul

Exemplo 3.20. Considere $f(x, y) = y^2 - x^2(x + 1)$, $(x, y) \in \mathbb{R}^2$.

$$df_{(x_0, y_0)} = \begin{bmatrix} -3x^2 - 2x & 2y \end{bmatrix}$$

Para $(x, y) \in f^{-1}(0, 0)$ temos que $df_{(x, y)} = 0$ somente para $(x, y) = (0, 0)$ e nesse ponto não podemos escrever $f^{-1}(0)$ como gráfico de uma função (veja Figura 3.4). Nos demais pontos podemos escrever $y = x\sqrt{x-1}$ ou $y = -x\sqrt{x-1}$.

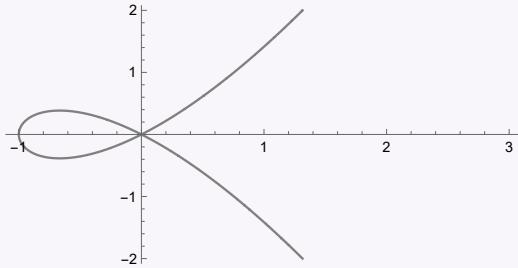


Figura 3.4: $y^2 = x^2(x + 1)$

Exemplo 3.21. O cone $z^2 - x^2 - y^2 = 0$ em \mathbb{R}^3 falha em ser localmente gráfico em $(0, 0, 0)$.

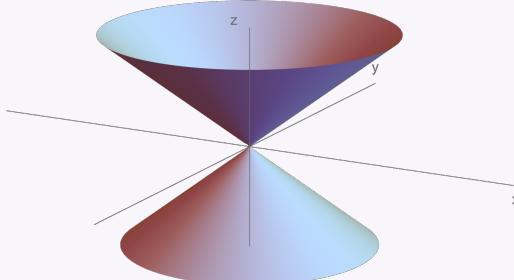


Figura 3.5: $z^2 - x^2 - y^2 = 0$

Teorema 3.22. Considere a função suave $f : U \rightarrow \mathbb{R}^q$, com $U \subset \mathbb{R}^{p+q}$ aberto. Suponha que $M := f^{-1}(0)$ é não vazio e suponha que df_x é sobrejetora para todo $x \in M$. Então, M é uma p -variedade e $T_x M = \ker df_x$ para $x \in f^{-1}(0)$. Adicionalmente, para $x \in f^{-1}(0)$ os vetores

$$\nabla f^1(x), \nabla f^2(x), \dots, \nabla f^q(x)$$

formam uma base de $(T_x M)^\perp$.

Demonstração. Tome $x_0 \in M$. Temos que df_{x_0} tem posto q , isto é, há índices j_1, \dots, j_q tais que os vetores

$$\partial_{j_1} f(x_0), \dots, \partial_{j_q} f(x_0) \in \mathbb{R}^q$$

são linearmente independentes. Sejam k_1, \dots, k_p os índices em $\{1, 2, \dots, p+q\} \setminus \{j_1, \dots, j_q\}$. Considere o isomorfismo linear $T : \mathbb{R}^{p+q} \rightarrow \mathbb{R}^{p+q}$ dado por

$$T(e_1) = e_{k_1}, \dots, T(e_p) = e_{k_p}, T(e_{p+1}) = e_{j_1}, \dots, T(e_{p+q}) = e_{j_q},$$

$g := f \circ T$ e $z_0 := T^{-1}(x_0)$. Temos que as q últimas colunas da matriz Jacobiana de dg_{z_0} são linearmente independentes.

Defina

$$\varphi(u, v) := (u^1, \dots, u^p, g^1(u, v), \dots, g^q(u, v))$$

com $(u, v) \in T^{-1}(U)$.

Temos que a matriz Jacobiana de φ em z_0 é

$$d\varphi_{z_0} = \left[\begin{array}{c|c} \text{Id}_{p \times p} & 0_{p \times q} \\ \hline \frac{\partial g^i}{\partial u^j}(z_0) & \frac{\partial g^i}{\partial v^j}(z_0) \end{array} \right],$$

tem determinante não nulo e pelo teorema da função inversa $\varphi \circ T^{-1}$ é difeomorfismo em uma vizinhança V de x_0 tal que $\varphi \circ T^{-1}(V)$ é aberto. Como

$$\varphi \circ T^{-1}(M \cap V) = \{x \in \mathbb{R}^{p+q} : x^{j_1} = \dots = x^{j_q} = 0\} \cap \varphi \circ T^{-1}(V),$$

temos que M é uma p -variedade.

Considere $v \in T_{x_0}M$. Existe uma curva γ em M tal que $\gamma(0) = x_0$ e $\dot{\gamma}(0) = v$. Assim, $f(\gamma(t)) = 0$ e como consequência $df_{x_0}(v) = 0$. Logo, $T_{x_0}M \subset \ker df_{x_0}$. Como

$$\dim T_{x_0}M = \dim \ker df_{x_0} = p$$

temos

$$T_{x_0}M = \ker df_{x_0}.$$

Como a matriz $(\partial_j f^i(x_0))$ tem posto q , as q -linhas da matriz são linearmente independentes. Adicionalmente, como $T_{x_0}M = \ker df_{x_0}$, temos que

$$df_{x_0}^i(v) = 0 \quad \forall v \in T_{x_0}M,$$

ou seja, $\nabla f^i(x_0) \perp T_{x_0}M$.

Note que o gradiente $[\partial_1 f^i(x_0) \ \partial_2 f^i(x_0) \ \dots \ \partial_{p+q} f^i(x_0)]$ é a i -ésima linha da matriz Jacobiana de f em x_0 . Como a matriz tem posto q , essas linhas são linearmente independentes. Assim, os gradientes $\nabla f^i(x_0)$, com $i = 1, \dots, q$, são linearmente independentes.

Como $T_{x_0}M$ tem dimensão p , segue que $(T_{x_0}M)^\perp$ tem dimensão q e, consequentemente,

$$\nabla f^1(x_0), \nabla f^2(x_0), \dots, \nabla f^q(x_0)$$

formam uma base de $(T_{x_0}M)^\perp$. ■

Exemplo 3.23. Considere $f(x) = -R^2 + \sum_{i=0}^p (x^i)^2$, $x \in \mathbb{R}^{p+1}$, onde $R > 0$. Para $x \in f^{-1}(0)$ temos que $\nabla f(x) \neq 0$, ou seja, df_x é sobrejetora.

Logo, $f^{-1}(0)$ é a esfera de raio R e tem dimensão p . Para $x \in f^{-1}(0)$ temos

$$T_x f^{-1}(0) = \ker df_x = \{h \in \mathbb{R}^{p+1} : \sum_{i=0}^p h^i x^i = 0\}.$$

Similarmente, se $g(x) = -R^2 + x_0^2 - \sum_{i=1}^p (x^i)^2$, para $x \in \mathbb{R}^{p+1}$ com $x^0 > 0$, temos que $g^{-1}(0)$ é uma p -variedade (um hiperbolóide) e

$$T_x g^{-1}(0) = \{h \in \mathbb{R}^{p+1} : -x^0 h^0 + \sum_{i=1}^p h^i x^i = 0\}.$$

Exemplo 3.24. Lembre-se que $\text{SL}(d, \mathbb{R}) = \det^{-1}(1)$ e $\text{GL}(d, \mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{0\})$. Defina $f : \text{GL}(d, \mathbb{R}) \rightarrow \mathbb{R}$ dado por $f(A) = \det(A) - 1$.

Tome $A \in f^{-1}(0)$ e a curva $\gamma(t) = e^t A$ passa por A com velocidade A em $t = 0$.

$$df_A(A) = \frac{d}{dt} \Big|_{t=0} e^{dt} = d \neq 0.$$

Portanto, $df_A \neq 0$ para todo $A \in f^{-1}(0)$.

Logo, $\text{SL}(d, \mathbb{R})$ é uma $d^2 - 1$ variedade em $M_d(\mathbb{R})$.

Para $H \in M_d(\mathbb{R})$ temos

$$\begin{aligned} df_A(H) &= \frac{d}{dt} \Big|_{t=0} f(A \exp(tA^{-1}H)) \\ &= \frac{d}{dt} \Big|_{t=0} \det(A \exp(tA^{-1}H)) \\ &= \frac{d}{dt} \Big|_{t=0} \det(\exp(tA^{-1}H)) \\ &= \frac{d}{dt} \Big|_{t=0} e^{\text{tr}(tA^{-1}H)} \\ &= \text{tr}(A^{-1}H). \end{aligned}$$

Logo, para $A \in \text{SL}(d, \mathbb{R})$ temos

$$T_A \text{SL}(d, \mathbb{R}) = \{H \in M_d(\mathbb{R}): \text{tr}(A^{-1}H) = 0\} = \{AH' \in M_d(\mathbb{R}): \text{tr}(H') = 0\}.$$

Exemplo 3.25. Lembre-se que $\text{O}(d) = \{A \in \text{GL}(d, \mathbb{R}): A^{-1} = A^T\}$. Seja V o espaço das matrizes simétricas. O espaço vetorial V tem base $\frac{E_{ij} + E_{ji}}{2}$ com $i \leq j$, onde E_{ij} é a matriz $d \times d$ com entrada ij igual a 1 e as demais entradas nulas. Assim, $\dim V = \frac{d(d+1)}{2}$.

Defina $f : \text{GL}(d, \mathbb{R}) \rightarrow V$ dada por $f(A) = AA^T - \text{Id}_{\mathbb{R}^d}$.

Considere $A \in \text{O}(d)$ e $H \in M_d(\mathbb{R})$. Temos que

$$\begin{aligned} df_A(H) &= \frac{d}{dt} \Big|_{t=0} f(A + tH) \\ &= AH^T + HA^T. \end{aligned}$$

Logo, para H simétrica,

$$df_A\left(\frac{1}{2}HA\right) = \frac{A(HA)^T + (HA)A^T}{2} = \frac{AA^TH^T + HAA^T}{2} = H.$$

Portanto, df_A é sobrejetora, $\text{O}(d)$ é variedade e tem dimensão

$$\dim \ker df_A = d^2 - \frac{d(d+1)}{2} = \frac{d(d-1)}{2}.$$

Além disso,

$$\begin{aligned} T_A(d) &= \{H \in M_d(\mathbb{R}): AH^T + HA^T = 0\} \\ &= \{HA \in M_d(\mathbb{R}): H + H^T = 0\} \\ &\simeq \{H \in M_d(\mathbb{R}): H \text{ é anti-simétrica}\}. \end{aligned}$$

Definição 3.26. Considere a função suave $f : U \rightarrow \mathbb{R}^q$, $U \subset \mathbb{R}^p$ aberto. Um ponto $s \in \mathbb{R}^q$ é valor regular se para todo $x \in f^{-1}(s)$ temos que $df_x : \mathbb{R}^p \rightarrow \mathbb{R}^q$ é sobrejetora.

Note que pode ocorrer de $f^{-1}(s)$ ser vazio. No entanto, se $f^{-1}(s)$ é não vazio, então tal conjunto é uma $(p-q)$ -variedade.

Para $f : U \rightarrow \mathbb{R}$ suave, $U \subset \mathbb{R}^p$ aberto, e $s \in \mathbb{R}$ o conjunto $f^{-1}(s)$ é um conjuntos de nível. Para os casos em que s é valor regular de f , temos que $f^{-1}(s)$ é uma $(p-1)$ -variedade e $\nabla f(x) \perp T_x f^{-1}(s)$ para $x \in f^{-1}(s)$. De fato, basta notar que

$$T_x f^{-1}(s) = \ker df_x = \{v \in \mathbb{R}^p: df_x(v) = 0\} = \{v \in \mathbb{R}^p: \langle \nabla f(x), v \rangle = 0\}.$$

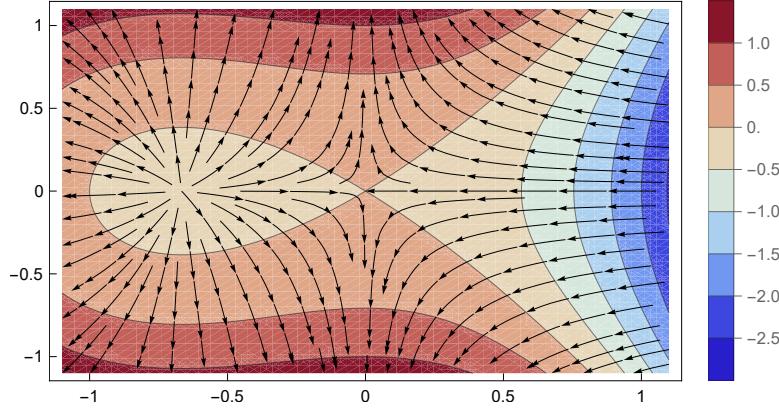


Figura 3.6: Curvas de nível de $f(x, y) = y^2 - x^2(x + 1)$ e o campo gradiente $\nabla f(x, y)$ normal as curvas de nível

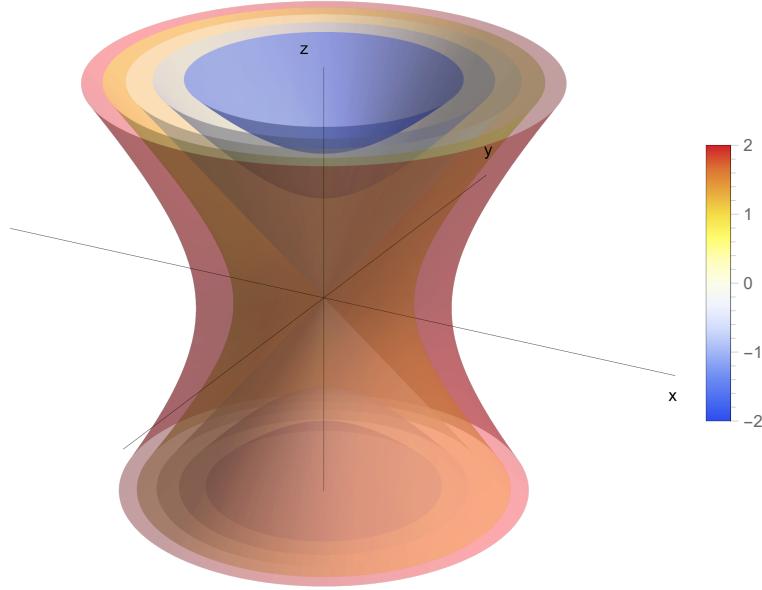


Figura 3.7: Curvas de nível de $f(x, y, z) = x^2 + y^2 - z^2$ e o campo gradiente $\nabla f(x, y)$ normal as superfícies de nível

Definição 3.27. Uma função $f : M \rightarrow N$ entre variedades, com $M \subset \mathbb{R}^p$ e $N \subset \mathbb{R}^q$, é suave se existe uma função suave $F : U \rightarrow \mathbb{R}^q$ definida em uma vizinhança U de M em \mathbb{R}^p tal que $f = F|_M$. Em outras palavras, f é suave se é restrição de uma função suave definida em uma vizinhança de M . A função suave f é difeomorfismo se for invertível e sua inversa for suave também

Exemplo 3.28. Seguindo o Exemplo 1.105, temos a projeção estereográfica

$$f : \mathbb{S}^d \setminus \{N\} \rightarrow \mathbb{R}^d$$

é dada por

$$f(x^1, x^2, \dots, x^d, x^{d+1}) = \left(\frac{x^1}{1-x^{d+1}}, \frac{x^2}{1-x^{d+1}}, \dots, \frac{x^d}{1-x^{d+1}} \right).$$

A função f é suave porque f é restrição de

$$F : \mathbb{R}^{d+1} \setminus \{N\} \rightarrow \mathbb{R}^d$$

é dada por

$$F(x^1, x^2, \dots, x^d, x^{d+1}) = \left(\frac{x^1}{1-x^{d+1}}, \frac{x^2}{1-x^{d+1}}, \dots, \frac{x^d}{1-x^{d+1}} \right).$$

Note também que a inversa de f é suave também. Portanto, f é difeomorfismo. O homeomorfismo no Exemplo 1.104 é difeomorfismo também.

Definição 3.29. Seja $f : M \rightarrow N$ suave. Definimos $df_x : T_x M \rightarrow T_{f(x)} N$ como sendo $df_x(v) = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t))$, onde γ é uma curva em M passando em x com velocidade v em $t = 0$.

Observação 3.30. Em uma vizinhança $U \subset \mathbb{R}^p$ de M temos uma função suave $F : U \rightarrow \mathbb{R}^q$ tal que $f = F|_M$. Assim

$$\left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = \left. \frac{d}{dt} \right|_{t=0} F(\gamma(t)) = dF_x(v)$$

e, consequentemente, $df_x(v)$ não depende da escolha da curva γ em M satisfazendo $\gamma(0) = x$ e $\dot{\gamma}(0) = v$.

Teorema 3.31. Sejam $f : M \rightarrow N$ e $g : N \rightarrow S$ funções suaves entre variedades. Então,

$$g \circ f : M \rightarrow S$$

é suave e

$$d(g \circ f)_x = dg_{f(x)} \circ df_x.$$

Demonstração. A primeira afirmação segue da definição, basta estender f e g para funções suaves definidas em vizinhanças abertas de x e $y := f(x)$.

Se $\gamma(t)$ é uma curva passando por x com velocidade v em $t = 0$, então $f \circ \gamma(t)$ passa com velocidade $df_x(v)$ por y em $t = 0$.

$$d(g \circ f)_x(v) = \left. \frac{d}{dt} \right|_{t=0} g(f(\gamma(t))) = dg_y \left(\left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) \right) = dg_y(df_x(v)).$$

■

Exemplo 3.32. (Mapa de Gauss e curvaturas) Sejam $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ suave e 0 um valor regular na imagem de f . Considere a superfície $M = f^{-1}(0)$. Temos que $n(x) := \frac{\nabla f(x)}{|\nabla f(x)|}$ é um vetor unitário normal a M .

A função $n : M \rightarrow \mathbb{S}^2$ é suave e se chama mapa de Gauss de M . Considere $v \in T_x M$ e uma curva γ em M satisfazendo $\gamma(0) = x$ e $\dot{\gamma}(0) = v$.

Temos que

$$0 = \left. \frac{d}{dt} \right|_{t=0} \langle n(\gamma(t)), n(\gamma(t)) \rangle = 2 \langle dn_x(v), n(x) \rangle,$$

ou seja, $dn_x(v) \in T_x M$.

Assim, a aplicação linear $dn_x : T_x M \rightarrow T_x M$ está bem definida. Além disso, o mapa de Gauss é auto-adjunto. De fato, considere $u, v \in T_x M$ e uma carta $\varphi : U \rightarrow \varphi(U)$. Podemos supor que $\varphi(x) = 0$ a menos de compor φ com uma translação e

$$\varphi(U \cap M) = (\mathbb{R}^2 \times \{0\}) \cap \varphi(U).$$

Temos a parametrização $r(s, t) := \varphi^{-1}(sd\varphi_x(u) + td\varphi_x(v))$ satisfazendo

$$\frac{\partial r}{\partial s}(0, 0) = u \quad \text{e} \quad \frac{\partial r}{\partial t}(0, 0) = v.$$

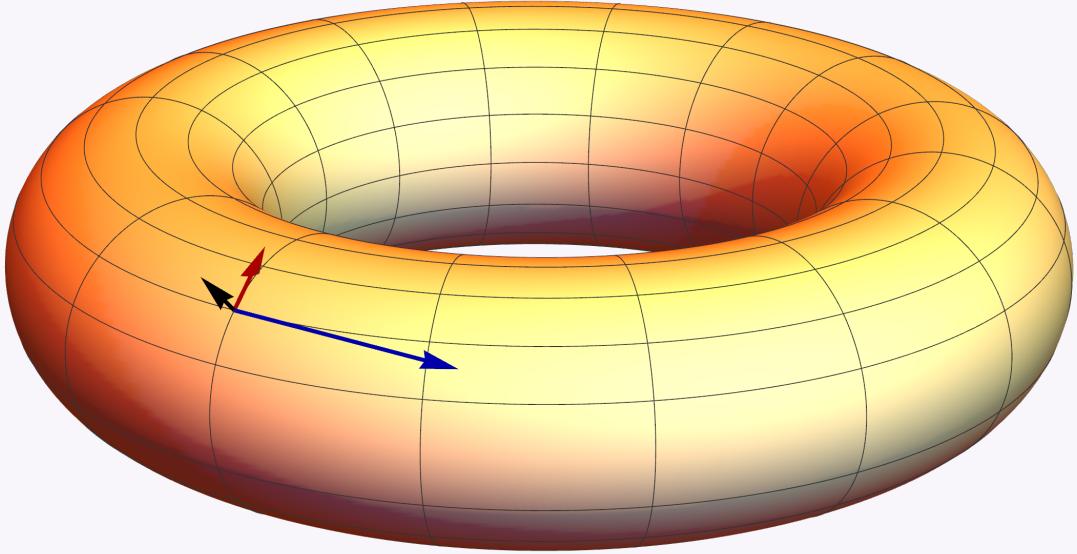


Figura 3.8: Toro parametrizado por $r(s, t) = ((3 + \cos(t)) \cos(s), (3 + \cos(t)) \sin(s), \sin(t))$ com campos vetoriais tangentes $\partial_s \mathbf{r}$ e $\partial_t \mathbf{r}$ e campo normal unitário $\mathbf{n}(s, t) = \frac{\partial_t \mathbf{r} \times \partial_s \mathbf{r}}{|\partial_t \mathbf{r} \times \partial_s \mathbf{r}|}$

$$\begin{aligned}\frac{\partial^2}{\partial s \partial t} \langle n(r(s, t)), r(s, t) \rangle &= \frac{\partial}{\partial s} \left\langle \frac{\partial}{\partial t} n(r(s, t)), r(s, t) \right\rangle + \langle n(r(s, t)), r_t(s, t) \rangle \\ &= \frac{\partial}{\partial s} \left(\left\langle \frac{\partial}{\partial t} n(r(s, t)), r(s, t) \right\rangle \right) \\ &= \left\langle \frac{\partial^2}{\partial s \partial t} n(r(s, t)), r(s, t) \right\rangle + \left\langle \frac{\partial}{\partial t} n(r(s, t)), r_s(s, t) \right\rangle\end{aligned}$$

Similarmente,

$$\frac{\partial^2}{\partial t \partial s} \langle n(r(s, t)), r(s, t) \rangle = \left\langle \frac{\partial^2}{\partial t \partial s} n(r(s, t)), r(s, t) \right\rangle + \left\langle \frac{\partial}{\partial s} n(r(s, t)), r_t(s, t) \right\rangle.$$

Como derivadas parciais comutam, temos que

$$\left\langle \frac{\partial}{\partial t} n(r(s, t)), r_s(s, t) \right\rangle = \left\langle \frac{\partial}{\partial s} n(r(s, t)), r_t(s, t) \right\rangle,$$

$$\langle d n_{r(s,t)}(r_t(s,t)), r_s(s,t) \rangle = \langle d n_{r(s,t)}(r_s(s,t)), r_t(s,t) \rangle$$

e fazendo $(s, t) = (0, 0)$ temos

$$\langle d n_x(v), u \rangle = \langle d n_x(u), v \rangle.$$

Portanto, $d n_x : T_x M \rightarrow T_x M$ é auto-adjunto e portanto diagonalizável.

A curvatura de Gauss $K(x)$ é o determinante de $d n_x$ e a curvatura média $H(x)$ é $\frac{1}{2} \operatorname{tr}(d n_x)$, ou seja, se λ_1, λ_2 são os autovalores de $d n_x$, que são reais porque esse operador é auto-adjunto, então

$$K(x) = \lambda_1 \lambda_2 \quad \text{e} \quad H(x) = \frac{\lambda_1 + \lambda_2}{2}.$$

3.3 Multiplicadores de Lagrange

Teorema 3.33. (Teorema dos multiplicadores de Lagrange) Considere duas funções suaves $f : U \rightarrow \mathbb{R}$ e $g : U \rightarrow \mathbb{R}^q$, onde $U \subset \mathbb{R}^p$ é aberto.

Suponha que 0 é valor regular de g e defina $M = g^{-1}(0)$. Considere $x^* \in M$ tal que x^* é

máximo (mínimo) local $f|_M$.

Temos que

$$\nabla f(x^*) \in \mathbb{R} \nabla g^1(x^*) \oplus \mathbb{R} \nabla g^2(x^*) \oplus \cdots \oplus \mathbb{R} \nabla g^q(x^*).$$

Demonstração. Seja γ uma curva suave em M passando por x^* com velocidade v em $t = 0$.

Temos que 0 é máximo (mínimo) local da função $t \mapsto f(\gamma(t))$. Consequentemente, temos

$$\frac{d}{dt} \Big|_{t=0} f(\gamma(t)) = 0,$$

ou seja, $df_{x^*}(v) = 0$. Portanto, $\nabla f(x^*) \perp T_{x^*} M$.

Pelo Teorema 3.22 temos que $\nabla f(x^*)$ é combinação linear da base

$$\nabla g^1(x^*), \nabla g^2(x^*), \dots, \nabla g^q(x^*)$$

de $(T_{x^*} M)^\perp$. ■

Exemplo 3.34. Considere o conjunto de números reais $S = \{x_1, \dots, x_n\}$. Suponha que temos a distribuição de probabilidade $P(X = x_i) = p_i$ para a variável aleatória X , com $p_i \geq 0$ e $\sum_{i=1}^n p_i = 1$.

A entropia da variável aleatória X é

$$H(p) = \sum_{i=1}^n -p_i \log(p_i),$$

onde $p = (p_1, \dots, p_n)$. Note que podemos definir $0 \log(0) = 0$ porque $\lim_{x \rightarrow 0^+} x \log(x) = 0$.

Procuremos a distribuição de probabilidade p que maximiza a entropia.

$$\begin{cases} \text{Maximizar: } & H(p) \\ \text{Restrições: } & \sum_{i=1}^n p_i = 1 \\ & p_1, \dots, p_n \geq 0 \end{cases}$$

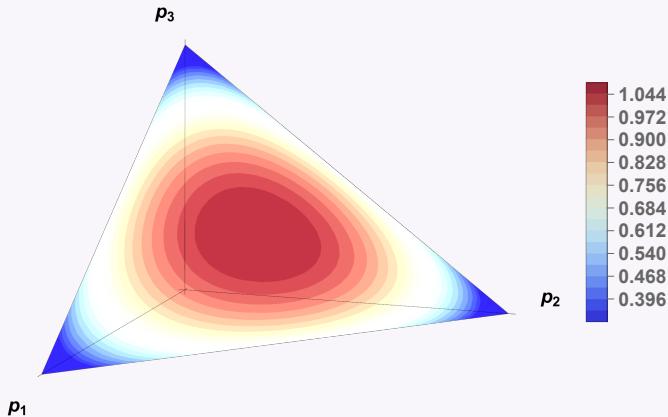


Figura 3.9: $H(p_1, p_2, p_3)$ com $p_1 + p_2 + p_3 = 1$ e $p_1, p_2, p_3 \geq 0$.

A função H admite um ponto de máximo p^* no convexo

$$\Delta := \left\{ p \in \mathbb{R}^n : \sum_{i=1}^n p_i = 1 \text{ e } p_1, \dots, p_n \geq 0 \right\}$$

por compacidade.

Suponha que p^* está no interior de Δ . Pelo teorema de multiplicadores de Lagrange existe $\lambda \in \mathbb{R}$ tal que

$$\nabla H(p^*) = \lambda(1, \dots, 1),$$

$$-\log(p_i^*) - 1 = \lambda \quad \text{para } i = 1, \dots, n,$$

onde usamos que

$$\frac{\partial}{\partial p_i} p_i \log(p_i) = \log(p_i) + 1.$$

Portanto, $\log(p_i^*) = \log(p_j^*)$ para $i \neq j$ e como consequência $p_i^* = \frac{1}{n}$, ou seja, X tem de ter distribuição uniforme.

Mostremos que $p^* = (\frac{1}{n}, \dots, \frac{1}{n})$ é ponto de máximo em Δ .

No aberto $(0, 1)^n$, o Hessiano de H é

$$\mathbf{Hess}H_p(v) = \sum_{i=1}^n -\frac{1}{p_i}(v_i)^2,$$

ou seja, $\mathbf{Hess}H_p(v) < 0$ para $v \neq 0$.

Tome $p \in \Delta$ diferente de p^* e defina $v = p - p^*$. Como

$$\frac{d^2}{dt^2} H(p^* + tv) = \mathbf{Hess}H_{p^*+tv}(v) < 0,$$

$$\left. \frac{d}{dt} \right|_{t=0} H(p^* + tv) = \langle \nabla H(p^*), v \rangle = \left(-\log\left(\frac{1}{n}\right) - 1 \right) \sum_{i=1}^n v_i = 0$$

temos que $\frac{d}{dt} H(p^* + tv) < 0$ para $t \in (0, 1)$ e, consequentemente, $t \mapsto H(p^* + tv)$ é estritamente decrescente.

Portanto, tomando $t \rightarrow 1^-$, temos $H(p) < H(p^*)$.

Exemplo 3.35. Considere

$$S = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1 \quad \text{e} \quad x_1, \dots, x_n \geq 0 \right\}.$$

O máximo de $x \mapsto \sqrt[n]{x_1 \cdots x_n}$ não ocorre na fronteira de S porque essa função é positiva no interior de S e nula na fronteira.

Procuremos o máximo dessa função. Maximizar tal função é o mesmo que maximizar $\log(\sqrt[n]{x_1 \cdots x_n}) = \sum_{i=1}^n \frac{1}{n} \log(x_i)$.

$$\begin{cases} \text{Maximizar:} & \sum_{i=1}^n \frac{1}{n} \log(x_i) \\ \text{Restrições:} & \sum_{i=1}^n x_i = 1 \\ & x_1, \dots, x_n > 0 \end{cases}$$

Considere $f(x) = \sum_{i=1}^n \frac{1}{n} \log(x_i)$ definida para $x_1, \dots, x_n > 0$.

Temos que

$$\nabla f(x) = \frac{1}{n} \left(\frac{1}{x_1}, \dots, \frac{1}{x_n} \right),$$

$$\mathbf{Hess}f_x(v) = -\sum_{i=1}^n \frac{1}{n} \frac{v_i^2}{x_i^2}.$$

Procuremos o ponto crítico x^* de f no interior de S . Existe $\lambda \in \mathbb{R}$ tal que

$$\nabla f(x^*) = \lambda \nabla_{x^*} \left(\sum_{i=1}^n x_i \right) = \lambda(1, \dots, 1),$$

de onde segue que $x^* = (\frac{1}{n}, \dots, \frac{1}{n})$. Mostremos que x^* é ponto de máximo de f . Considere x no interior de S diferente de x^* . Existe $\theta \in (0, 1)$ tal que

$$f(x) = f(x^*) + \langle \nabla f(x^*), x - x^* \rangle + \frac{1}{2} \mathbf{Hess}_{x^*+\theta(x-x^*)} f(x - x^*),$$

$$f(x) = \log\left(\frac{1}{n}\right) + \frac{1}{n^2} \sum_{i=1}^n (x_i - x_i^*) + \frac{1}{2} \mathbf{Hess}_{x^* + \theta(x-x^*)} f(x - x^*) \geq \log\left(\frac{1}{n}\right),$$

$$f(x) = \log\left(\frac{1}{n}\right) + \frac{1}{2} \mathbf{Hess}_{x^* + \theta(x-x^*)} f(x - x^*) > \log\left(\frac{1}{n}\right),$$

Logo,

$$\frac{1}{n} > \sqrt[n]{x_1 \cdots x_n}$$

para $x \in S$. Portanto,

$$\frac{1}{n} \geq \sqrt[n]{x_1 \cdots x_n}$$

com a igualdade valendo somente quando

$$x_1 = \cdots = x_n = \frac{1}{n}.$$

Considere $u_1, \dots, u_n \geq 0$. Se algum u_i se anula, temos

$$\frac{u_1 + \cdots + u_n}{n} \geq \sqrt[n]{u_1 \cdots u_n}.$$

Se todos os u_i 's são positivos então definindo $x_i = \frac{u_i}{u_1 + \cdots + u_n}$ temos que

$$\frac{u_1 + \cdots + u_n}{n} \geq \sqrt[n]{u_1 \cdots u_n}$$

com a igualdade ocorrendo somente quando os u_i 's são todos iguais.

Exemplo 3.36. (Teorema espectral) Seja V um espaço vetorial real de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$. Seja $A : V \rightarrow V$ uma transformação linear auto-adjunta, isto é,

$$\langle Au, v \rangle = \langle u, Av \rangle \quad u, v \in V.$$

Provemos que V admite uma base ortonormal de autovetores para A com autovalores reais. Faremos um argumento por indução. O resultado é trivialmente verdadeiro para $\dim V = 1$. Suponha que o resultado vale para dimensão n e provemos que vale para dimensão $n+1$. Seja V um espaço vetorial de dimensão $n+1$. Podemos supor que $V = \mathbb{R}^{n+1}$ com o produto interno canônico a menos de escolher uma base ortonormal para V .

Considere a esfera

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : \langle x, x \rangle = 1\}$$

e a função suave $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ dada por $f(x) := \langle Ax, x \rangle$.

Como

$$f(x+v) = \langle Ax, x \rangle + \langle Ax, v \rangle + \langle x, Av \rangle + o(|v|),$$

$$f(x+v) = \langle Ax, x \rangle + 2\langle Ax, v \rangle + o(|v|),$$

temos $df_x(v) = 2\langle Ax, v \rangle$. Assim, $\nabla f(x) = 2Ax$.

O gradiente de $g(x) = \langle x, x \rangle$ é $2x$.

Como a esfera $\mathbb{S}^n = g^{-1}(1)$ é compacta, existe $b_{n+1} \in \mathbb{S}^n$ que é máximo de $f|_{\mathbb{S}^n}$. Em outras palavras, estamos resolvendo o problema

$$\begin{cases} \text{Maximizar:} & f(x) \\ \text{Restrição:} & g(x) = 1 \end{cases}$$

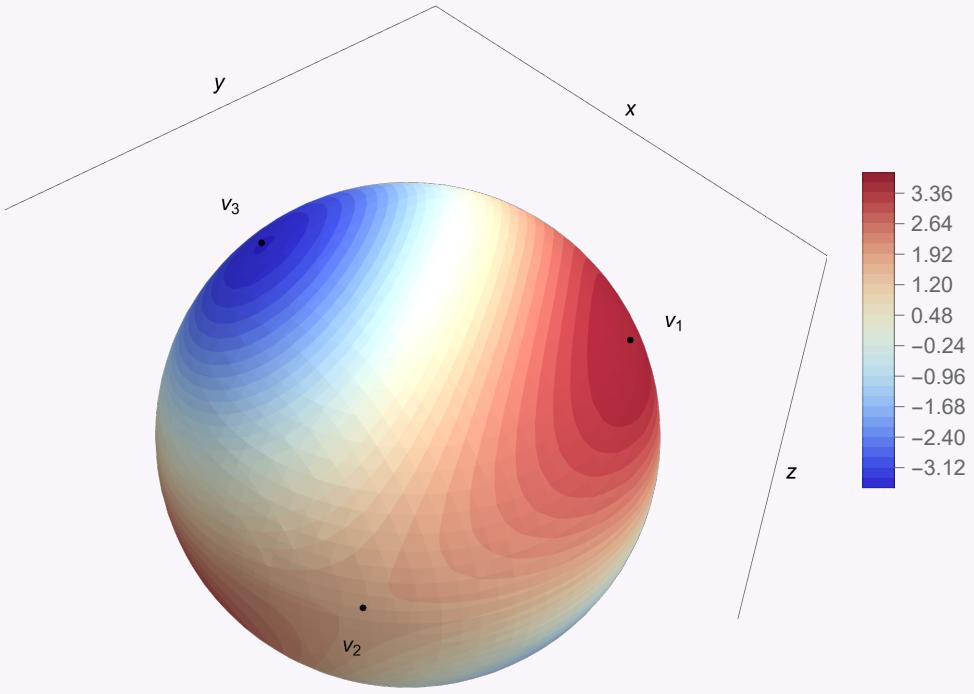


Figura 3.10: Para $A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \\ 3 & 2 & -1 \end{bmatrix}$ temos as curvas de nível de $f(x, y, z) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \\ 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ sobre a esfera. Os vetores unitários v_1, v_2, v_3 são pontos críticos de f e formam uma base ortonormal de autovetores para A : v_1 é ponto de máximo, v_2 é ponto de sela, e v_3 é ponto de mínimo.

Pelo teorema dos múltiplos de Lagrange, temos que existe $\lambda_{n+1} \in \mathbb{R}$ tal que

$$\nabla f(b_{n+1}) = \lambda_{n+1} \nabla g(b_{n+1}),$$

$$Ab_{n+1} = \lambda_{n+1} b_{n+1}$$

Portanto, b_{n+1} é autovetor de A com autovalor real λ_{n+1} . O espaço vetorial $W = b_{n+1}^\perp$ tem dimensão n e $AW \subset W$. De fato, se $x \in W$, então

$$\langle Ax, b_{n+1} \rangle = \langle x, Ab_{n+1} \rangle = \lambda_{n+1} \langle x, b_{n+1} \rangle = 0,$$

ou seja, $Ax \in W$. Além disso, $A|_W : W \rightarrow W$ é auto-adjunto. Por hipótese de indução, há uma base ortonormal b_1, \dots, b_n para W com autovalores $\lambda_1, \dots, \lambda_n$.

Logo, há uma base ortonormal b_1, \dots, b_{n+1} de autovetores para A com autovalores $\lambda_1, \dots, \lambda_{n+1}$ reais.

Capítulo 4

Convergência uniforme

Sejam X um espaço métrico e V um espaço vetorial de dimensão finita. Como V tem dimensão finita, as normas são equivalentes entre si. Além disso, V é completo com qualquer norma.

Seja $\text{BC}(X, V)$ o espaço das funções contínuas de X em V limitadas, que é um espaço vetorial real. Note que se X é compacto, então $\text{BC}(X, V) = \text{C}(X, V)$ porque funções contínuas definidas em um compacto são limitadas.

Considere a norma

$$\|f\|_\infty = \sup_{z \in X} |f(z)|.$$

Essa norma se chama norma do supremo. A convergência nessa norma leva o nome de convergência uniforme. Se $f_n \rightarrow f$, então dizemos que f_n converge uniformemente para f .

Assim, $\text{BC}(X, V)$ é um espaço vetorial normado. Se $V = \mathbb{R}$, então escrevemos $B(X, \mathbb{R})$ como $\text{BC}(X)$.

Teorema 4.1. $\text{BC}(X, V)$ é espaço de Banach, isto é, um espaço vetorial normado completo.

Demonstração. Seja f_n uma sequência de Cauchy de $\text{BC}(X, V)$. Para cada $x \in X$ temos

$$|f_n(x) - f_m(x)| \leq \sup_{z \in X} |f_n(z) - f_m(z)| = \|f_n - f_m\|_\infty \rightarrow 0$$

quando $n, m \rightarrow \infty$.

Logo, $f_n(x)$ é sequência de Cauchy em V . Defina $f(x) := \lim_{n \rightarrow \infty} f_n(x)$.

Mostremos que f é contínua. Tome $x_0 \in X$. Dado $\varepsilon > 0$ existe N tal que

$$|f_N(x_0) - f(x_0)| < \varepsilon \quad \text{e} \quad \|f_n - f_N\|_\infty < \varepsilon$$

para $n \geq N$. Como f_N é contínua, existe $\delta > 0$ tal que $|f_N(x) - f_N(x_0)| < \varepsilon$ para $|x - x_0| < \delta$.

Logo, para $|x - x_0| < \delta$, temos

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &\leq |f(x) - f_n(x)| + \|f_n - f_N\|_\infty + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< |f(x) - f_n(x)| + 3\varepsilon. \end{aligned}$$

Fazendo $n \rightarrow \infty$, obtemos $|f(x) - f(x_0)| \leq 3\varepsilon$. Portanto, f é contínua.

Mostremos que f é limitada. Fixe $x_0 \in X$. Existe N tal que $\|f_n - f_N\|_\infty \leq 1$ para $n \geq N$. Logo,

$$|f(x) - f_N(x)| \leq \lim_{n \rightarrow \infty} |f_n(x) - f_N(x)| \leq 1,$$

ou seja, f é limitada.

Falta mostrar que $f_n \rightarrow f$ uniformemente. Dado $\varepsilon > 0$ existe N tal que $\|f_n - f_m\|_\infty < \varepsilon$ para $n, m \geq N$.

Portanto, para qualquer $z \in X$ e $m \geq N$

$$|f(z) - f_m(z)| = \lim_{n \rightarrow \infty} |f_n(z) - f_m(z)| \leq \varepsilon,$$

$$\|f - f_m\|_\infty \leq \varepsilon.$$

■

Exemplo 4.2. (Teorema de Existência e unicidade para EDO's) Sejam $x_0 \in \mathbb{R}^p$ e f uma função definida em uma vizinhança de x_0 com valores em \mathbb{R}^p . Considere a equação diferencial ordinária

$$\begin{cases} \dot{x} = f(x) \\ x(0) = x_0 \end{cases} \quad (4.3)$$

Suponha que f é Lipschitz em uma vizinhança de x_0 . Existem $r > 0$ e $L > 0$ tal que

$$|f(u) - f(v)| \leq L|u - v|$$

para $u, v \in \overline{B(x_0, r)}$.

Mostremos que existe uma solução $x(t)$ para equação diferencial tal que $x(0) = x_0$. Considere o número $\delta > 0$ a ser definido posteriormente. O subconjunto

$$X = \{x : [-\delta, \delta] \rightarrow \mathbb{R}^p : x(0) = x_0 \text{ e } x([- \delta, \delta]) \subset \overline{B(x_0, r)}\}.$$

de $C([-\delta, \delta], \mathbb{R}^p)$ é fechado e, em particular, X é completo com a métrica

$$d(x, y) := \sup_{[-\delta, \delta]} |x(s) - y(s)|.$$

Considere o operador

$$\Lambda(x)(t) := x_0 + \int_0^t f(x(s))ds$$

definido para funções $x \in X$. Queremos que $\Lambda(X) \subset X$ e que Λ seja contração. Para isso, devemos tomar um $\delta > 0$ conveniente. Note que para $x, y \in X$ e $0 \leq t \leq \delta$ temos

$$|\Lambda(x)(t) - x_0| \leq \int_0^t |f(x(s))|ds \leq \max_{\overline{B(x_0, r)}} |f(u)|\delta,$$

$$|\Lambda(x)(t) - \Lambda(y)(t)| \leq \int_0^t |f(x(s)) - f(y(s))|ds \leq L\delta \sup_{[-\delta, \delta]} |x(s) - y(s)|.$$

Similarmente, para $x, y \in X$ e $-\delta \leq t \leq 0$ temos

$$|\Lambda(x)(t) - x_0| \leq \int_t^0 |f(x(s))|ds \leq \max_{\overline{B(x_0, r)}} |f(u)|\delta,$$

$$|\Lambda(x)(t) - \Lambda(y)(t)| \leq \int_t^0 |f(x(s)) - f(y(s))|ds \leq L\delta \sup_{[-\delta, \delta]} |x(s) - y(s)|.$$

Tomando $\delta > 0$ tal que

$$\max_{\overline{B(x_0, r)}} |f(u)|\delta \leq r \quad \text{e} \quad 2L\delta \leq 1$$

temos que Λ leva X em X e é uma contração:

$$\max_{[-\delta, \delta]} |\Lambda(x)(s) - \Lambda(y)(s)| \leq \frac{1}{2} \max_{[-\delta, \delta]} |x(s) - y(s)|.$$

Pelo Teorema de ponto fixo de Banach 3.1, há $\hat{x} \in X$ tal que

$$\hat{x}(t) = x_0 + \int_0^t f(\hat{x}(s))ds,$$

ou seja,

$$\dot{\hat{x}} = f(\hat{x}) \quad \text{e} \quad \hat{x}(0) = x_0.$$

Portanto, a equação diferenciável ordinária (4.3) admite uma solução. Como a derivada de \hat{x} se iguala a uma função contínua, temos que \hat{x} é de classe C^1 .

Agora mostremos a unicidade da solução: considere duas soluções y, z para a equação diferenciável ordinária (4.3), definidas em uma vizinhança de $t = 0$ e satisfazendo $y(0) = z(0) = x_0$. Para $\delta > 0$ suficientemente pequeno temos que y, z estão definidas em $[-\delta, \delta]$,

$$y(t), z(t) \in \overline{B(x_0, r)} \quad \forall t \in [-\delta, \delta],$$

$$\max_{B(x_0, r)} |f(u)|\delta \leq r,$$

$$2L\delta \leq 1$$

e, consequentemente, Λ é contração em

$$\{x : [-\delta, \delta] \rightarrow \mathbb{R}^d : x(0) = x_0 \quad \text{e} \quad x([-\delta, \delta]) \subset \overline{B(x_0, r)}\}.$$

Temos que $y = z$ em $[-\delta, \delta]$ porque y e z , quando restritas a tal intervalo, são pontos fixos de Λ e há apenas um ponto fixo para Λ pelo Teorema de ponto fixo de Banach 3.1.

Observação 4.4. Sejam x_1, x_2 soluções de $\dot{x} = f(x)$ definidas em uma vizinhança de t_0 tais que $x_1(t_0) = x_2(t_0)$, então $x_1(t) = x_2(t)$ em uma vizinhança de t_0 porque $z_1(t) = x_1(t - t_0)$ e $z_2(t) = x_2(t - t_0)$ são soluções de

$$\begin{cases} \dot{z} = f(z) \\ z(0) = x_1(t_0) \end{cases}$$

Observação 4.5. Para f de classe C^1 , $k \geq 1$, temos que a função f é localmente Lipschitz e o teorema de existência e unicidade se aplica. Além disso, como

$$\dot{x} = f(x(t)),$$

obtemos que x é de classe C^{k+1} .

Exemplo 4.6. Seguindo o exemplo anterior, considere $y_0 \in \overline{B(x_0, \frac{r}{2})} \subset \mathbb{R}^p$ e defina

$$X = \{y \in C([-\delta, \delta] \rightarrow \mathbb{R}^p) : y(0) = y_0, \quad y([-\delta, \delta]) \subset \overline{B(x_0, r)}\}$$

com a métrica

$$d(y, y') := \max_{[-\delta, \delta]} |y(s) - y'(s)|.$$

Temos que X é completo. Tome $\delta > 0$ tal que

$$\max_{B(x_0, r)} |f(u)|\delta \leq \frac{r}{2} \quad \text{e} \quad 2L\delta \leq 1$$

e

$$\Lambda(y)(t) := y_0 + \int_0^t f(y(s))ds.$$

Temos $\Lambda(y)(0) = y_0$,

$$|\Lambda(x)(t) - x_0| \leq |x_0 - y_0| + \max_{B(x_0, r)} |f(u)|\delta \leq r \quad \text{e} \quad |\Lambda(y)(t) - \Lambda(y')(t)| \leq \frac{1}{2} \max_{[-\delta, \delta]} |y(s) - y'(s)|.$$

Assim, $\Lambda(X) \subset X$ e Λ tem um ponto fixo $y : [-\delta, \delta] \rightarrow \overline{B(x_0, r)}$ que é solução da EDO

$$\begin{cases} \dot{y} = f(y) \\ y(0) = y_0 \end{cases}$$

para cada $y_0 \in \overline{B(x_0, \frac{r}{2})}$.

Defina $g : [-\delta, \delta] \times B(x_0, \frac{r}{2}) \rightarrow \overline{B(x_0, r)}$, onde $g(t, u)$ é solução da EDO

$$\begin{cases} \partial_t g(t, u) = f(g(t, u)) \\ g(0, u) = u \end{cases}$$

Mostremos que

$$|g(t, u') - g(t, u)| \leq |u' - u|e^{Lt}$$

para $u', u \in \overline{B(x_0, \frac{r}{2})}$.

Sem perda de generalidade podemos assumir $u \neq u'$. Considere $\phi(t) := \frac{1}{2} \frac{|g(t, u') - g(t, u)|^2}{|u' - u|^2 e^{2Lt}}$. Temos

$$\begin{aligned} \dot{\phi}(t) &= \frac{\langle \partial_t g(t, u') - \partial_t g(t, u), g(t, u') - g(t, u) \rangle - L(g(t, u') - g(t, u))^2}{|u' - u|^2 e^{2Lt}} \\ &= \frac{\langle f(g(t, u')) - f(g(t, u)), g(t, u') - g(t, u) \rangle - L(g(t, u') - g(t, u))^2}{|u' - u|^2 e^{2Lt}} \\ &\leq \frac{|f(g(t, u')) - f(g(t, u))| |g(t, u') - g(t, u)| - L(g(t, u') - g(t, u))^2}{|u' - u|^2 e^{2Lt}} \\ &= \frac{(|f(g(t, u')) - f(g(t, u))| - L|g(t, u') - g(t, u)|) |g(t, u') - g(t, u)|}{|u' - u|^2 e^{2Lt}} \\ &\leq 0 \end{aligned}$$

onde usamos a desigualdade de Cauchy-Schwarz.

Assim, ϕ é decrescente com $\phi(0) = \frac{1}{2}$. Logo,

$$|g(t, u') - g(t, u)| \leq |u' - u|e^{Lt}, \quad 0 \leq t \leq \delta.$$

Definindo $h(t, u) := g(-t, u)$, temos que

$$\begin{cases} \partial_t h(t, u) = -f(h(t, u)) \\ h(0, u) = u \end{cases}$$

Portanto,

$$|h(t, u') - h(t, u)| \leq |u' - u|e^{Lt}, \quad 0 \leq t \leq \delta,$$

$$|g(-t, u') - g(-t, u)| \leq |u' - u|e^{Lt}, \quad 0 \leq t \leq \delta,$$

$$|g(t, u') - g(t, u)| \leq |u' - u|e^{-Lt}, \quad -\delta \leq t \leq 0.$$

Em resumo, temos

$$|g(t, u') - g(t, u)| \leq |u' - u|e^{L|t|}, \quad -\delta \leq t \leq \delta.$$

Da desigualdade triangular obtemos

$$|g(t', u') - g(t, u)| \leq |g(t', u') - g(t', u)| + |g(t', u) - g(t, u)| \leq |u' - u|e^{L|t'|} + |g(t', u) - g(t, u)|$$

e consequentemente

$$\lim_{(t', u') \rightarrow (t, u)} |g(t', u') - g(t, u)| = 0.$$

Portanto, $g : [-\delta, \delta] \times \overline{B(x_0, \frac{r}{2})} \rightarrow \overline{B(x_0, r)}$ é contínua. Desta forma, obtemos que as soluções dependem continuamente das condições iniciais. Observe também que se $u' \rightarrow u$, obtemos que $g(-, u') \rightarrow g(-, u)$ uniformemente no intervalo $[-\delta, \delta]$ porque

$$\max_{[-\delta, \delta]} |g(s, u') - g(s, u)| \leq |u' - u|e^{L\delta} \rightarrow 0.$$

Corolário 4.7. (Teste de convergência de Weierstrass) Seja g_n uma sequência em $\text{BC}(X, V)$. Suponha que

$$\sum_{n=1}^{\infty} \|g_n\|_{\infty} < \infty.$$

Então $f = \sum_{n=1}^{\infty} g_n$ está bem definida, é contínua e a convergência é uniforme.

Demonstração. Basta notar que $f_n = g_1 + \dots + g_n$ é uma sequência de Cauchy em $\text{BC}(X)$. ■

Exemplo 4.8. A função $\exp : M_d(\mathbb{R}) \rightarrow M_d(\mathbb{R})$ dada por

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

converge pontualmente e uniformemente em cada bola $B(0, R)$ em M_d com respeito a norma do operador $|\cdot|$. De fato, para $|A| < R$ temos

$$\sum_{n=0}^{\infty} \frac{|A^n|}{n!} \leq \sum_{n=0}^{\infty} \frac{|A|^n}{n!} \leq \sum_{n=0}^{\infty} \frac{R^n}{n!} < \infty.$$

Desta forma, a exponencial $\exp : M_d(\mathbb{R}) \rightarrow M_d(\mathbb{R})$ é contínua.

Teorema 4.9. Seja $f_n : [a, b] \rightarrow V$ uma sequência de funções contínuas convergindo uniformemente para $f : [a, b] \rightarrow V$. Então

$$\int_a^b f(t) dt = \lim_{n \rightarrow \infty} \int_a^b f_n(t) dt.$$

Demonstração. Basta notar que

$$\left| \int_a^b f(t) - f_n(t) dt \right| \leq \int_a^b |f(t) - f_n(t)| dt \leq \|f - f_n\|_{\infty} \rightarrow 0.$$

■

Corolário 4.10. Seja $f_n : [a, b] \rightarrow V$ uma sequências de funções de classe C^1 que converge pontualmente para uma função $f : [a, b] \rightarrow V$. Suponha que f'_n é sequência de Cauchy em $C([a, b], V)$. Então f é de classe C^1 e

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x).$$

Demonstração. Seja $g \in C([a, b])$ o limite de f'_n na convergência uniforme.

Pelo Teorema 4.9, temos que

$$\int_a^x g(t) dt = \lim_{n \rightarrow \infty} \int_a^x f'_n(t) dt = \lim_{n \rightarrow \infty} f_n(x) - f_n(a) = f(x) - f(a),$$

$$f(x) - f(a) = \int_a^x g(t) dt.$$

Portanto, f é de classe C^1 e $f' = g$.

■

Exemplo 4.11. Seja $f : (-R, R) \rightarrow \mathbb{R}$ uma função e suponha que

$$f(t) = \sum_{k=0}^{\infty} a_k t^k,$$

onde a série acima converge para todo $t \in (-R, R)$. Defina $p_k(t) = a_k t^k$.

Tome $r \in (0, R)$. Mostremos que a série acima converge uniformemente em $[-r, r]$ com

$0 < r < R$. Tome s tal que $r < s < R$. Temos que

$$\sum_{k=0}^{\infty} p_k(s)$$

converge, ou seja, existe $k_0 \in \mathbb{Z}_{\geq 1}$ tal que para $k \geq k_0$ temos

$$|p_k(s)| < 1,$$

$$|a_k| < \frac{1}{s^k},$$

ou seja, para $|t| \leq r$ temos

$$|p_k(t)| < \left(\frac{r}{s}\right)^k.$$

Portanto, $\sup_{|t| \leq r} |p_k(t)| \leq \left(\frac{r}{s}\right)^k$ para $k \geq k_0$. Logo, pelo Corolário 4.7 temos que a série que define f converge uniformemente em $[-r, r]$. Portanto, f é contínua pois r é arbitrário.

Similarmente, para $k \geq k_0$ e $|t| \leq r$ temos

$$|p'_k(t)| = |ka_k t^{k-1}| < \frac{kr^{k-1}}{s^k} = \frac{k}{r} \frac{r^k}{s^k},$$

e consequentemente

$$\sum_{k=0}^{\infty} p'_k$$

converge uniformemente em $[-r, r]$. Pelo Corolário 4.10, temos que f é de classe C^1 e

$$f' = \sum_{k=0}^{\infty} p'_k.$$

Recursivamente, f é C^∞ .

Similar ao que foi feito anteriormente, podemos provar que séries de potências em várias variáveis em \mathbb{R} (e \mathbb{C}) são suaves.

Teorema 4.12. Seja $f : B(0, R) \rightarrow \mathbb{R}$ com $B(0, R)$ sendo uma bola em \mathbb{R}^p com respeito a norma L^∞ , isto é, $|x|_\infty := \max_i |x^i|$.

Considere o polinômio homogêneo de grau k

$$p_k(x) = \sum_{i_1 + \dots + i_p = k} a_{i_1 \dots i_p} (x^1)^{i_1} \cdots (x^p)^{i_p},$$

onde i_1, \dots, i_p são inteiros não negativos, e suponha que

$$f(x) = \sum_{k=0}^{\infty} p_k(x)$$

para todo $x \in B(0, R)$, ou seja, estamos supondo que tal série converge para todo x satisfazendo $|x| < R$.

Para $0 < r < R$, temos que a série acima converge uniformemente em $\overline{B(0, r)}$.

Além disso,

$$\partial_a f(x) = \sum_{k=0}^{\infty} \partial_a p_k(x)$$

para $x \in B(0, R)$.

Dessas duas observações, segue que f é C^∞ .

Observação 4.13. Uma função $f : U \rightarrow \mathbb{R}$, com $U \subset \mathbb{R}^p$ aberto, é analítica real se para cada

$x_0 \in U$ existe $r > 0$ tal que

$$f(x) = \sum_{k=0}^{\infty} p_k(x)$$

em $B(x_0, r) \subset U$, onde

$$p_k(x) = \sum_{i_1+\dots+i_p=k} a_{i_1\dots i_p} (x^1 - x_0^1)^{i_1} \cdots (x^p - x_0^p)^{i_p}.$$

Pelo teorema acima, f é suave.

Demonstração. Fixe $r \in (0, R)$. Tome s tal que $r < s < R$.

Dado ε existe k_0 tal que para $i_1 + \dots + i_p \geq k_0$ temos

$$|a_{i_1\dots i_p} s^{i_1} \cdots s^{i_p}| < 1$$

porque os coeficientes de uma série convergente tendem a zero e no ponto (s, s, \dots, s) a série definindo f converge.

Logo, para $z \in \overline{B(0, r)}$ e $i_1 + \dots + i_p \geq k_0$ temos

$$|a_{i_1\dots i_p} (z^1)^{i_1} \cdots (z^p)^{i_p}| < \left(\frac{r}{s}\right)^{i_1+\dots+i_p},$$

ou seja, para $k \geq k_0$

$$|p_k(z)| \leq \binom{p+k-1}{p-1} \left(\frac{r}{s}\right)^k,$$

pois

$$\{(i_1, \dots, i_p) \in (\mathbb{Z}_{\geq 0})^p \mid i_1 + \dots + i_p = k\}$$

tem $\binom{p+k-1}{p-1}$ elementos (veja: [https://pt.wikipedia.org/wiki/Mál'_todo_bola_traseira](https://pt.wikipedia.org/wiki/M%C3%A1l'_todo_bola_traseira)).

Logo,

$$\sup_{|z| \leq r} |p_k(z)| \leq \binom{p+k-1}{p-1} \left(\frac{r}{s}\right)^k$$

e como a série

$$\sum_{k=0}^{\infty} \binom{p+k-1}{p-1} \left(\frac{r}{s}\right)^k$$

converge pelo teste da razão temos que

$$\sum_{k=0}^{\infty} p_k$$

converge uniformemente para f em $\overline{B(0, r)}$ pelo Corolário 4.7.

Portanto, se f é dada por uma série de potências em uma bola $B(0, R)$, então f é contínua e a série converge uniformemente em bolas compactas $\overline{B(0, r)}$ com $r < R$.

Agora considere $g_a(t) = f(x + te_a)$. Temos que g_a é uma série de potências na variável t . Sabemos do Exemplo 4.11 que a derivada de g_a é dada por

$$g'(t) = \sum_{k=0}^{\infty} \frac{d}{dt} p_k(x + te_a).$$

Fazendo $t = 0$ obtemos

$$\partial_a f(x) = \sum_{k=0}^{\infty} \partial_a p_k(x).$$

Logo, para todo $x \in B(0, r)$ temos que

$$\partial_a f(x) = \sum_{k=0}^{\infty} \partial_a p_k(x),$$

uma série de potências. Logo, $\partial_a f$ é contínua para todo a e assim, f é de classe C^1 .

Pelo mesmo argumento, cada $\partial_a f$ é de classe C^1 , ou seja, f é de classe C^2 . Indutivamente, se prova que f é C^∞ . ■

Exemplo 4.14. A exponencial $\exp : M_d(\mathbb{R}) \rightarrow M_d(\mathbb{R})$ é C^∞ : A exponencial $\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$ tem coordenadas $\exp(A)_{ij}$. A função $A \mapsto \exp(A)_{ij}$ vai de $M_d(\mathbb{R})$ em \mathbb{R} e é dada por uma série de potências em termos das entradas da matriz A . Logo, $\exp(A)_{ij}$ é C^∞ e, consequentemente, \exp é C^∞ .

Exercício 4.15. Mostre que se f é analítica real, então para $x_0 \in U$ temos

$$f(x) = \sum_{k=0}^{\infty} \sum_{i_1+\dots+i_p=k} \frac{1}{i_1! \cdots i_p!} \partial_{i_1} \cdots \partial_{i_p} f(x_0) (x^1 - x_0^1)^{i_1} \cdots (x^p - x_0^p)^{i_p}$$

em uma bola $B(x_0, r) \subset U$ centrada em x_0 , onde a bola é com respeito à norma $|\cdot|_\infty$.

Observação 4.16. Há várias formas padrões de se obter funções analíticas reais ou complexas, uma é usar a fórmula de Taylor e mostrar que o resto tende a zero. Mais precisamente, se temos $f : B(0, R) \rightarrow \mathbb{R}$ de classe C^∞ e para todo x temos que o resto tende a zero, então f é analítica. Por exemplo,

$$\sin(x^1 + x^2) = \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!} (x^1 + x^2)^{2k+1}$$

é analítica real e obtida a partir da fórmula de Taylor. Outra forma é via análise complexa, onde se prova a analiticidade via EDP, resolvendo as equações de Cauchy-Riemann. Essa segunda abordagem é mais simples. Funções analíticas reais surgem também em outras EDPs ao usarmos o teorema de Cauchy-Kovalevskaya, que diz que uma EDP linear com coeficientes analíticos tem solução analítica.

Capítulo 5

Integral de Riemann

5.1 Propriedades básicas da integral de Riemann.

Definição 5.1. O volume de uma caixa $B = [a_1, b_1] \times \cdots \times [a_p, b_p]$ em \mathbb{R}^p é

$$m(B) := (b_1 - a_1) \cdots (b_p - a_p).$$

Um conjunto $A \subset \mathbb{R}^p$ tem medida zero se dado $\varepsilon > 0$ há caixas B_1, B_2, \dots tais que

$$A \subset \bigcup_{i=1}^{\infty} B_i \quad \text{e} \quad \sum_{i=1}^{\infty} m(B_i) < \varepsilon.$$

Proposição 5.2. União de enumeráveis conjuntos de medida zero tem medida zero.

Demonstração. Considere os conjuntos $A_1, A_2, A_3, \dots \subset \mathbb{R}^p$ de medida zero. Dado $\varepsilon > 0$ existem caixas B_{ij} , com $i, j \in \mathbb{Z}_{\geq 1}$, tais que

$$A_i \subset \bigcup_{j=1}^{\infty} B_{ij} \quad \text{e} \quad \sum_{j=1}^{\infty} m(B_{ij}) < \frac{\varepsilon}{2^i}.$$

Assim, temos que o conjunto $A = \bigcup_{i=1}^{\infty} A_i$ é coberto pelas caixas B_{ij} e

$$\sum_{i,j \geq 1} m(B_{ij}) = \sum_{i \geq 1} \sum_{j \geq 1} m(B_{ij}) < \sum_{i \geq 1} \frac{\varepsilon}{2^i} = \varepsilon.$$

Portanto, A tem medida zero. ■

Exemplo 5.3. Todo conjunto enumerável tem medida zero porque se A é enumerável com elementos x_1, x_2, x_3, \dots então dado $\varepsilon > 0$ consideramos caixas B_i com $x_i \in B_i$ tais que $m(B_i) < \frac{\varepsilon}{2^i}$. Assim,

$$A \subset \bigcup_{i \geq 1} B_i \quad \text{e} \quad \sum_{i \geq 1} m(B_i) < \sum_{i \geq 1} \frac{\varepsilon}{2^i} = \varepsilon.$$

Exemplo 5.4. Considere $K \subset \mathbb{R}^p$ compacto e $f : K \rightarrow \mathbb{R}$ contínua. Dado $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x) - f(y)| < \varepsilon$ para $|x - y|_{\infty} < \delta$. Considere um cubo B contendo K (um cubo é uma caixa com todos lados de mesmo comprimento). Divida B em n^p cubos B_i com diâmetro menor que δ tais que $m(B_i) = \frac{m(B)}{n^p}$. Isso pode ser feito dividindo cada lado em n intervalos. A menos de descartar as caixas B_i 's que não intersectam K , podemos supor que cada B_i intersecta K . Para cada caixa B_i escolha $x_i \in K \cap B_i$. Temos que o gráfico de $f|_{B_i}$ está contido em $B_i \times [f(x_i) - \varepsilon, f(x_i) + \varepsilon]$, que tem volume $2\varepsilon m(B_i) = \frac{2\varepsilon m(B)}{n^p}$.

Assim, temos que o gráfico de $f : K \rightarrow \mathbb{R}$ é coberto pelas caixas $B_i \times [f(x_i) - \varepsilon, f(x_i) + \varepsilon]$ e o

número de caixas na cobertura é no máximo n^p . Portanto,

$$\sum_i m(B_i \times [f(x_i) - \varepsilon, f(x_i) + \varepsilon]) \leq 2m(B)\varepsilon.$$

Logo, o gráfico de $f : K \rightarrow \mathbb{R}$ tem medida zero.

Agora considere $g : U \rightarrow \mathbb{R}$ contínua, com $U \subset \mathbb{R}^p$ aberto. O gráfico de g tem medida zero também pois U pode ser escrito como união enumerável de compactos. Por exemplo,

$$K_n = \left\{ x \in \mathbb{R}^p : d(x, \mathbb{R}^p \setminus U) \leq \frac{1}{n} \text{ e } |x| \leq n \right\}.$$

O mesmo vale para funções contínuas com contradomínio em \mathbb{R}^q em vez de \mathbb{R} .

Lema 5.5. Dado $A \subset \mathbb{R}^p$ com medida zero. Dado $\varepsilon > 0$ existem bolas B_1, B_2, \dots com a norma L^∞ (que são caixas com todos lados de mesmo comprimento) que cobrem A e satisfaz

$$\sum_{i=1}^{\infty} m(B_i) < \varepsilon.$$

Chamemos essas caixas de cubos (nomenclatura não padrão).

Demonstração. Considere caixas C_1, C_2, \dots que cobrem A e satisfaz $\sum_i m(C_i) < \frac{\varepsilon}{2}$.

Dada a caixa C_i cubra ela com cubos B_{ij} tal que

$$-m(C_i) + \sum_j m(B_{ij}) < \frac{\varepsilon}{2^{i+1}}.$$

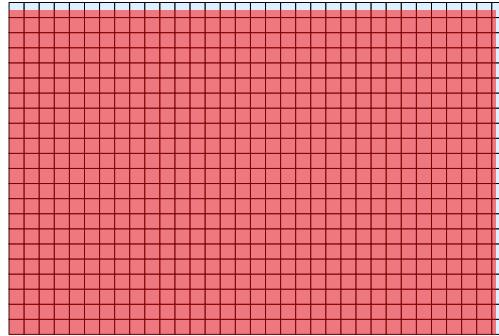


Figura 5.1: Aproximando caixas por cubos.

Temos que os cubos B_{ij} cobrem A e

$$\sum_{i,j} m(B_{ij}) < \varepsilon.$$

Proposição 5.6. Seja $f : U \rightarrow \mathbb{R}^q$ Lipschitz, onde $U \subset \mathbb{R}^p$ é aberto. Se A tem medida zero, então $f(A)$ tem medida zero. ■

Demonstração. Existe uma constante $c > 0$ tal que $|f(x) - f(y)|_\infty \leq c|x - y|_\infty$

Seja B um cubo de lado r em U e centro x . Temos que $f(B)$ está contido no cubo B' de centro $f(x)$ e lado rc .

Agora basta usar o Lema 5.5. ■

Note que todo aberto de \mathbb{R}^p pode ser escrito como união enumerável de bolas abertas com centro centro em \mathbb{Q}^p e raios da forma $\frac{1}{n}$, n inteiro positivo. Assim, todo aberto é união de enumerável de bolas. Como corolário do teorema acima temos o seguinte resultado.

Lema 5.7. Seja $M \subset \mathbb{R}^p$ subconjunto. Seja $\{U_\lambda\}_{\lambda \in \Lambda}$ uma cobertura de M . Existem índices $\lambda_1, \lambda_2, \dots$ em Λ tais que

$$M \subset \bigcup_{i=1}^{\infty} U_{\lambda_i}.$$

Demonastração. Considere a família \mathcal{F} das bolas de centro em \mathbb{Q}^p e raios $\frac{1}{n}$. O conjunto dessas bolas é enumerável. Cada U_λ é união de uma subfamília \mathcal{F}_λ de \mathcal{F} . Temos que $\cup_{\lambda \in \Lambda} \mathcal{F}_\lambda$ é enumerável por ser subconjunto de \mathcal{F} .

Sejam B_1, B_2, \dots os elementos de $\cup_{\lambda \in \Lambda} \mathcal{F}_\lambda$. Temos que as bolas B_i 's cobrem M .

Dado $i \in \mathbb{Z}_{\geq 1}$ existe $\lambda_i \in \Lambda$ tal que $B_i \subset U_{\lambda_i}$. Assim,

$$M \subset \bigcup_{i=1}^{\infty} B_i \subset \bigcup_{i=1}^{\infty} U_{\lambda_i}.$$

■

Corolário 5.8. Se $f : U \rightarrow \mathbb{R}^q$ é localmente Lipschitz, então $f(A)$ tem medida zero quando $A \subset U$ tem medida zero.

Em particular, mapas de classe C^1 satisfazem tal propriedade por serem localmente Lipschitz.

Corolário 5.9. Se $M \subset \mathbb{R}^p$ é uma variedade de dimensão n com $n < p$, então M tem medida zero.

Demonastração. Pelo Lema 5.7 podemos cobrir M com uma família enumerável de cartas.

Considere uma família enumerável $\varphi_i : U_i \rightarrow V_i$ de cartas tal que os abertos U_i 's cobrem M . Temos que

$$M = \bigcup_{i=1}^{\infty} \varphi_i^{-1}(V_i \cap (\mathbb{R}^n \times \{0\})),$$

uma união enumerável de conjuntos de medida zero e, portanto, M tem medida zero. De fato, cada conjunto $\varphi_i^{-1}(V_i \cap (\mathbb{R}^n \times \{0\}))$ tem medida zero porque $V_i \cap (\mathbb{R}^n \times \{0\}) \subset \mathbb{R}^p$ tem medida zero (por ser gráfico de uma função contínua) e φ_i^{-1} é localmente Lipschitz. ■

Definição 5.10. Seja $B \subset \mathbb{R}^p$ uma caixa. Uma partição \mathcal{P} de B é uma família finita de caixas contidas em B tal que

$$B = \bigcup_{C \in \mathcal{P}} C$$

e os interiores de C e C' não se intersectam para $C \neq C'$.

Um refinamento de \mathcal{P} é uma partição \mathcal{P}' tal que cada caixa de \mathcal{P}' está contida numa caixa de \mathcal{P} .

Note que duas partições \mathcal{P}_1 e \mathcal{P}_2 admite um refinamento comum \mathcal{Q} feito pelas caixas $B_1 \cap B_2$ com $B_i \in \mathcal{P}_i$ tais que $\overset{\circ}{B}_1 \cap \overset{\circ}{B}_2 \neq \emptyset$.

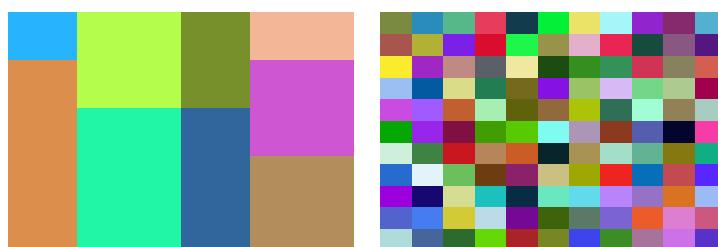


Figura 5.2: Exemplos de partições.

Exemplo 5.11. Dividindo os lados de B em n intervalos de mesmo comprimento obtemos n^p caixas B_i tal que $m(B_i) = \frac{m(B)}{n^p}$.

Definição 5.12. Seja $B \subset \mathbb{R}^p$ uma caixa e seja $f : B \rightarrow \mathbb{R}$ uma função. Dada uma partição \mathcal{P} de B temos

$$S(f, \mathcal{P}) = \sum_{C \in \mathcal{P}} \sup_C f(z)m(C),$$

$$s(f, \mathcal{P}) = \sum_{C \in \mathcal{P}} \inf_C f(z)m(C).$$

A integral superior de f é

$$\overline{\int}_B f(x)dx := \inf_{\mathcal{P}} S(f, \mathcal{P})$$

e a integral inferior de f é

$$\underline{\int}_B f(x)dx := \sup_{\mathcal{P}} s(f, \mathcal{P}).$$

Exercício 5.13. Se \mathcal{Q} é refinamento de \mathcal{P} , então

$$S(f, \mathcal{P}) \geq S(f, \mathcal{Q}) \quad \text{e} \quad s(f, \mathcal{P}) \leq s(f, \mathcal{Q}).$$

Proposição 5.14. Sejam $f, g : B \rightarrow \mathbb{R}$ limitadas e $\alpha \in \mathbb{R}$.

1. $\underline{\int}_B f(x)dx \leq \overline{\int}_B f(x)dx$;
2. $\overline{\int}_B (f(x) + g(x))dx \geq \overline{\int}_B f(x)dx + \overline{\int}_B g(x)dx$;
3. $\underline{\int}_B (f(x) + g(x))dx \leq \underline{\int}_B f(x)dx + \underline{\int}_B g(x)dx$;
4. para $\alpha \geq 0$ temos $\overline{\int}_B \alpha f(x)dx = \alpha \overline{\int}_B f(x)dx$ e $\underline{\int}_B \alpha f(x)dx = \alpha \underline{\int}_B f(x)dx$;
5. para $\alpha \leq 0$ temos $\overline{\int}_B \alpha f(x)dx = \alpha \underline{\int}_B f(x)dx$.

Demonstração. 1. Segue de $s(f, \mathcal{Q}) \leq S(f, \mathcal{Q})$.

2. Considere uma partição \mathcal{Q} .

$$S(f + g, \mathcal{Q}) = \sum_{C \in \mathcal{Q}} \sup_C f(z)m(C) + \sum_{C \in \mathcal{Q}} \sup_C g(z)m(C) \geq \overline{\int}_B f(x)dx + \overline{\int}_B g(x)dx.$$

Portanto,

$$\overline{\int}_B (f(x) + g(x))dx = \inf_{\mathcal{Q}} S(f + g, \mathcal{Q}) \geq \overline{\int}_B f(x)dx + \overline{\int}_B g(x)dx.$$

3. Similar ao caso anterior.

4. A partir de $S(\alpha f, \mathcal{Q}) = \alpha S(f, \mathcal{Q})$ obtemos

$$\begin{aligned} \overline{\int}_B \alpha f(x)dx &= \inf_{\mathcal{Q}} S(\alpha f, \mathcal{Q}) \\ &= \inf_{\mathcal{Q}} \alpha S(f, \mathcal{Q}) \\ &= \alpha \inf_{\mathcal{Q}} S(f, \mathcal{Q}) \\ &= \alpha \overline{\int}_B f(x)dx. \end{aligned}$$

5. Como $S(\alpha f, \mathcal{Q}) = \alpha s(f, \mathcal{Q})$, obtemos

$$\begin{aligned} \overline{\int}_B \alpha f(x)dx &= \inf_{\mathcal{Q}} S(\alpha f, \mathcal{Q}) \\ &= \inf_{\mathcal{Q}} \alpha s(f, \mathcal{Q}) \end{aligned}$$

$$\begin{aligned}
&= \alpha \sup_{\mathcal{Q}} s(f, \mathcal{Q}) \\
&= \alpha \int_{\underline{B}} f(x) dx.
\end{aligned}$$

■

Lema 5.15. Seja $A \subset \mathbb{R}^p$ um conjunto de medida zero. Dado $\varepsilon > 0$ existem caixas B_1, B_2, B_3, \dots tais que

$$A \subset \bigcup_i \overset{\circ}{B}_i$$

$$\sum_i m(B_i) < \varepsilon.$$

Demonstração. Considere caixas C_i que cobrem A tais que

$$\sum_i m(C_i) < \frac{\varepsilon}{2}.$$

Para cada i considere uma caixa B_i tal que $C_i \subset \overset{\circ}{B}_i$ e $m(B_i) - m(C_i) < \frac{\varepsilon}{2^{i+1}}$. Temos que os abertos $\overset{\circ}{B}_i$ cobrem A e

$$\sum_i m(B_i) < \varepsilon.$$

■

Lema 5.16. Considere a caixa fechada B e a função limitada $f : B \rightarrow \mathbb{R}$. A oscilação de f em x é o limite

$$\omega(f, x) := \lim_{\delta \rightarrow 0^+} \left(\sup_{B(x, \delta)} f(z) - \inf_{B(x, \delta)} f(z) \right).$$

Temos que f é contínua em x se, e só se, $\omega(f, x) = 0$ e $\{x \in B : \omega(x, f) \geq c\}$ é fechado em \mathbb{R}^p para todo $c > 0$.

Observação: O limite acima existe porque a função

$$\delta \mapsto \sup_{B(x, \delta)} f(z) - \inf_{B(x, \delta)} f(z)$$

é decrescente.

Demonstração. Se f é contínua em x , então dado $\varepsilon > 0$ existe $\delta > 0$ tal que

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

$$f(x) > f(z) - \varepsilon \quad \text{e} \quad f(x) < f(z) + \varepsilon.$$

Assim,

$$f(x) \geq \sup_{B(x, \delta)} f(z) - \varepsilon,$$

$$f(x) \leq \inf_{B(x, \delta)} f(z) + \varepsilon,$$

e, consequentemente,

$$\sup_{B(x, \delta)} f(z) - \inf_{B(x, \delta)} f(z) \leq 2\varepsilon.$$

Logo,

$$\lim_{\delta \rightarrow 0^+} \left(\sup_{B(x, \delta)} f(z) - \inf_{B(x, \delta)} f(z) \right) = 0.$$

Para mostrar que $\omega(f, x) = 0$ implica que f é contínua em x segue de argumento similar.

Agora provemos que $\{x \in B : \omega(f, x) \geq c\}$ é fechado. Considere $U = \{x \in B : \omega(f, x) < c\}$. Tome $x \in U$. Existe $\delta > 0$ tal que

$$\sup_{B(x, \delta)} f(z) - \inf_{B(x, \delta)} f(z) < c.$$

Dado $x' \in B(x, \frac{\delta}{2})$. Como $B(x', \frac{\delta}{2}) \subset B(x, \delta)$ obtemos

$$\sup_{B(x', \frac{\delta}{2})} f(z) < \sup_{B(x, \delta)} f(z),$$

$$\inf_{B(x', \frac{\delta}{2})} f(z) > \inf_{B(x, \delta)} f(z),$$

e, consequentemente,

$$\omega(f, x') \leq \sup_{B(x', \frac{\delta}{2})} f(z) - \inf_{B(x', \frac{\delta}{2})} f(z) < c.$$

Portanto, $B(x, \frac{\delta}{2}) \subset U$. Assim, U é aberto em B e $\{x \in B : \omega(f, x) \geq c\}$ é fechado. \blacksquare

Definição 5.17. Seja $B \subset \mathbb{R}^p$ uma caixa e seja $f : B \rightarrow \mathbb{R}$ uma função. A função f é (Riemann) integrável se f é limitada e $\{x \in B : f \text{ é descontínua em } x\}$ tem medida zero.

Teorema 5.18. A função $f : B \rightarrow \mathbb{R}$ é integrável se, e só se,

$$\underline{\int}_B f(x) dx = \overline{\int}_B f(x) dx.$$

Demonstração.

Parte 1. Mostremos que se f é integrável, então $\underline{\int}_B f(x) dx = \overline{\int}_B f(x) dx$.

Seja L uma constante tal que $|f(x)| \leq L$ para todo $x \in B$.

Considere $\varepsilon > 0$. Temos que

$$K := \{x \in B : \omega(f, x) \geq \varepsilon\}$$

é compacto. Como

$$K \subset \{x \in B : f \text{ é descontínua em } x\}$$

temos que K tem medida zero. Pelo Lema 5.15, existe uma família enumerável de caixas $\{B_i\}_{i=1}^\infty$ tal que

$$K \subset \bigcup_{i=1}^\infty \overset{\circ}{B}_i \quad \text{e} \quad \sum_{i=1}^\infty m(B_i) < \varepsilon.$$

Como K é compacto, existe N tal que

$$K \subset \bigcup_{i=1}^N \overset{\circ}{B}_i \quad \text{e} \quad \sum_{i=1}^N m(B_i) < \varepsilon.$$

Dado $x \in B \setminus \bigcup_{i=1}^N \overset{\circ}{B}_i$ temos que $\omega(x) < \varepsilon$, ou seja, existe $r_x > 0$ tal que

$$\sup_{B(x, r_x)} f(z) - \inf_{B(x, r_x)} f(z) < \varepsilon.$$

Temos que as bolas $B(x, r_x)$ cobrem o compacto $B \setminus \bigcup_{i=1}^N \overset{\circ}{B}_i$. Pelo Lema 1.140 existe um número de Lebesgue $\delta > 0$ tal que cada bola $B(x, \delta)$ está contida em algum $B(x', r_{x'})$. Assim,

$$\sup_{B(x, \delta)} f(z) - \inf_{B(x, \delta)} f(z) \leq \sup_{B(x', r_{x'})} f(z) - \inf_{B(x', r_{x'})} f(z) < \varepsilon.$$

Considere uma partição \mathcal{Q} de B forma que todas suas caixas tenham diâmetro menor que menor que $\frac{\delta}{2}$ e parte de suas caixas formam $\bigcup_{i=1}^N B_i$. Podemos escrever $\mathcal{Q} = \mathcal{Q}_1 \cup \mathcal{Q}_2$,

onde \mathcal{Q}_1 é formada pelas caixas C satisfazendo $C \subset \cup_{i=1}^N B_i$ e $\mathcal{Q}_2 = \mathcal{Q} \setminus \mathcal{Q}_1$. Assim, temos que para $C \in \mathcal{Q}_1$

$$\sup_C f(z) - \inf_C f(z) \leq 2L$$

e para $C \in \mathcal{Q}_2$ temos $C \cap K = \emptyset$ e, consequentemente,

$$\sup_C f(z) - \inf_C f(z) < \varepsilon.$$

Portanto,

$$\begin{aligned} S(f, \mathcal{Q}) - s(f, \mathcal{Q}) &= \sum_{C \in \mathcal{Q}} \left(\sup_C f(z) - \inf_C f(z) \right) m(C) \\ &= \sum_{C \in \mathcal{Q}_1} \left(\sup_C f(z) - \inf_C f(z) \right) m(C) + \sum_{C \in \mathcal{Q}_2} \left(\sup_C f(z) - \inf_C f(z) \right) m(C) \\ &\leq 2L \sum_{C \in \mathcal{Q}_1} m(C) + \varepsilon \sum_{C \in \mathcal{Q}_2} m(C) \\ &= 2L \sum_{i=1}^n m(B_i) + \varepsilon \sum_{C \in \mathcal{Q}_2} m(C) \\ &\leq 2L\varepsilon + m(B)\varepsilon \\ &= (2L + m(B))\varepsilon. \end{aligned}$$

Portanto,

$$0 \leq \overline{\int}_B f(x) dx - \underline{\int}_B f(x) dx \leq S(f, \mathcal{Q}) - s(f, \mathcal{Q}) \leq (2L + m(B))\varepsilon$$

e, fazendo $\varepsilon \rightarrow 0^+$, obtemos

$$\overline{\int}_B f(x) dx = \underline{\int}_B f(x) dx.$$

Parte 2. Mostremos que se as integrais superiores e inferiores coincidem, então a função é integrável. Considere $K_n = \{x \in B : \omega(f, x) \geq \frac{1}{n}\}$.

Dado $\varepsilon > 0$ existe uma partição \mathcal{P} de B tal que

$$S(f, \mathcal{P}) - s(f, \mathcal{P}) < \frac{\varepsilon}{n},$$

ou seja,

$$\sum_{C \in \mathcal{P}} \left(\sup_C f(z) - \inf_C f(z) \right) m(C) < \frac{\varepsilon}{n}.$$

Seja \mathcal{A} a família de caixas $C \in \mathcal{P}$ satisfazendo $\dot{C} \cap K_n \neq \emptyset$.

Para $x \in \dot{C} \cap K_n$ temos

$$\sup_{z \in C} f(z) - \inf_{z \in C} f(z) \geq \omega(x) \geq \frac{1}{n}.$$

$$\begin{aligned} \sum_{C \in \mathcal{A}} \frac{m(C)}{n} &\leq \sum_{C \in \mathcal{A}} \left(\sup_{z \in C} f(z) - \inf_{z \in C} f(z) \right) m(C) \\ &\leq \sum_{C \in \mathcal{P}} \left(\sup_{z \in C} f(z) - \inf_{z \in C} f(z) \right) m(C) \\ &< \frac{\varepsilon}{n}. \end{aligned}$$

Portanto, as caixas de \mathcal{A} cobrem

$$K_n \setminus \bigcup_{D \in \mathcal{P}} \partial D$$

e $\sum_{C \in \mathcal{A}} m(C) < \varepsilon$, onde ∂D é a fronteira da caixa D .

Por outro lado, $\bigcup_{D \in \mathcal{P}} \partial D$ tem medida zero, ou seja, há caixas $B_i \subset B$, $i \in \mathbb{Z}_{>0}$, tais que $\bigcup_{D \in \mathcal{P}} \partial D \subset \bigcup_i B_i$ e $\sum_i m(B_i) < \varepsilon$.

Assim, o conjunto K_n tem medida zero.

Como

$$\{x \in B : f \text{ é descontínua em } x\} = \bigcup_{n=1}^{\infty} K_n,$$

temos que esse conjunto tem medida zero e, consequentemente, f é integrável. ■

Definição 5.19. Considere a caixa $B \subset \mathbb{R}^p$. O espaço das funções $B \rightarrow \mathbb{R}$ integráveis será denotado por $\mathcal{R}(B)$. Para $f : B \rightarrow \mathbb{R}$ integrável denotaremos

$$\overline{\int}_B f(x) dx = \int_{-B} f(x) dx$$

por

$$\int_B f(x) dx.$$

Proposição 5.20. Considere $f, g \in \mathcal{R}(B)$ e $\alpha \in \mathbb{R}$. Temos que

1. αf , $f + g$ e fg são integráveis;
2. Se $\phi : I \rightarrow \mathbb{R}$ é função contínua, com I intervalo, e se $f(B) \subset I$, então $\phi \circ f \in \mathcal{R}(B)$;
3. $\int_B (f(x) + g(x)) dx = \int_B f(x) dx + \int_B g(x) dx$;
4. $\int_B \alpha f(x) dx = \alpha \int_B f(x) dx$.

Portanto, $\mathcal{R}(B)$ é uma álgebra real e $\int_B : \mathcal{R}(B) \rightarrow \mathbb{R}$ é um funcional linear.

Demonstração. Verifiquemos o primeiro item. Seja Z_f o conjunto de pontos de B onde f é descontínua. Assim, os conjuntos $Z_{\alpha f}$, Z_{f+g} e Z_{fg} tem medida zero porque

$$Z_{\alpha f} \subset Z_f, \quad Z_{f+g} \subset Z_f \cup Z_g, \quad Z_{fg} \subset Z_f \cup Z_g.$$

Portanto, $\alpha f, f + g, fg \in \mathcal{R}(B)$.

O segundo item segue de $Z_{\phi \circ f} \subset Z_f$.

Os itens 3 e 4 seguem da Proposição 5.14. ■

Exercício 5.21. Seja $f : B \rightarrow \mathbb{R}$ integrável. Mostre que

$$\left| \int_B f(x) dx \right| \leq \int_B |f(x)| dx.$$

Teorema 5.22. (Teorema de Fubini) Considere as caixas $C \subset \mathbb{R}^p$ e $D \subset \mathbb{R}^q$ caixas, a função integrável $f : C \times D \rightarrow \mathbb{R}$, e as coordenadas (x, y) de $\mathbb{R}^p \times \mathbb{R}^q$ com $x \in \mathbb{R}^p$ e $y \in \mathbb{R}^q$. Suponha que para cada x a função $y \mapsto f(x, y)$ é integrável.

Então

$$x \mapsto \int_D f(x, y) dy$$

é integrável e

$$\int_{C \times D} f(x, y) dx dy = \int_C \left(\int_D f(x, y) dy \right) dx.$$

Demonstração. Defina $g(x) = \int_D f(x, y) dy$ e considere $\varepsilon > 0$.

Considere as partições $\{C_i\}_i$ de C e $\{D_j\}_j$ de D tais que

$$\sum_{i,j} \left(\sup_{C_i \times D_j} f(x, y) - \inf_{C_i \times D_j} f(x, y) \right) m(C_i \times D_j) < \varepsilon.$$

Para $z \in C_i$ temos

$$\sum_j \sup_{C_i \times D_j} f(x, y)m(D_j) \geq \sum_j \sup_{D_j} f(z, y)m(D_j) \geq \sum_j \int_{D_j} f(z, y)dy = g(z),$$

$$\sum_j \sup_{C_i \times D_j} f(x, y)m(D_j) \geq \sup_{C_i} g(x).$$

Similarmente,

$$\sum_j \inf_{C_i \times D_j} f(x, y)m(D_j) \leq \inf_{C_i} g(x).$$

Logo,

$$\sum_i \left(\sup_{C_i} g(x) - \inf_{C_i} g(x) \right) m(C_i) \leq \sum_{i,j} \left(\sup_{C_i \times D_j} f(x, y) - \inf_{C_i \times D_j} f(x, y) \right) m(C_i \times D_j) < \varepsilon,$$

e, consequentemente,

$$\int_C \left(\int_D f(x, y)dy \right) dx = \int_{C \times D} f(x, y)dxdy.$$

■

Similarmente, se para todo y a função $x \mapsto f(x, y)$ é integrável, então

$$y \mapsto \int_C f(x, y)dx$$

é integrável e

$$\int_{C \times D} f(x, y)dxdy = \int_D \left(\int_C f(x, y)dx \right) dy.$$

Se $B = [a^1, b^1] \times \cdots \times [a^p, b^p]$ e $f : B \rightarrow \mathbb{R}$ é contínua, então

$$\int_B f(x^1, \dots, x^p)dx^1 \cdots dx^p = \int_{a^1}^{b^1} \cdots \left(\int_{a^{p-1}}^{b^{p-1}} \left(\int_{a^p}^{b^p} f(x^1, \dots, x^p)dx^p \right) dx^{p-1} \right) \cdots dx^p$$

e podemos permutar a ordem das integrais.

5.2 Volume e Mudança de variável

Seja $f : A \rightarrow \mathbb{R}$ uma função limitada e com domínio limitado. Estenda f para \mathbb{R}^p definindo $f = 0$ em $\mathbb{R}^p \setminus A$. Dizemos que f é integrável se

$$\{x \in \mathbb{R}^p : f \text{ é descontínua em } x\}$$

tem medida zero.

Se f é Riemann integrável, então definimos

$$\int_{\mathbb{R}^p} f(x)dx := \int_B f(x)dx.$$

Se B' é outra caixa contendo A , então

$$\int_B f(x)dx = \int_{B \cap B'} f(x)dx = \int_{B'} f(x)dx.$$

Dado um subconjunto $A \subset \mathbb{R}^p$ a função característica de A é

$$\chi_A(x) := \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

Definição 5.23. Um subconjunto $A \subset \mathbb{R}^p$ limitado é Jordan mensurável se χ_A é integrável. A medida/volume de A é

$$m(A) := \int_{\mathbb{R}^p} \chi_A(x) dx.$$

Observação 5.24. Segue da definição de integrável que se A_1, A_2 são Jordan mensuráveis, então $A_1 \cap A_2$ e $A_1 \cup A_2$ são Jordan mensuráveis porque

$$\chi_{A_1 \cap A_2} = \chi_{A_1} \chi_{A_2} \quad \text{e} \quad \chi_{A_1 \cup A_2} = \chi_{A_1} + \chi_{A_2} - \chi_{A_1 \cap A_2}$$

são integráveis.

Exemplo 5.25. Considere o simplexo padrão $\Delta_p = \{(x^1, \dots, x^p) \in [0, 1]^p : x^1 + \dots + x^p \geq 1\}$. Computemos a volume de Δ_p . Note que

$$\chi_{\Delta_p}(x^1, \dots, x^p) = \chi_{\Delta_{p-1}}(x^1, \dots, x^{p-1}) \chi_{[0, 1-x^1-\dots-x^{p-1}]}(x^p).$$

Assim, temos

$$\begin{aligned} & \int_{[0,1]^p} \chi_{\Delta_p}(x^1, \dots, x^p) dx^1 \cdots dx^p = \\ & \int_{[0,1]^{p-1}} \left(\int_0^1 \chi_{\Delta_p}(x^1, \dots, x^p) dx^p \right) dx^1 \cdots dx^{p-1} = \\ & \int_{[0,1]^{p-1}} \left(\int_0^1 \chi_{\Delta_{p-1}}(x^1, \dots, x^{p-1}) \chi_{[0, 1-x^1-\dots-x^{p-1}]}(x^p) dx^p \right) dx^1 \cdots dx^{p-1} = \\ & \int_{[0,1]^{p-1}} (1 - x^1 - \cdots - x^{p-1}) \chi_{\Delta_{p-1}}(x^1, \dots, x^{p-1}) dx^1 \cdots dx^{p-1}. \end{aligned}$$

Similarmente,

$$\begin{aligned} & \int_{[0,1]^{p-1}} (1 - x^1 - \cdots - x^{p-1}) \chi_{\Delta_{p-1}}(x^1, \dots, x^{p-1}) dx^1 \cdots dx^{p-1} = \\ & \int_{[0,1]^{p-2}} \left(\int_0^{1-x^1-\dots-x^{p-2}} (1 - x^1 - \cdots - x^{p-1}) \chi_{\Delta_{p-2}}(x^1, \dots, x^{p-2}) dx^{p-1} \right) dx^1 \cdots dx^{p-2} = \\ & \int_{[0,1]^{p-2}} \frac{(1 - x^1 - \cdots - x^{p-2})^2}{2!} \chi_{\Delta_{p-2}}(x^1, \dots, x^{p-2}) dx^1 \cdots dx^{p-2}. \end{aligned}$$

Recursivamente,

$$\begin{aligned} & \int_{[0,1]^p} \chi_{\Delta_p}(x^1, \dots, x^p) dx^1 \cdots dx^p = \\ & \int_{[0,1]^k} \frac{(1 - x^1 - \cdots - x^k)^{(p-k)}}{(p-k)!} \chi_{\Delta_k}(x^1, \dots, x^k) dx^1 \cdots dx^k. \end{aligned}$$

Daí, segue que

$$\int_{[0,1]^p} \chi_{\Delta_p}(x^1, \dots, x^p) dx^1 \cdots dx^p = \int_0^1 \frac{(1 - x^1)^{p-1}}{(p-1)!} \chi_{\Delta_1}(x^1) dx^1 = \frac{1}{p!}.$$

Portanto, o volume de Δ_p .

Exercício 5.26. Seja $B \subset \mathbb{R}^p$ uma caixa. Se $f : B \rightarrow \mathbb{R}$ é integrável e

$$\{x \in B : f(x) \neq 0\}$$

tem medida zero, então

$$\int_B f(x)dx = 0.$$

Exercício 5.27. Se A_1, A_2 são Jordan mensuráveis e $A_1 \cap A_2$ tem medida zero, então $m(A_1 \cup A_2) = m(A_1) + m(A_2)$.

5.2.1 Volume de paralelogramos

Considere \mathbb{R}^p e p -vetores v_1, \dots, v_p . Os vetores v_j tem coordenadas v_j^i .

$$v_j = \begin{bmatrix} v_j^1 \\ v_j^2 \\ \vdots \\ v_j^p \end{bmatrix}.$$

O paralelogramo dado por v_1, \dots, v_p é o conjunto

$$A = \{t_1 v_1 + \dots + t_p v_p \in \mathbb{R}^p : t_1, \dots, t_p \in [0, 1]\}.$$

Mostraremos que $m(A)$ é o valor absoluto do determinante

$$\det \begin{bmatrix} v_1^1 & v_2^1 & \cdots & v_p^1 \\ v_1^2 & v_2^2 & \cdots & v_p^2 \\ \vdots & \vdots & \ddots & \vdots \\ v_1^p & v_2^p & \cdots & v_p^p \end{bmatrix}.$$

Lema 5.28. Seja

$$T = (T_j^i)$$

uma matriz $p \times p$ invertível. Temos que T se escreve como produto de transformações lineares da forma

1. $E_1(x^1, \dots, x^p) = (x^1, \dots, x^p) + \lambda x^j e_k, \quad \lambda \in \mathbb{R}, \quad j \neq k,$
2. $E_2(x^1, \dots, x^i, \dots, x^j, \dots, x^p) = (x^1, \dots, x^j, \dots, x^i, \dots, x^p), \quad i < j,$
3. $E_3(x^1, \dots, x^i, \dots, x^p) = (x^1, \dots, \lambda x^i, \dots, x^p), \quad \lambda \neq 0.$

O primeiro tipo de transformação linear tem determinante 1. O segundo tipo de transformação linear tem determinante -1. O terceiro tipo de transformação linear tem determinante λ .

Demonstração. Se E_1 é uma matriz do primeiro tipo então TE_1 é a matriz T trocando a coluna

$$\begin{bmatrix} t_j^1 \\ t_j^2 \\ \vdots \\ t_j^p \end{bmatrix}$$

por

$$\begin{bmatrix} t_j^1 + \lambda t_k^1 \\ t_j^2 + \lambda t_k^2 \\ \vdots \\ t_j^p + \lambda t_k^p \end{bmatrix}.$$

Similarmente, $E_1 T$ faz o mesmo com as linhas.

Se E_2 é uma matriz do segundo tipo então TE_2 permuta as colunas i, j e $E_2 T$ permuta as linhas i, j .

Já TE_3 multiplica a coluna i por λ . Similarmente, E_3T multiplica a linha i por λ .

O restante do teorema segue da eliminação de Gauss. Basta aplicar matrizes dos três tipos até obtermos a identidade. ■

Proposição 5.29. Se T é invertível, seja A um paralelogramo gerado por v_1, \dots, v_p .

$$\int_{\mathbb{R}^p} \chi_A(T^{-1}(y)) dy = \int_{\mathbb{R}^p} \chi_A(x) |\det(T)| dx.$$

Demonstração. Usaremos as transformações lineares de a cordo com o Lema 5.28. O resultado é trivialmente verdadeiro para transformações do segundo tipo. Sem perda de generalidade, para transformações do primeiro tipo temos $T(x) = (\lambda x^1, x^2, x^3, \dots, x^p)$ com $\lambda \neq 0$.

$$\begin{aligned} \int_{\mathbb{R}^{p-1}} \left(\int_{\mathbb{R}} \chi_A(\lambda^{-1}x^1, x^2, \dots, x^p) dy^1 \right) dx^2 \cdots dx^p &= \int_{\mathbb{R}^{p-1}} \left(\int_{\mathbb{R}} |\lambda| \chi_A(y^1, x^2, \dots, x^p) dy^1 \right) dx^2 \cdots dx^p \\ &= \int_{\mathbb{R}^{p-1}} \left(\int_{\mathbb{R}} |\lambda| \chi_A(y^1, x^2, \dots, x^p) dy^1 \right) dx^2 \cdots dx^p \\ &= \int_{\mathbb{R}^p} |\lambda| \chi_A(y^1, x^2, \dots, x^p) dy^1 dx^2 \cdots dx^p, \end{aligned}$$

onde fizemos $y^1 = \lambda^{-1}x^1$. Note que usamos o seguinte: para integrais em uma variável com $\lambda > 0$ temos

$$\int_{-\infty}^{\infty} \phi(\lambda^{-1}x) dx = \int_{-\infty}^{\infty} \lambda \phi(y) dy = \int_{-\infty}^{\infty} |\lambda| \phi(y) dy$$

e para $\lambda < 0$ temos

$$\int_{-\infty}^{\infty} \phi(\lambda^{-1}x) dx = \int_{\infty}^{-\infty} \lambda \phi(y) dy = \int_{-\infty}^{\infty} (-\lambda) \phi(y) dy = \int_{-\infty}^{\infty} |\lambda| \phi(y) dy.$$

Sem perda de simplicidade, considere agora $T(x) = (x^1 + \lambda x^2, x^2, \dots, x^p)$. Temos

$$\begin{aligned} \int_{\mathbb{R}^{p-1}} \left(\int_{\mathbb{R}} \chi_A(x^1 + \lambda x^2, x^2, \dots, x^p) dy^1 \right) dx^2 \cdots dx^p &= \int_{\mathbb{R}^{p-1}} \left(\int_{\mathbb{R}} \chi_A(y^1, x^2, \dots, x^p) dy^1 \right) dx^2 \cdots dx^p \\ &= \int_{\mathbb{R}^{p-1}} \left(\int_{\mathbb{R}} \chi_A(y^1, x^2, \dots, x^p) dy^1 \right) dx^2 \cdots dx^p \\ &= \int_{\mathbb{R}^p} \chi_A(y^1, x^2, \dots, x^p) dy^1 dx^2 \cdots dx^p, \end{aligned}$$

onde fizemos a troca de variável $y^1 = x^1 + \lambda x^2$. Agora use o Lema 5.28. ■

Como $\chi_A \circ T^{-1} = \chi_{T(A)}$ temos

$$m(T(A)) = |\det(T)| m(A)$$

onde A é um paralelogramo. Se v_1, v_2, \dots, v_p são linearmente dependentes, então estão contidos em um subespaço linear de dimensão $p-1$, ou seja, o paralelogramo gerado por tais vetores tem volume 0. Se v_1, v_2, \dots, v_p são linearmente independentes, então definindo $A = [0, 1]^p$ e $T = (v_j^i)$ então $T(A)$ é o paralelogramo definido por v_1, \dots, v_p e $m(T(A)) = |\det(T)|$.

Portanto, o volume de

$$\{\alpha^1 v_1 + \cdots + \alpha^p v_p \in \mathbb{R}^p : \alpha^1, \dots, \alpha^p \in [0, 1]\}$$

é $|\det(v_j^i)|$.

5.3 Teorema de mudança de variável

Lema 5.30. Seja $\phi : U \rightarrow V$ difeomorfismo com $U, V \subset \mathbb{R}^p$ aberto. Considere uma caixa $B \subset U$. Temos que

$$m(\phi(B)) = \int_B |\det(d\phi_x)| dx$$

Demonstração.

Parte 1. Suponha que B é um cubo. Considere a norma $|\cdot|_\infty$ em \mathbb{R}^p e a norma de operador

$$|A| = \max_{|x|_\infty=1} |Ax|_\infty.$$

Dado $\varepsilon \in (0, 1)$ existe $\delta > 0$ tal que

$$|d\phi_{x'}^{-1} d\phi_x - \text{Id}| < \varepsilon \quad \text{e} \quad |\det(d\phi_{x'}) - \det(d\phi_x)| < \varepsilon$$

para $|x - x'|_\infty < \delta$ com $x, x' \in B$.

Considere uma partição B_1, \dots, B_n de B feita de cubos com diâmetro menor que δ . Seja $x_k \in B_k$ o centro do cubo B_k .

Note que

$$\phi(x) - \phi(x_k) = \int_0^1 d\phi_{x_k+t(x-x_k)}(x - x_k) dt,$$

$$\phi(x) - \phi(x_k) - d\phi_{x_k}(x - x_k) = \int_0^1 (d\phi_{x_k+t(x-x_k)} - d\phi_{x_k})(x - x_k) dt,$$

$$d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k)) - (x - x_k) = \int_0^1 (d\phi_{x_k}^{-1} d\phi_{x_k+t(x-x_k)} - \text{Id})(x - x_k) dt.$$

Podemos escrever $B_k = \overline{B(x_k, r_k)}$ onde a bola é com respeito norma $|\cdot|_\infty$ e $r_k < \delta$.

Note que

$$|d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k)) - (x - x_k)|_\infty \leq \varepsilon|x - x_k|_\infty,$$

$$|d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k))|_\infty - |(x - x_k)|_\infty \leq \varepsilon|x - x_k|_\infty,$$

$$|d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k))|_\infty \leq (1 + \varepsilon)|x - x_k|_\infty$$

e

$$|d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k)) - (x - x_k)|_\infty \leq \varepsilon|x - x_k|_\infty,$$

$$-|d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k))|_\infty + |(x - x_k)|_\infty \leq \varepsilon|x - x_k|_\infty,$$

$$(1 - \varepsilon)|x - x_k|_\infty \leq |d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k))|_\infty,$$

ou seja,

$$(1 - \varepsilon)|x - x_k|_\infty \leq |d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k))|_\infty \leq (1 + \varepsilon)|x - x_k|_\infty.$$

Se $x \in x_k + \overline{B(0, r_k)}$, temos

$$d\phi_{x_k}^{-1} (\phi(x) - \phi(x_k)) \in \overline{B(0, (1 + \varepsilon)r_k)}$$

e, consequentemente, temos

$$\phi(x) \in \phi(x_k) + d\phi_{x_k} \left(\overline{B(0, (1 + \varepsilon)r_k)} \right).$$

Assim

$$\phi(B_k) \subset \phi(x_k) + d\phi_{x_k} \left(\overline{B(0, (1 + \varepsilon)r_k)} \right)$$

e, consequentemente,

$$m(\phi(B_k)) \leq m \left(\phi(x_k) + d\phi_{x_k} \left(\overline{B(0, (1 + \varepsilon)r_k)} \right) \right)$$

$$\begin{aligned}
&= |\det(d\phi_{x_k})|(1 + \varepsilon)^p r_k^p \\
&= (1 + \varepsilon)^p |\det(d\phi_{x_k})| m(B_k).
\end{aligned}$$

Similarmente, se

$$x \in \phi^{-1} \left(\phi(x_k) + d\phi_{x_k} \left(\overline{B(0, (1 - \varepsilon)r_k)} \right) \right),$$

então

$$\begin{aligned}
d\phi_{x_k}^{-1}(\phi(x) - \phi(x_k)) &\in \overline{B(0, (1 - \varepsilon)r_k)}, \\
|x - x_k|_\infty &\leq \frac{1}{(1 - \varepsilon)} |d\phi_{x_k}^{-1}(\phi(x) - \phi(x_k))|_\infty \leq r_k,
\end{aligned}$$

ou seja,

$$\begin{aligned}
\phi^{-1} \left(\phi(x_k) + d\phi_{x_k} \left(\overline{B(0, (1 - \varepsilon)r_k)} \right) \right) &\subset B_k, \\
\phi(x_k) + d\phi_{x_k} \left(\overline{B(0, (1 - \varepsilon)r_k)} \right) &\subset \phi(B_k), \\
(1 - \varepsilon)^p |\det(d\phi_{x_k})| m(B_k) &\leq m(\phi(B_k)).
\end{aligned}$$

Em resumo

$$\frac{m(\phi(B_k))}{(1 + \varepsilon)^p} \leq |\det(d\phi_{x_k})| m(B_k) \leq \frac{m(\phi(B_k))}{(1 - \varepsilon)^p}.$$

e como

$$m(\phi(B)) = \sum_k m(\phi(B_k)),$$

obtemos

$$\frac{m(\phi(B))}{(1 + \varepsilon)^p} \leq \sum_k |\det(d\phi_{x_k})| m(B_k) \leq \frac{m(\phi(B))}{(1 - \varepsilon)^p}.$$

Note que

$$|\det(d\phi_{x_k})| - \varepsilon \leq |\det(d\phi_x)| \leq |\det(d\phi_{x_k})| + \varepsilon$$

para $x \in B_k$ porque $|\det(d\phi_x) - \det(d\phi_{x_k})| < \varepsilon$. Dai segue que

$$\begin{aligned}
-\varepsilon m(B) &\leq \int_B |\det(d\phi_x)| dx - \sum_k |\det(d\phi_{x_k})| m(B_k) \leq \varepsilon m(B), \\
-\varepsilon m(B) + \frac{m(\phi(B))}{(1 + \varepsilon)^p} &\leq \int_B |\det(d\phi_x)| dx \leq \varepsilon m(B) + \frac{m(\phi(B))}{(1 - \varepsilon)^p}.
\end{aligned}$$

Finalmente, fazendo $\varepsilon \rightarrow 0^+$, obtemos

$$m(\phi(B)) = \int_B |\det(d\phi_x)| dx.$$

Parte 2. Suponha que $B \subset U$ é uma caixa. Sejam $B' \subset U$ um caixa com $B \subset \overset{\circ}{B'}$ e $\delta > 0$ tal que $\delta < d(B, \partial B')$.

Considere o reticulado $\delta\mathbb{Z}^p$. Sejam \mathcal{S}_1 a família dos cubos do reticulado contidos em B e \mathcal{S}_2 a família dos cubos que intersectam B (veja Figura 5.1). Dado $\varepsilon > 0$ temos que tomando δ suficientemente pequeno, podemos supor que

$$\begin{aligned}
m(B) - \sum_{C \in \mathcal{S}_1} m(C) &\leq \varepsilon, \\
\sum_{C \in \mathcal{S}_2} m(C) - m(B) &\leq \varepsilon.
\end{aligned}$$

Assim,

$$\begin{aligned}
\int_B |\det(d\phi_x)| dx &\leq \sum_{C \in \mathcal{S}_1} \int_C |\det(d\phi_x)| dx + \max_{x \in B'} |\det(d\phi_x)| \varepsilon \\
&\leq \max_{x \in B'} |\det(d\phi_x)| \varepsilon + \sum_{C \in \mathcal{S}_1} m(\phi(C)) dx
\end{aligned}$$

$$\leq \max_{x \in B'} |\det(d\phi_x)|\varepsilon + m(\phi(B))$$

e

$$\begin{aligned} \int_B |\det(d\phi_x)|dx &\geq \sum_{C \in \mathcal{S}_2} \int_C |\det(d\phi_x)|dx - \max_{x \in B'} |\det(d\phi_x)|\varepsilon \\ &\geq -\max_{x \in B'} |\det(d\phi_x)|\varepsilon + \sum_{C \in \mathcal{S}_2} m(\phi(C))dx \\ &\geq -\max_{x \in B'} |\det(d\phi_x)|\varepsilon + m(\phi(B)). \end{aligned}$$

$$-\max_{x \in B'} |\det(d\phi_x)|\varepsilon + m(\phi(B)) \leq \int_B |\det(d\phi_x)|dx \leq \max_{x \in B'} |\det(d\phi_x)|\varepsilon + m(\phi(B)).$$

Fazendo $\varepsilon \rightarrow 0^+$, obtemos

$$m(\phi(B)) = \int_B |\det(d\phi_x)|dx.$$

■

Teorema 5.31. (Teorema de mudança de variável) Sejam $\phi : U \rightarrow V$ difeomorfismo com $U, V \subset \mathbb{R}^p$ abertos, e $f : A \rightarrow \mathbb{R}$ integrável, onde $A \subset \mathbb{R}^p$ é Jordan-mensurável e \bar{A} está contido em U . Temos que

$$\int_{\phi(A)} f(\phi^{-1}(y))dy = \int_A f(x)|\det(d\phi_x)|dx$$

Demonstração. Seja W uma vizinhança de \bar{A} tal que $\bar{A} \subset W \subset \overline{W} \subset U$.

Dado $\varepsilon > 0$ considere uma partição \mathcal{P} de uma caixa contendo A . Sejam \mathcal{S}_1 as caixas de \mathcal{P} contidas em A e \mathcal{S}_2 as caixas de \mathcal{P} que intersectam A . A menos de refinar a partição \mathcal{P} , podemos supor que todas as caixas de \mathcal{S}_2 estão contidas em W .

Refinando novamente se necessário, podemos supor que

$$\int_A f(x)dx - \sum_{B \in \mathcal{S}_1} \inf_B f(z)m(B) < \varepsilon,$$

$$\sum_{B \in \mathcal{S}_2} \sup_B f(z)m(B) - \int_A f(x)dx < \varepsilon,$$

$$\sum_{B \in \mathcal{S}_2} m(B) - \sum_{B \in \mathcal{S}_1} m(B) < \varepsilon,$$

Seja $M > 0$ tal que $\sup_W |\det(d\phi_x)| < M$ e $\sup_W |f(x)| < M$.

$$\begin{aligned} \int_{\phi(A)} f(\phi^{-1}(y))dy &\leq \sum_{B \in \mathcal{S}_2} \sup_{\phi(B)} f(\phi^{-1}(y))m(\phi(B)) \\ &= \sum_{B \in \mathcal{S}_2} \sup_B f(x) \int_B |\det(d\phi_x)|dx \\ &\leq \sum_{B \in \mathcal{S}_1} \sup_B f(x) \int_B |\det(d\phi_x)|dx + \sum_{B \in \mathcal{S}_2 \setminus \mathcal{S}_1} \sup_B |f(x)| \int_B |\det(d\phi_x)|dx \end{aligned}$$

Similarmente,

$$\begin{aligned} \int_{\phi(A)} f(\phi^{-1}(y))dy &\geq \sum_{B \in \mathcal{S}_1} \inf_{\phi(B)} f(\phi^{-1}(y))m(\phi(B)) \\ &= \sum_{B \in \mathcal{S}_1} \inf_B f(x) \int_B |\det(d\phi_x)|dx. \end{aligned}$$

Da mesma forma,

$$\sum_{B \in \mathcal{S}_1} \inf_B f(x) \int_B |\det(d\phi_x)| dx \leq \int_A f(x) |\det(d\phi_x)| dx \leq \sum_{B \in \mathcal{S}_2} \sup_B f(x) \int_B |\det(d\phi_x)| dx.$$

Portanto,

$$\begin{aligned} \left| \int_A f(x) |\det(d\phi_x)| dx - \int_{\phi(A)} f(\phi^{-1}(y)) dy \right| &\leq \\ \sum_{B \in \mathcal{S}_2 \setminus \mathcal{S}_1} \sup_B |f(z)| \int_B |\det(d\phi_x)| dx + \sum_{B \in \mathcal{S}_1} \left(\sup_B f(z) - \inf_B f(z) \right) \int_B |\det(d\phi_x)| dx \end{aligned}$$

Como

$$\sum_{B \in \mathcal{S}_2 \setminus \mathcal{S}_1} \sup_B |f(z)| \int_B |\det(d\phi_x)| dx \leq M^2 \varepsilon$$

e

$$\sum_{B \in \mathcal{S}_1} \left(\sup_B f(z) - \inf_B f(z) \right) \int_B |\det(d\phi_x)| dx \leq M \varepsilon.$$

Concluímos que

$$\left| \int_A f(x) |\det(d\phi_x)| dx - \int_{\phi(A)} f(\phi^{-1}(y)) dy \right| \leq (M^2 + M) \varepsilon.$$

Fazendo $\varepsilon \rightarrow 0^+$ obtemos

$$\int_{\phi(A)} f(\phi^{-1}(y)) dy = \int_A f(x) |\det(d\phi_x)| dx.$$

■

Exemplo 5.32. Seja $f : B(0, R) \setminus \{0\} \rightarrow \mathbb{R}$ integrável, onde $B(0, R)$ é a bola em \mathbb{R}^2 centrada em 0 com raio R . Temos que para

$$A_\varepsilon := \{(r \cos(\theta), r \sin(\theta)) \in \mathbb{R}^2 : \varepsilon \leq r < R \text{ e } \varepsilon \leq \theta \leq 2\pi - \varepsilon\}$$

$$\begin{aligned} \int_{A_\varepsilon} f(x, y) dx dy &= \int_\varepsilon^{2\pi - \varepsilon} \int_\varepsilon^R f(r \cos(\theta), r \sin(\theta)) \left| \det \begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} \right| dr d\theta \\ &= \int_\varepsilon^{2\pi - \varepsilon} \int_\varepsilon^R f(r \cos(\theta), r \sin(\theta)) r dr d\theta. \end{aligned}$$

Fazendo $\varepsilon \rightarrow 0^+$ obtemos

$$\int_{B(0, R)} f(x, y) dx dy = \int_0^{2\pi} \int_0^R f(r \cos(\theta), r \sin(\theta)) r dr d\theta,$$

porque f é limitada,

$$m(B(0, R) \setminus A_\varepsilon) = \pi R^2 - \left(\frac{1}{2} (2\pi - 2\varepsilon) R^2 - \frac{1}{2} (2\pi - 2\varepsilon) \varepsilon^2 \right) \rightarrow 0$$

e

$$m([0, 2\pi] \times [0, R] \setminus [\varepsilon, 2\pi - \varepsilon] \times [\varepsilon, R]) = 2\pi R - (2\pi - 2\varepsilon)(R - \varepsilon) \rightarrow 0.$$

Observação 5.33. (Integral imprópria) Fixe o aberto $U \subset \mathbb{R}^p$. Considere uma família de conjuntos compactos mensuráveis F_k tais que $F_k \subset \overset{\circ}{F}_{k+1}$ e $U = \cup_k F_k$.

Por exemplo, seja C_n a união das caixas

$$\left[\frac{l^1}{2^n}, \frac{l^1 + 1}{2^n} \right] \times \cdots \times \left[\frac{l^p}{2^n}, \frac{l^p + 1}{2^n} \right] \subset U \cap [-2^n, 2^n]^p$$

com inteiros l^1, \dots, l^p (Veja Figura 5.3). Seja n_1 o primeiro índice com $C_{n_1} \neq \emptyset$. Existe $n_2 > n_1$ tal que $C_{n_1} \subset \dot{C}_{n_2}$. Recursivamente, temos $C_{n_1}, C_{n_2}, \dots, C_{n_k}, \dots$ e definindo $F_k := C_{n_k}$ obtemos a sequência de compactos encaixados tal que $F_k \subset \dot{F}_{k+1}$ e $U = \cup_k F_k$.

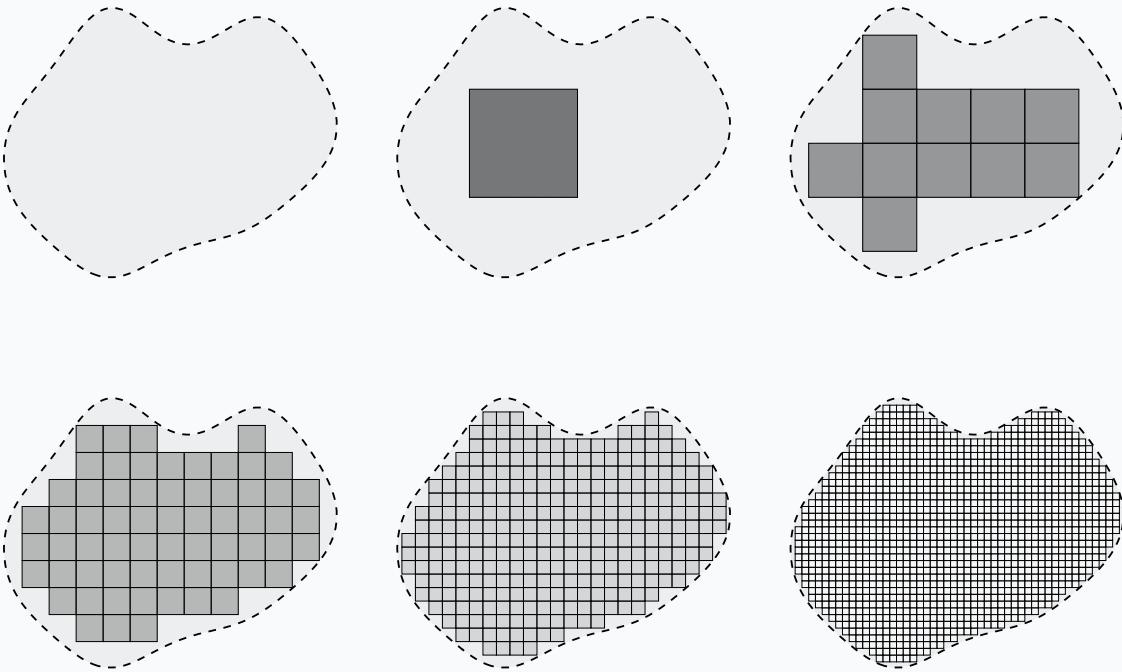


Figura 5.3: Exaurindo um aberto com caixas com lados $\frac{1}{2^n}$.

Seja $f : U \rightarrow \mathbb{R}$ contínua e não negativa.

Temos que

$$\int_{F_k} f(x) dx$$

é não decrescente e portanto definimos

$$\int_U f(x) dx := \lim_{k \rightarrow \infty} \int_{F_k} f(x) dx.$$

Tal limite pode ser finito ou infinito.

Se G_n é outra família de conjuntos compactos mensuráveis com as propriedades $G_n \subset \dot{G}_{n+1}$ e $U = \cup_n G_n$, então temos que os abertos \dot{F}_k cobrem o compacto G_n e assim existe k' tal que $G_n \subset \dot{F}_{k'}$. Assim,

$$\begin{aligned} \int_{G_n} f(x) dx &\leq \lim_{k \rightarrow \infty} \int_{F_k} f(x) dx, \\ \lim_{n \rightarrow \infty} \int_{G_n} f(x) dx &\leq \lim_{k \rightarrow \infty} \int_{F_k} f(x) dx. \end{aligned}$$

Analogamente,

$$\lim_{k \rightarrow \infty} \int_{F_k} f(x) dx \leq \lim_{n \rightarrow \infty} \int_{G_n} f(x) dx.$$

Assim,

$$\int_U f(x) dx = \lim_{n \rightarrow \infty} \int_{G_n} f(x) dx$$

Portanto, a integral

$$\int_U f(x) dx$$

independe da escolha da família de compactos encaixados crescentes como os descritos acima, ou seja, a integral de funções contínuas não negativas está definida em abertos.

Agora, considere $g : U \rightarrow \mathbb{R}$ contínua. Podemos escrever $g = g^+ - g^-$ e $|g| = g^+ + g^-$, onde $g^+ = \frac{|g|+g}{2}$ e $g^- = \frac{|g|-g}{2}$ são funções não negativas e contínuas.

Dizemos que $g : U \rightarrow \mathbb{R}$ é integrável se

$$\int_U |g(x)|dx < \infty.$$

Nesse caso, temos que

$$\int_U g^+(x)dx < \infty \quad \text{e} \quad \int_U g^-(x)dx < \infty$$

porque

$$\int_U |g(x)|dx = \int_U g^+(x)dx + \int_U g^-(x)dx.$$

Assim, a integral de g é definida como

$$\int_U g(x)dx := \int_U g^+(x)dx - \int_U g^-(x)dx.$$

Portanto, temos que g é integrável (como integral imprópria) com integral

$$\int_U g(x)dx = \lim_{n \rightarrow \infty} \int_{F_n} g(x)dx$$

quando $\int_U |g(x)|dx < \infty$. O mesmo argumento vale para funções g integráveis no sentido que o conjunto

$$\{x \in U : g \text{ é descontínua em } x\}$$

tem medida zero.

Em particular, se temos uma função $f : \mathbb{R}^p \rightarrow \mathbb{R}$ integrável no sentido que

$$\int_{\mathbb{R}^p} |f(x)|dx := \lim_{R \rightarrow \infty} \int_{|x|_\infty \leq R} |f(x)|dx < \infty,$$

então

$$\int_{\mathbb{R}^p} f(x)dx := \lim_{R \rightarrow \infty} \int_{|x|_\infty \leq R} f(x)dx.$$

Se $f : \mathbb{R}^p \rightarrow \mathbb{R}$ é contínua, então podemos aplicar o teorema de Fubini com integrais impróprias:

$$\int_{\mathbb{R}^p} f(x)dx = \int_{-\infty}^{\infty} \cdots \left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x^1, \dots, x^p) dx^1 \right) dx^2 \right) \cdots dx^p.$$

Exemplo 5.34. Mostremos que

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Usando coordenadas polares, em \mathbb{R}^2 temos a integral

$$\int_{x^2+y^2 < r^2} e^{-x^2-y^2} dxdy = \int_0^r \int_0^{2\pi} e^{-s^2} s d\theta ds = 2\pi \int_0^r e^{-s^2} s ds = 2\pi \left(\frac{1}{2} - \frac{e^{-r^2}}{2} \right) \rightarrow \pi.$$

Por outro lado,

$$\int_{|x| < r, |y| < r} e^{-x^2-y^2} dxdy = \left(\int_{-r}^r e^{-x^2} dx \right) \left(\int_{-r}^r e^{-y^2} dy \right) = \left(\int_{-r}^r e^{-x^2} dx \right)^2.$$

Como

$$\int_{|x| < \frac{\sqrt{2}}{2}r, |y| < \frac{\sqrt{2}}{2}r} e^{-x^2-y^2} dxdy \leq \int_{x^2+y^2 < r^2} e^{-x^2-y^2} dxdy \leq \int_{|x| < r, |y| < r} e^{-x^2-y^2} dxdy$$

temos que

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \int_{\mathbb{R}^2} e^{-x^2-y^2} dxdy = \pi$$

Exemplo 5.35. Similar ao exemplo anterior, se $f : B(0, R) \setminus \{0\} \rightarrow \mathbb{R}$ é integrável, onde $B(0, R)$ é a bola de centro 0 e raio R em \mathbb{R}^3 , então em coordenadas esféricas

$$\begin{cases} x = r \cos(\theta) \cos(\phi) \\ y = r \sin(\theta) \cos(\phi) \\ z = r \sin(\phi) \end{cases}$$

com $0 < \theta < 2\pi$, $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$ e $0 < r < R$, obtemos

$$\begin{aligned} \int_{B(0, R)} f(x, y, z) dxdydz &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} \int_0^R f(r \cos(\theta) \cos(\phi), r \sin(\theta) \cos(\phi), r \sin(\phi)) \left| \det \begin{pmatrix} \frac{\partial_r x}{\partial_r y} & \frac{\partial_\theta x}{\partial_\theta y} & \frac{\partial_\phi x}{\partial_\phi y} \\ \frac{\partial_r z}{\partial_r y} & \frac{\partial_\theta z}{\partial_\theta y} & \frac{\partial_\phi z}{\partial_\phi y} \end{pmatrix} \right| dr d\theta d\phi \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} \int_0^R f(r \cos(\theta) \cos(\phi), r \sin(\theta) \cos(\phi), r \sin(\phi)) r \sin(\phi) dr d\theta d\phi \end{aligned}$$

Exemplo 5.36. Seja $A \subset U$ um conjunto Jordan mensurável com $\bar{A} \subset U$. Temos que

$$\begin{aligned} m(\phi(A)) &= \int_V \chi_{\phi(A)}(y) dy = \int_V \chi_A(\phi^{-1}(y)) dy = \int_U \chi_A(x) |\det(d\phi_x)| dx, \\ m(\phi(A)) &= \int_U \chi_A(x) |\det(d\phi_x)| dx. \end{aligned}$$

Considere os vetores linearmente independentes v_1, \dots, v_p em \mathbb{R}^p . Considere o isomorfismo linear $\phi(x) = (v_j^i)x$. Seja Δ_p o simplexo padrão definido em 5.25. Temos que simplexo

$$\Pi := \phi(\Delta_p) = \left\{ x^1 v_1 + \dots + x^p v_p : x_1, \dots, x_p \geq 0, \sum_{i=1}^p x_i \leq 1 \right\}$$

tem volume

$$\begin{aligned} m(\Pi) &= m(\phi(\Delta_p)) = \int_{\mathbb{R}^p} \chi_{\Delta_p}(x) |\det(\phi)| dx = |\det(v_j^i)| m(\Delta_p) = \frac{|\det(v_j^i)|}{p!}, \\ m(\Pi) &= \frac{|\det(v_j^i)|}{p!}, \end{aligned}$$

onde usamos que $m(\Delta_p) = \frac{1}{p!}$, o que foi provado no Exemplo 5.25.

5.4 Teorema da divergência

Abordaremos essa seção sem muito rigor.

Considere um hiperplano $\langle x - x_0, n \rangle = 0$ em \mathbb{R}^p , onde x_0 é um ponto do hiperplano e n é um vetor unitário normal ao hiperplano. Suponha que $n^p \neq 0$ e considere $v_1, \dots, v_{p-1} \in n^\perp$. O volume do paralelogramo

$$x + \{t^1 v_1 + \dots + t^{p-1} v_{p-1} \in \mathbb{R}^p : t^1, \dots, t^{p-1} \in [0, 1]\}$$

é

$$|\det(v_1, \dots, v_{p-1}, n)|$$

porque esse é o volume de um sólido com base $\{t^1 v_1 + \dots + t^{p-1} v_{p-1} \in \mathbb{R}^p : t^1, \dots, t^{p-1} \in [0, 1]\}$ e altura 1.

Seja $\pi(u) = u - \langle u, e_p \rangle e_p$ a projeção de vetores em $\mathbb{R}^{p-1} \times \{0\}$.

$$\begin{aligned} \det(\pi(v_1), \dots, \pi(v_{p-1}), e_p) &= \det(v_1, \dots, v_{p-1}, e_p) \\ &= \det(v_1, \dots, v_{p-1}, e_p - \langle e_p, n \rangle n) + \det(v_1, \dots, v_{p-1}, \langle e_p, n \rangle n) \\ &= \det(v_1, \dots, v_{p-1}, \langle e_p, n \rangle n) \\ &= n^p \det(v_1, \dots, v_{p-1}, n), \end{aligned}$$

ou seja,

$$|\det(\pi(v_1), \dots, \pi(v_{p-1}), e_p)| = |n^p| |\det(v_1, \dots, v_{p-1}, n)|.$$

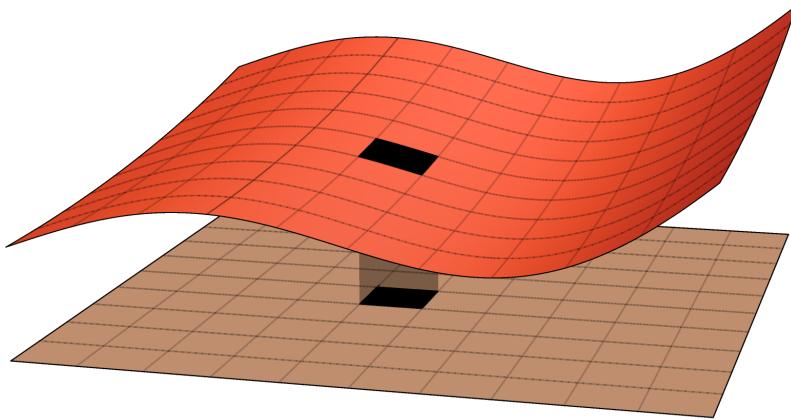


Figura 5.4: Área de superfície.

Assim, a relação entre o elemento de área $d\mathcal{O}$ no plano $\langle x - x_0, n \rangle = 0$ e o elemento de área em $\mathbb{R}^{p-1} \times \{0\}$ é dada por $d\mathcal{O} = \frac{dx^1 \cdots dx^{p-1}}{|n^p|}$.

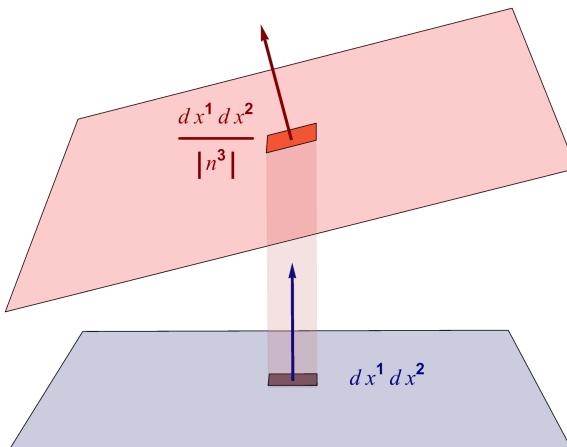


Figura 5.5: Elemento de área $d\mathcal{O}$ pode ser escrito como $\frac{dx^1 \cdots dx^{p-1}}{|n^p|}$ se $n^p \neq 0$.

Exemplo 5.37. Considere $x^p = f(x^1, \dots, x^{p-1})$. O vetor normal ao gráfico é

$$n(x^1, \dots, x^{p-1}) = \frac{(-\partial_1 f, \dots, -\partial_{p-1} f, 1)}{\sqrt{(\partial_1 f)^2 + \dots + (\partial_{p-1} f)^2 + 1}}.$$

O elemento de área na superfície é

$$\begin{aligned} d\mathcal{O} &= \frac{dx^1 \cdots dx^{p-1}}{|n^p|} \\ &= \frac{dx^1 \cdots dx^{p-1}}{\sqrt{(\partial_1 f)^2 + \dots + (\partial_{p-1} f)^2 + 1}}. \end{aligned}$$

Assim, para $D \subset \mathbb{R}^{p-1}$ onde f está definida, a área de $S = \{(x, f(x)) \in \mathbb{R}^p : x \in D\}$ é

$$\mathcal{O}(S) = \int_D \frac{dx^1 \cdots dx^{p-1}}{\sqrt{(\partial_1 f)^2 + \dots + (\partial_{p-1} f)^2 + 1}}$$

Por exemplo, para esfera de centro 0 e raio R em \mathbb{R}^3 , temos que a calota superior é o gráfico de $z = \sqrt{R^2 - x^2 - y^2}$. Temos que

$$\sqrt{(\partial_x z)^2 + (\partial_y z)^2 + 1} = \frac{R}{\sqrt{R^2 - x^2 - y^2}}.$$

Usando coordenadas polares $x = r \cos(\theta)$ e $y = r \sin(\theta)$, temos que a área da calota superior é

$$\begin{aligned} \int_{x^2+y^2 < R^2} \frac{\sqrt{R^2 - x^2 - y^2}}{R} dx dy &= \int_0^{2\pi} \int_0^R \frac{\sqrt{R^2 - r^2}}{R} r dr d\theta \\ &= 2\pi \int_0^R \frac{\sqrt{R^2 - r^2}}{R} r dr \\ &= \frac{2\pi}{3} R^2. \end{aligned}$$

Portanto, a área da esfera é $\frac{4\pi}{3} R^3$.

Sejam $T \subset \mathbb{R}^p$ um p -simplexo não degenerado e $f : T \rightarrow \mathbb{R}$ uma função contínua e de classe C^1 em $\overset{\circ}{T}$.

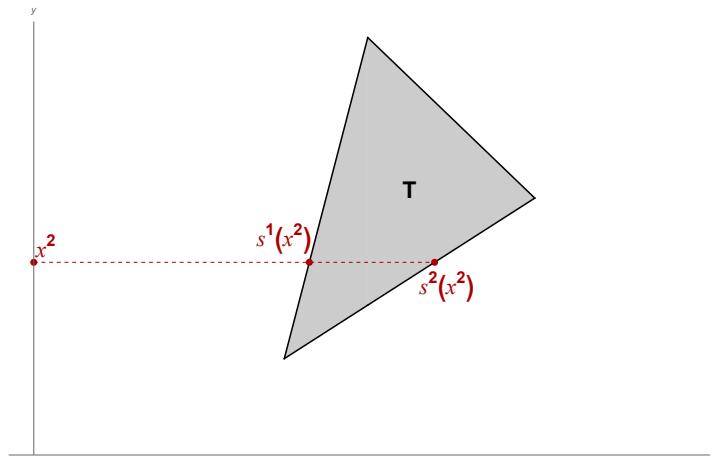


Figura 5.6: Simplexo em \mathbb{R}^2

Seja Δ a projeção de T no hiperplano $\{0\} \times \mathbb{R}^{p-1}$. Para cada $(x^2, \dots, x^p) \in \Delta$ temos

$$(s^1(x^2, \dots, x^p), x^2, \dots, x^p), \quad (s^2(x^2, \dots, x^p), x^2, \dots, x^p) \in \partial T$$

tais que $s^1(x^2, \dots, x^p) \leq s^2(x^2, \dots, x^p)$. Seja T_i o gráfico de s_i .

$$\begin{aligned}
& \int_T \frac{\partial f}{\partial x^1}(x^1, \dots, x^p) dx^1 \cdots dx^p = \\
& \int_{\Delta} f(s^2(x^2, \dots, x^p), x^2, \dots, x^p) dx^2 \cdots dx^p - \int_{\Delta} f(s^1(x^2, \dots, x^p), x^2, \dots, x^p) dx^2 \cdots dx^p = \\
& \int_{\Delta} f(s^2(x^2, \dots, x^p), x^2, \dots, x^p) n^1(s^2(x^2, \dots, x^p), x^2, \dots, x^p) \frac{dx^2 \cdots dx^p}{|n^1(s^2(x^2, \dots, x^p), x^2, \dots, x^p)|} + \\
& \int_{\Delta} f(s^1(x^2, \dots, x^p), x^2, \dots, x^p) n^1(s^1(x^2, \dots, x^p), x^2, \dots, x^p) \frac{dx^2 \cdots dx^p}{|n^1(s^1(x^2, \dots, x^p), x^2, \dots, x^p)|},
\end{aligned}$$

onde usamos que $n^1(s^1(x^2, \dots, x^p))$ é negativo e $n^1(s^2(x^2, \dots, x^p))$ é positivo. Repare que os vetores normais não estão definidos em todos os pontos, mas os pontos onde tais vetores não estão definidos são negligenciáveis, ou seja, eles podem ser ignorados na integração.

Como $d\mathcal{O} = \frac{dx^2 \cdots dx^p}{|n^1(x)|}$ temos

$$\int_{T_1} f(x) n^1(x) d\mathcal{O} + \int_{T_2} f(x) n^1(x) d\mathcal{O} = \int_{\partial T} f(x) n^1(x) d\mathcal{O}.$$

Em geral, para um campo vetorial $F = (F^1, \dots, F^p)$ temos

$$\int_T \left(\frac{\partial F^1}{\partial x^1}(x) + \cdots + \frac{\partial F^p}{\partial x^p}(x) \right) dx^1 \cdots dx^p = \int_{\partial T} (F^1(x) n^1(x) + \cdots + F^p(x) n^p(x)) d\mathcal{O}.$$

Definição 5.38. O divergente de um campo F de classe C^1 é

$$\operatorname{div} F = \frac{\partial F^1}{\partial x^1} + \cdots + \frac{\partial F^p}{\partial x^p}.$$

Portanto,

$$\int_T \operatorname{div} F(x) dx^1 \cdots dx^p = \int_{\partial T} \langle F(x), n(x) \rangle d\mathcal{O}.$$

Teorema 5.39. (Teorema da divergência) Considere um aberto D limitado em \mathbb{R}^p e suponha que ∂D é uma $(p-1)$ -variedade suave. Seja $n : \partial D \rightarrow \mathbb{R}^p$ o campo unitário normal a ∂D apontando para fora de D . Seja $F : \bar{D} \rightarrow \mathbb{R}^p$ contínua e suponha que $F|_D$ de classe C^1 . Então,

$$\int_D \operatorname{div} F(x) dx^1 \cdots dx^p = \int_{\partial D} \langle F(x), n(x) \rangle d\mathcal{O}.$$

Isto é verdade porque podemos aproximar D e sua fronteira por um modelo discreto (um politopo, que é uma generalização de poliedro), isto é, por um modelo formado por p -simplexos (veja Figura 5.9). Note que estamos usando que se T e T' são dois p -simplexos adjacentes, então as integrais nas faces comuns se cancelam porque os vetores normais apontam em direções opostas porque o vetor normal em ∂T aponta para fora de T e o vetor normal em $\partial T'$ aponta para fora de T' .

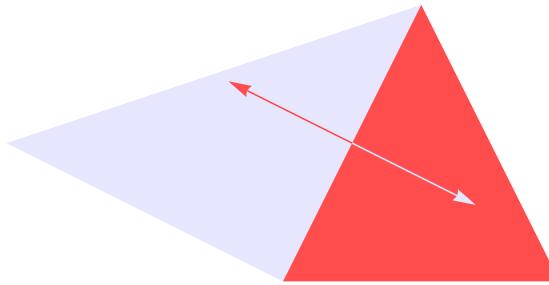


Figura 5.7: Triângulos adjacentes e seus vetores normais.

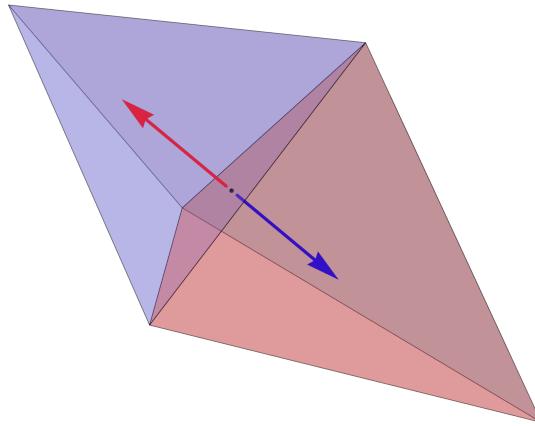


Figura 5.8: Tetraedros adjacentes e seus vetores normais.

Assim,

$$\int_{\partial T} \langle F(x), n(x) \rangle d\mathcal{O} + \int_{\partial T'} \langle F(x), n(x) \rangle d\mathcal{O} = \int_{\partial(T \cup T')} \langle F(x), n(x) \rangle d\mathcal{O}.$$

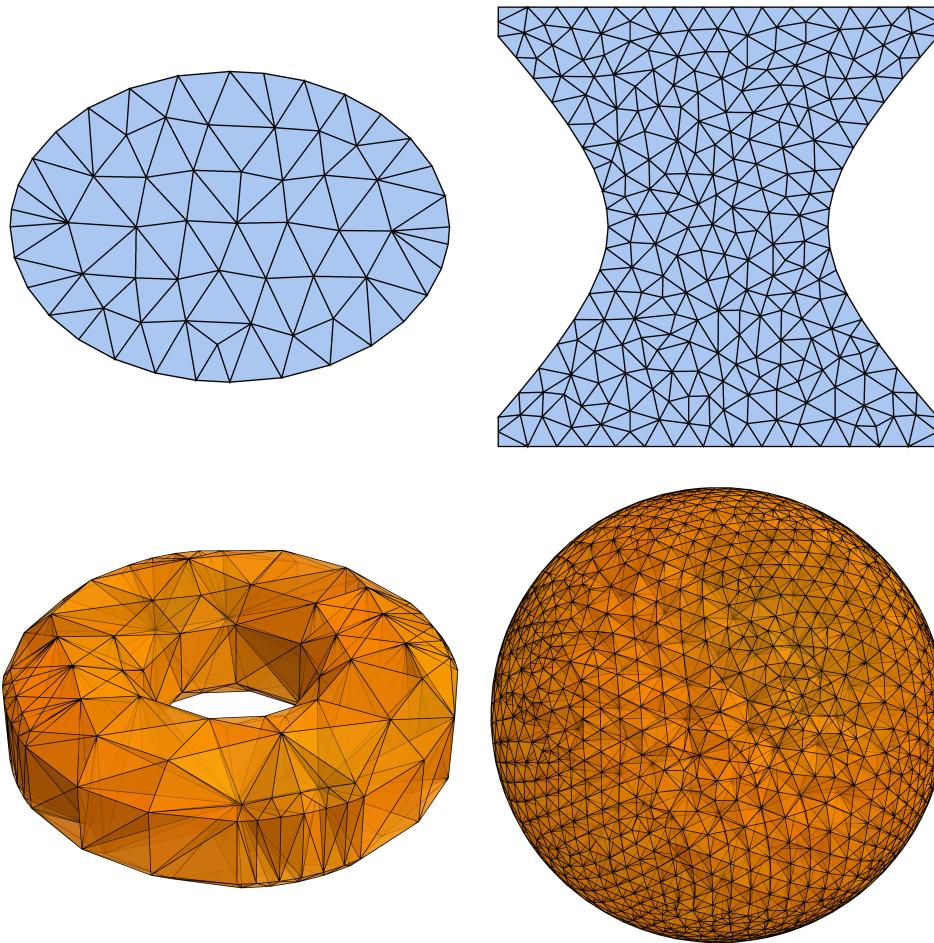


Figura 5.9: Triangulações de variedades e suas fronteiras.

Teorema 5.40. (Integral em coordenadas polares para funções de classe C^1)

Considere $f : B(0, R) \rightarrow \mathbb{R}$ de classe C^1 . Temos que

$$\int_{B(0,R)} f(x) dx = \int_0^R \left(\int_{\mathbb{S}^{p-1}} f(ru) d\mathcal{O}(u) \right) r^{n-1} dr.$$

Observação 5.41. O teorema acima vale para f contínua. Isso pode ser provado aproximando f por funções de classe C^1 uniformemente.

Demonstração. Defina

$$g(r) = \int_{B(0,r)} f(y) dy.$$

Pelo teorema de mudança de variáveis,

$$\int_{B(0,r)} f(y) dy = r^p \int_{B(0,1)} f(rx) dx.$$

Assim,

$$\begin{aligned} g'(r) &= pr^{p-1} \int_{B(0,1)} f(rx) dx + r^p \int_{B(0,1)} \sum_{i=1}^p \partial_i f(rx) x^i dx \\ &= pr^{p-1} \int_{B(0,1)} f(rx) dx + r^{p-1} \int_{B(0,1)} \sum_{i=1}^p r \partial_i f(rx) x^i dx. \end{aligned}$$

Como

$$\partial_i(f(rx)x^i) = r\partial_i f(rx)x^i + f(rx),$$

temos

$$\begin{aligned} \int_{B(0,1)} \sum_{i=1}^p r \partial_i f(rx) x^i dx &= \int_{B(0,1)} \sum_{i=1}^p \partial_i(f(rx)x^i) dx - \int_{B(0,1)} \sum_{i=1}^p f(x) dx \\ &= \int_{B(0,1)} \sum_{i=1}^p \partial_i(f(rx)x^i) dx - p \int_{B(0,1)} f(rx) dx. \end{aligned}$$

Pelo teorema da divergência

$$\int_{B(0,1)} \sum_{i=1}^p \partial_i(f(rx)x^i) dx = \int_{\mathbb{S}^{p-1}} \sum_{i=1}^p f(ru)(u^i)^2 d\mathcal{O}(u) = \int_{\mathbb{S}^{p-1}} f(ru) d\mathcal{O}(u).$$

Portanto,

$$\int_{B(0,1)} \sum_{i=1}^p s \partial_i f(rx) x^i dx = \int_{\mathbb{S}^{p-1}} f(ru) d\mathcal{O}(u) - p \int_{B(0,1)} f(rx) dx$$

e, consequentemente,

$$g'(r) = r^{p-1} \int_{\mathbb{S}^{p-1}} f(ru) d\mathcal{O}(u).$$

Como $g(0) = 0$, temos que

$$\int_{B(0,R)} f(x) dx = g(R) = \int_0^R \left(\int_{\mathbb{S}^{p-1}} f(ru) d\mathcal{O}(u) \right) r^{p-1} ds.$$

■

Exemplo 5.42. (Integral em coordenadas polares/caso geral/brute force)

Consideremos coordenadas polares em \mathbb{R}^p

$$\begin{cases} x^1 = r \cos(\theta^1) \\ x^2 = r \sin(\theta^1) \cos(\theta^2) \\ x^3 = r \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \\ \vdots \\ x^{p-1} = r \sin(\theta^1) \sin(\theta^2) \sin(\theta^3) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \\ x^p = r \sin(\theta^1) \sin(\theta^2) \sin(\theta^3) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \end{cases}$$

com $r \in \mathbb{R}_{>0}$, $0 < \theta^1, \dots, \theta^{p-2} < \pi$ e $0 < \theta^{p-1} < 2\pi$. Para $x = (x^1, \dots, x^p)$ obtemos

$$\begin{aligned} \partial_r x &= \begin{bmatrix} \cos(\theta^1) \\ \sin(\theta^1) \cos(\theta^2) \\ \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \\ \vdots \\ \sin(\theta^1) \sin(\theta^2) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \\ \sin(\theta^1) \sin(\theta^2) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \end{bmatrix} \\ \partial_{\theta^1} x &= \begin{bmatrix} -r \sin(\theta^1) \\ r \cos(\theta^1) \cos(\theta^2) \\ r \cos(\theta^1) \sin(\theta^2) \cos(\theta^3) \\ \vdots \\ r \cos(\theta^1) \sin(\theta^2) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \\ r \cos(\theta^1) \sin(\theta^2) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \end{bmatrix} \\ \partial_{\theta^2} x &= \begin{bmatrix} 0 \\ -r \sin(\theta^1) \sin(\theta^2) \\ r \sin(\theta^1) \cos(\theta^2) \cos(\theta^3) \\ \vdots \\ r \sin(\theta^1) \cos(\theta^2) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \\ r \sin(\theta^1) \cos(\theta^2) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \end{bmatrix} \\ \partial_{\theta^3} x &= \begin{bmatrix} 0 \\ 0 \\ -r \sin(\theta^1) \sin(\theta^2) \sin(\theta^3) \\ \vdots \\ r \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \\ r \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \end{bmatrix} \\ &\vdots \\ \partial_{\theta^{p-1}} x &= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -r \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \\ r \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \end{bmatrix} \end{aligned}$$

Os vetores acima são ortogonais e tem as seguintes normas

$$\begin{aligned} |\partial_r x| &= 1 \\ |\partial_{\theta^1} x| &= 1, \\ |\partial_{\theta^2} x| &= r \sin(\theta^1), \\ |\partial_{\theta^3} x| &= r \sin(\theta^1) \sin(\theta^2), \\ &\vdots \\ |\partial_{\theta^{p-1}} x| &= r \sin(\theta^1) \sin(\theta^2) \cdots \sin(\theta^{p-2}). \end{aligned}$$

Assim, o Jacobiano de x é $[\partial_r x \ \partial_{\theta^1} x \ \dots \ \partial_{\theta^{p-1}} x]$ e o valor absoluto de seu determinante é

$$r^{p-1} \sin(\theta^1)^{p-2} \sin(\theta^2)^{p-3} \cdots \sin(\theta^{p-2}),$$

o volume da caixa com lados $\partial_r x, \partial_{\theta^1} x, \dots, \partial_{\theta^{p-1}} x$.

Assim, escrevendo $\theta = (\theta^1, \dots, \theta^{p-1})$, $\Pi := [0, \pi]^{p-2} \times [0, 2\pi]$ e

$$\left\{ \begin{array}{l} u^1 = \cos(\theta^1) \\ u^2 = \sin(\theta^1) \cos(\theta^2) \\ u^3 = \sin(\theta^1) \sin(\theta^2) \cos(\theta^3) \\ \vdots \\ u^{p-1} = \sin(\theta^1) \sin(\theta^2) \sin(\theta^3) \cdots \sin(\theta^{p-2}) \cos(\theta^{p-1}) \\ u^p = \sin(\theta^1) \sin(\theta^2) \sin(\theta^3) \cdots \sin(\theta^{p-2}) \sin(\theta^{p-1}) \end{array} \right.$$

obtemos a parametrização da esfera $u(\theta)$ e

$$\int_{B(0,R)} f(x) dx = \int_0^R \left(\int_{\Pi} f(ru(\theta)) \sin(\theta^1)^{p-2} \sin(\theta^2)^{p-3} \cdots \sin(\theta^{p-2}) d\theta^1 \cdots d\theta^{p-1} \right) r^{p-1} dr.$$

Como $\partial_{\theta^1} u, \dots, \partial_{\theta^{p-1}} u$ são tangentes a esfera \mathbb{S}^{p-1} em $u(\theta)$, dois a dois ortogonais e

$$\begin{aligned} |\partial_{\theta^1} u(\theta)| &= 1, \\ |\partial_{\theta^2} u(\theta)| &= \sin(\theta^1), \\ &\vdots \\ |\partial_{\theta^{p-1}} u(\theta)| &= \sin(\theta^1) \sin(\theta^2) \cdots \sin(\theta^{p-2}). \end{aligned}$$

obtemos que o elemento de volume na esfera \mathbb{S}^{p-1} é

$$d\mathcal{O} = \sin(\theta^1)^{p-2} \sin(\theta^2)^{p-3} \cdots \sin(\theta^{p-2}) d\theta^1 \cdots d\theta^{p-2}.$$

Assim,

$$\int_{B(0,R)} f(x) dx = \int_0^R \left(\int_{\mathbb{S}^{p-1}} f(ru) d\mathcal{O}(u) \right) r^{p-1} dr.$$

Exercício 5.43. A função Gamma é definida como

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt, \quad x \in \mathbb{R}_{\geq 1}.$$

Mostre que $\Gamma(x+1) = x\Gamma(x)$ e $\Gamma(1) = 1$. Em particular, $\Gamma(n+1) = n!$ para $n \in \mathbb{Z}_{\geq 0}$.

Exemplo 5.44. (Volume de esferas)

Após aplicar integração em coordenadas polares e mandar o raio da bola para infinito, obtemos

$$\int_{\mathbb{R}^p} e^{-(x^1)^2 - \cdots - (x^p)^2} dx^1 \cdots dx^p = \mathcal{O}(\mathbb{S}^{p-1}) \int_0^\infty e^{-r^2} r^{p-1} dr.$$

Aplicando o teorema de Fubini e a integral no exemplo 5.34, obtemos

$$\int_{\mathbb{R}^p} e^{-(x^1)^2 - \cdots - (x^p)^2} dx^1 \cdots dx^p = \left(\int_{-\infty}^\infty e^{-x^2} dx \right)^p = \pi^{\frac{p}{2}}.$$

Fazendo $r = t^{\frac{1}{2}}$ obtemos

$$\int_0^\infty e^{-r^2} r^{p-1} dr = \frac{1}{2} \int_0^\infty e^{-t} t^{\frac{p-1}{2}} t^{-\frac{1}{2}} dt = \frac{1}{2} \int_0^\infty e^{-t} t^{\frac{p}{2}-1} dt = \frac{1}{2} \Gamma\left(\frac{p}{2}\right).$$

Portanto,

$$\mathcal{O}(\mathbb{S}^{p-1}) = \frac{2\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}\right)}.$$

O volume da bola de raio 1 em \mathbb{R}^p é

$$m(B(0, 1)) = \int_{B(0, 1)} 1 dx = \int_0^1 \left(\int_{S^{p-1}} 1 d\mathcal{O} \right) r^{p-1} dr = \frac{\mathcal{O}(S^{p-1})}{p}.$$

Teorema 5.45. (Teorema de Green) Considere um aberto limitado $D \subset \mathbb{R}^2$ com fronteira ∂D . Suponha que ∂D é uma 1-variedade. Seja $N = (n^1, n^2)$ o campo normal unitário apontando para fora. O campo tangente a $\tau = (\tau^1, \tau^2)$ que percorre ∂D no sentido anti-horário é $\tau = (-n^2, n^1)$. Seja dl a medida de comprimento em ∂D . Temos

$$\int_D (\partial_1 f^2(x) - \partial_2 f^1(x)) dx^1 dx^2 = \int_{\partial D} \langle f(x), \tau(x) \rangle dl$$

para $f : \overline{D} \rightarrow \mathbb{R}$ contínua e de classe C^1 em D .

Demonstração. Segue do teorema da divergência

$$\begin{aligned} \int_D \partial_1 f^2 + \partial_2(-f^1) dx^1 dx^2 &= \int_{\partial D} (f^2 n^1 + (-f^1) n^2) dl, \\ \int_D (\partial_1 f^2 - \partial_2 f^1) dx^1 dx^2 &= \int_{\partial D} (f^2 n^1 - f^1 n^2) dl, \\ \int_D (\partial_1 f^2 - \partial_2 f^1) dx^1 dx^2 &= \int_{\partial D} (f^1 \tau^1 + f^2 \tau^2) dl. \end{aligned}$$

■

Exemplo 5.46. Considere uma curva $\gamma : [a, b] \rightarrow \mathbb{R}^2$ regular simples tal que $\dot{\gamma}(a) = \dot{\gamma}(b)$. Suponha que γ percorre um caminho no sentido anti-horário e limita um aberto D . Temos que $\tau = \frac{\dot{\gamma}(t)}{|\dot{\gamma}(t)|}$ e $dl = |\dot{\gamma}(t)| dt$.

Assim,

$$\int_D (\partial_1 f^2(x^1, x^2) - \partial_2 f^1(x^1, x^2)) dx^1 dx^2 = \int_a^b (f^1(\gamma(t)) \dot{\gamma}^1(t) + f^2(\gamma(t)) \dot{\gamma}^2(t)) dt.$$

Exemplo 5.47. A menos de reparametrizar uma curva regular, podemos supor que ela está definida no intervalo $[0, 1]$ e tem velocidade constante.

Considere uma curva regular simples $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$ de classe C^∞ tal que $\gamma(t) = \gamma(t+1)$ e percorre no sentido anti-horário. Suponha que γ limita um aberto D e tem velocidade $v \in \mathbb{R}_{>0}$. Para $f^1(x^1, x^2) = -\frac{x^2}{2}$, $f^2(x^1, x^2) = \frac{x^1}{2}$ obtemos que a área de D é dada por

$$m(D) = \int_D (\partial_1 f^2(x^1, x^2) - \partial_2 f^1(x^1, x^2)) dx^1 dx^2 = \frac{1}{2} \int_0^1 (-\gamma^2(t) \dot{\gamma}^1(t) + \gamma^1(t) \dot{\gamma}^2(t)) dt.$$

Identificando \mathbb{R}^2 com \mathbb{C} , obtemos $\gamma(t) = \gamma^1(t) + i\gamma^2(t)$.

Escrevendo γ em termos de séries de Fourier, obtemos

$$\gamma(t) = \sum_{k \in \mathbb{Z}} c_k e^{2\pi i k t},$$

onde

$$c_k = \int_0^1 \gamma(t) e^{-2\pi i k t} dt.$$

Pela identidade de Plancherel, obtemos

$$\int_0^1 |\gamma(t)|^2 dt = \sum_{k \in \mathbb{Z}} |c_k|^2.$$

Temos que

$$\dot{\gamma}(t) = \sum_{k \in \mathbb{Z}} 2\pi i k c_k e^{2\pi i k t},$$

$$v^2 = \int_0^1 |\dot{\gamma}(t)|^2 dt = \sum_{k \in \mathbb{Z}} (2\pi k)^2 |c_k|^2.$$

Note que

$$\begin{aligned} \gamma(t)\overline{\dot{\gamma}(t)} &= (\gamma^1(t) + i\gamma^2(t))(\dot{\gamma}^1(t) - i\dot{\gamma}^2(t)) = (\gamma^1(t)\dot{\gamma}^1(t) + \gamma^2(t)\dot{\gamma}^2(t)) + i(-\gamma^1(t)\dot{\gamma}^2(t) + \gamma^2(t)\dot{\gamma}^1(t)), \\ \operatorname{Re} \left(i\gamma(t)\overline{\dot{\gamma}(t)} \right) &= (\gamma^1(t)\dot{\gamma}^2(t) - \gamma^2(t)\dot{\gamma}^1(t)) \end{aligned}$$

A menos de translação, podemos supor que $c_0 = 0$ ao substituir γ por $\gamma - c_0$.

Portanto,

$$(2\pi)^2 \int_0^1 |\gamma(t)|^2 dt = (2\pi)^2 \sum_{k \in \mathbb{Z}} |c_k|^2 \leq \sum_{k \in \mathbb{Z}} (2\pi k)^2 |c_k|^2 = \int_0^1 |\dot{\gamma}(t)|^2 dt = v^2$$

comparando as séries acima.

No espaço $C([0, 1], \mathbb{C})$ temos o produto interno

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)} dt$$

e pela desigualdade de Cauchy-Schwarz

$$2m(D) = \operatorname{Re} \left(\int_0^1 i\gamma(t)\overline{\dot{\gamma}(t)} dt \right) \leq \left| \int_0^1 i\gamma(t)\overline{\dot{\gamma}(t)} dt \right| \leq \left(\int_0^1 |i\gamma(t)|^2 dt \right)^{\frac{1}{2}} \left(\int_0^1 |\dot{\gamma}(t)|^2 dt \right)^{\frac{1}{2}} \leq \frac{v^2}{2\pi}.$$

Portanto,

$$2m(D) \leq \frac{v^2}{2\pi}.$$

O perímetro de γ é $L = \int_0^1 |\dot{\gamma}(t)| dt = v$, ou seja,

$$m(D) \leq \frac{L^2}{4\pi} \quad (\textbf{Desigualdade isoperimétrica}).$$

A desigualdade acima se torna igualdade quando

$$(2\pi)^2 \int_0^1 |\gamma(t)|^2 dt = \int_0^1 |\dot{\gamma}(t)|^2 dt,$$

ou seja, $c_k = 0$ para $|k| \geq 2$, e $i\gamma(t) = \lambda\dot{\gamma}(t)$ para uma constante $\lambda \in \mathbb{C}$ para que a desigualdade de Cauchy-Schwarz se torne uma igualdade.

Assim,

$$\gamma(t) = c_{-1}e^{-2\pi it} + c_1e^{2\pi it}$$

e

$$ic_{-1}e^{-2\pi it} + ic_1e^{2\pi it} = \lambda(-2\pi ic_{-1}e^{-2\pi it} + 2\pi ic_1e^{2\pi it}),$$

ou seja,

$$c_{-1} = -2\pi\lambda c_{-1},$$

$$c_1 = 2\pi\lambda c_1.$$

Logo, $c_1 = 0$ ou $c_{-1} = 0$. Portanto, γ é um círculo. Como γ percorre no sentido anti-horário, devemos ter $c_{-1} = 0$ e $\gamma(t) = c_1e^{2\pi it}$.

Referências Bibliográficas

[Aur] Leandro F. Aurichi. Espaços métricos.

<https://sites.icmc.usp.br/aurichi/dokuwiki/doku.php?id=curso:metricos2021>

[Imb] Roberto Imbuzeiro Oliveira. Análise no \mathbb{R}^n . Notas de aula.

https://www.dropbox.com/s/1tq3zosp5s7lm5n/notas_Rn_2021.pdf?e=1&dl=0

[Lag1] Elon Lages Lima. Análise no espaço \mathbb{R}^n . Coleção Matemática Universitária, Rio de Janeiro, IMPA, 2004.

[Lag2] Elon Lages Lima. Análise real. Vol. 2 - Coleção Matemática Universitária, Rio de Janeiro, IMPA, 2004.

[Lan] Serge Lang. Undergraduate analysis. New York: Springer - Verlag, 1983.

[Spi] Michael Spivak. Calculus on Manifolds. CRC Press, 1^a edição. 1965.