

Capítulo 3

Funções reais de n variáveis

1 Derivadas parciais

Definição 1.1. Seja $f : U \rightarrow \mathbb{R}$ uma função real definida num subconjunto aberto $U \subset \mathbb{R}^n$. Dado $a \in U$, a *i -ésima derivada parcial de f no ponto a* , $1 \leq i \leq n$, é o limite

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

quando tal limite existe. Usa-se também a notação $\partial_i f(a)$.

Observação 1.1. Dados o ponto $a \in U$ e $i \in \{1, \dots, n\}$, a imagem do caminho de classe C^∞ $\lambda : \mathbb{R} \rightarrow \mathbb{R}^n$, $\lambda(t) = a + te_i$, é a reta que passa por a e é paralela ao i -ésimo eixo. Como U é aberto e $a \in U$, existe $\varepsilon > 0$ tal que $t \in (-\varepsilon, \varepsilon) \implies \lambda(t) = a + te_i \in U$.

A i -ésima derivada parcial de f no ponto a é, portanto, a derivada da função $f \circ \lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ no ponto $t = 0$, ou seja, $\frac{\partial f}{\partial x_i}(a) = (f \circ \lambda)'(0)$, pois

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

• Assim, o cálculo prático da i -ésima derivada parcial de uma função real $f(x_1, \dots, x_n)$ se faz considerando todas as variáveis como se fossem constantes, exceto a i -ésima, e aplicando as regras usuais de derivação em relação a esta variável.

Observação 1.2. Quando $n = 2$, o gráfico de f , $G = \{(x, y, f(x, y)) \mid (x, y) \in \text{Dom}(f)\}$ é uma "superfície" em $\mathbb{R}^3$, e a restrição de f ao segmento de reta que passa por $c = (a, b)$ e é paralelo ao eixo das abscissas tem como gráfico uma curva plana $x \mapsto (x, b, f(x, b))$ obtida na superfície fazendo y constante igual a b . Portanto, $\frac{\partial f}{\partial x}(a, b)$ é a inclinação da reta tangente a esta curva, no ponto $(a, b, f(a, b))$, em relação ao plano horizontal, uma vez que:

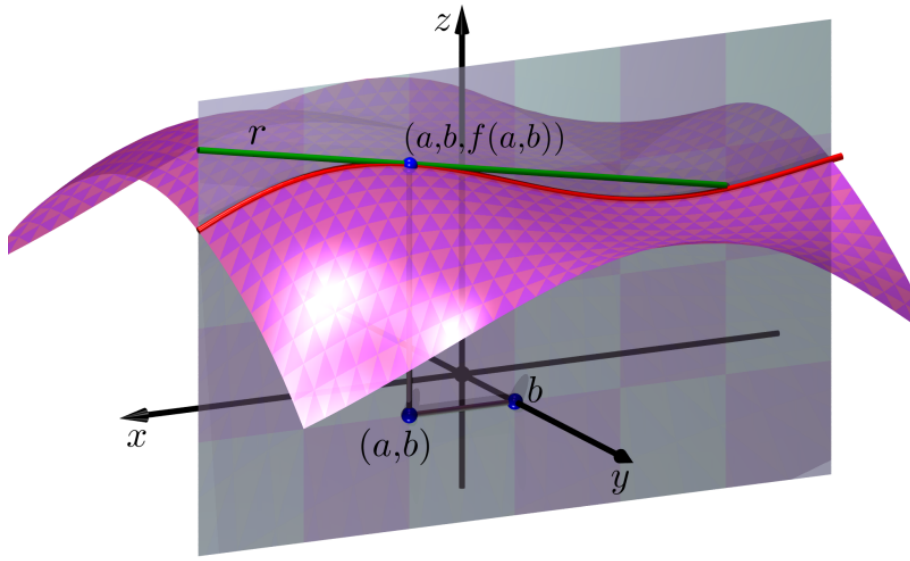


Fig. 1: $\frac{\partial f}{\partial x}(a, b)$ é a inclinação da reta r

$$r = \left\{ \left(1, 0, \frac{\partial f}{\partial x}(a, b) \right) t + (a, b, f(a, b)) \mid t \in \mathbb{R} \right\} = \left\{ \left(x, b, \frac{\partial f}{\partial x}(a, b)(x - a) + f(a, b) \right) \mid x \in \mathbb{R} \right\}.$$

Observação 1.3. A i -ésima derivada parcial $\frac{\partial f}{\partial x_i}$ dá informações sobre o comportamento de f ao longo de um segmento de reta contido em U e paralelo ao i -ésimo eixo.

- Por exemplo, se $f : U \rightarrow \mathbb{R}$ está definida num aberto $U \subset \mathbb{R}^2$, $J = \{(a, t) \mid t \in [0, 1]\} \subset U$ e $\frac{\partial f}{\partial y}(a, t) > 0$ para todo $t \in [0, 1]$, então f é crescente ao longo de J , ou seja, $0 \leq s < t \leq 1 \implies f(a, s) < f(a, t)$.

Definição 1.2. Dizemos que uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ *não depende da i -ésima variável* quando $a, b \in U$, $b = a + te_i \implies f(a) = f(b)$.

Neste caso, existe $\frac{\partial}{\partial x_i} f(a)$ em todos os pontos $a \in U$ e é igual a zero. Mas a recíproca nem sempre é verdadeira, como veremos abaixo.

Definição 1.3. Um conjunto $U \subset \mathbb{R}^n$ é chamado *i -convexo* ($1 \leq i \leq n$) quando:

$$a, b \in U, b = a + te_i \implies [a, b] = \{a + se_i \mid s \in [0, t]\} \subset U.$$

- Assim, se $U \subset \mathbb{R}^n$ é um aberto i -convexo e $f : U \rightarrow \mathbb{R}$ é uma função tal que $\frac{\partial f}{\partial x_i}(a) = 0$ para todo $a \in U$, então f independe da i -ésima variável.

De fato, se $a, b \in U$, $b = a + t_0 e_i$, então $\lambda(s) = a + se_i \in U$, para todo $s \in [0, t_0]$, e, portanto, existe $\varepsilon > 0$ tal que $\lambda(s) \in U$ para todo $s \in (-\varepsilon, t_0 + \varepsilon)$.

Além disso, como $f \circ \lambda$ é derivável em $(-\varepsilon, t_0 + \varepsilon)$ e $(f \circ \lambda)'(s) = \frac{\partial f}{\partial x_i}(a + se_i) = 0$ para todo $s \in (-\varepsilon, t_0 + \varepsilon)$, então $f \circ \lambda(s) = f \circ \lambda(0)$ para todo $s \in (-\varepsilon, t_0 + \varepsilon)$. Logo $f(b) = f(a)$.

Observação 1.4. Em $\mathbb{R}^2$, dizemos *horizontalmente* e *verticalmente convexo*, em vez de 1-convexo e 2-convexo, respectivamente.

Exemplo 1.1. Seja $\Gamma = \{(x, 0) \in \mathbb{R}^2 \mid x \geq 0\}$ o semi-eixo positivo fechado das abscissas. Então $U = \mathbb{R}^2 - \Gamma$ é aberto, horizontalmente convexo, mas não é verticalmente convexo.

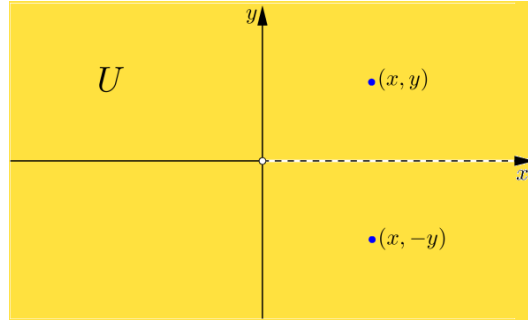


Fig. 2: $U = \mathbb{R}^2 - \Gamma$

Seja $f : U \rightarrow \mathbb{R}$ a função definida por $f(x, y) = x^2$, se $x > 0$ e $y > 0$, e $f(x, y) = 0$, se $x \leq 0$ ou $y \leq 0$. Então f possui derivada parcial $\frac{\partial f}{\partial y}(p) = 0$ para todo ponto $p \in U$, pois:

- $f|_{r_0^+} \equiv 0$, onde $r_0^+ = \{(0, t) \mid t > 0\}$;
- $f|_{r_0^-} \equiv 0$, onde $r_0^- = \{(0, t) \mid t < 0\}$;
- $f|_{r_{x_0}} \equiv 0$, onde $r_{x_0} = \{(x_0, t) \mid t \in \mathbb{R}\}$ e $x_0 < 0$;
- $f|_{r_{x_0}^+} \equiv x_0^2$ e $f|_{r_{x_0}^-} \equiv 0$, onde $r_{x_0}^+ = \{(x_0, t) \mid t > 0\}$, $r_{x_0}^- = \{(x_0, t) \mid t < 0\}$ e $x_0 > 0$.

Mas f não é independente da segunda variável, pois se $x > 0$ e $y > 0$, então $f(x, y) = x^2 > 0$ e $f(x, -y) = 0$. $\square$

Observação 1.5. A existência apenas das derivadas parciais não permite conclusões sobre o comportamento *n-dimensional* da função. Por exemplo, a existência de todas as derivadas parciais num ponto não implica a continuidade da função nesse ponto.

Exemplo 1.2. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por $f(x, y) = \frac{xy}{x^2 + y^2}$, se $(x, y) \neq (0, 0)$, e $f(0, 0) = 0$.

Se $z = (x, y) \neq (0, 0)$, temos que:

$$\frac{\partial f}{\partial x}(z) = \frac{y(x^2 + y^2) - xy(2x)}{(x^2 + y^2)^2} = \frac{y^3 - x^2y}{(x^2 + y^2)^2} \quad \text{e} \quad \frac{\partial f}{\partial y}(z) = \frac{x(x^2 + y^2) - xy(2y)}{(x^2 + y^2)^2} = \frac{x^3 - xy^2}{(x^2 + y^2)^2}.$$

E, na origem:

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

Assim, f possui derivadas parciais em todos os pontos de $\mathbb{R}^2$. Mas f não é contínua na origem.

Mais ainda, não existe $\lim_{(x,y) \rightarrow (0,0)} f(x,y)$, pois $f(at, bt) = \frac{ab}{a^2 + b^2}$, para todo $t \in \mathbb{R}$ e todo $(a, b) \neq (0, 0)$, e, portanto, $\lim_{t \rightarrow 0} f(t, t) = \frac{1}{2} \neq \frac{2}{5} = \lim_{t \rightarrow 0} f(t, 2t)$, por exemplo. $\square$

2 Derivadas direcionais

Definição 2.1. Sejam $f : U \rightarrow \mathbb{R}$ uma função definida no aberto $U \subset \mathbb{R}^n$, $a \in U$ e $v \in \mathbb{R}^n$. A *derivada direcional de f no ponto a segundo o vetor v* é o limite:

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

quando tal limite existe.

Observação 2.1. Se $v = 0$, então $\frac{\partial f}{\partial v}(a) = 0$ para todo $a \in U$.

Observação 2.2. As derivadas parciais são casos particulares das derivadas direcionais, pois: $\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i}(a)$ é a derivada direcional de f no ponto a segundo o vetor e_i .

Observação 2.3. Dados $a \in U$ e $v \in \mathbb{R}^n$, existe $\varepsilon > 0$ tal que $a + tv \in U$ para todo $t \in (-\varepsilon, \varepsilon)$. Assim, se $\lambda : (-\varepsilon, \varepsilon) \rightarrow U$ é o caminho retilíneo, com $\lambda(0) = a$ e $\lambda'(t) = v$ para todo $t \in (-\varepsilon, \varepsilon)$, temos que: $\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$.

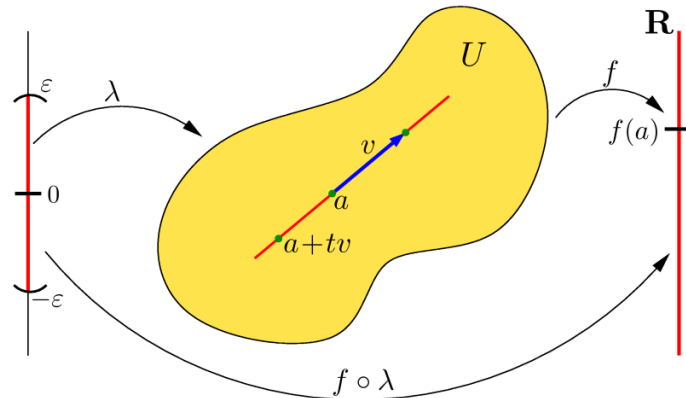


Fig. 3: f ao longo do caminho retilíneo λ

Exemplo 2.1. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função dada por $f(x, y) = \frac{xy}{x^2 + y^2}$, $(x, y) \neq (0, 0)$, e $f(0, 0) = 0$. Então f possui as derivadas direcionais $\frac{\partial f}{\partial v}(0, 0)$ para todo $v = (\alpha, 0)$ ou $v = (0, \beta)$, as quais são nulas, mas f não possui derivada direcional na origem segundo um vetor $v = (\alpha, \beta)$, com $\alpha \neq 0$ e $\beta \neq 0$, pois:

- $\frac{\partial f}{\partial v}(0,0) = \lim_{t \rightarrow 0} \frac{f(t\alpha, 0) - f(0,0)}{t} = 0, \quad v = (\alpha, 0)$
- $\frac{\partial f}{\partial v}(0,0) = \lim_{t \rightarrow 0} \frac{f(0, t\beta) - f(0,0)}{t} = 0, \quad v = (0, \beta),$

e o limite

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

não existe. $\square$

Observação 2.4. Se $\alpha \in \mathbb{R} - \{0\}$, então existe $\frac{\partial f}{\partial v}(a)$ num ponto a se, e somente se, existe $\frac{\partial f}{\partial(\alpha v)}(a)$ e, no caso afirmativo, temos:

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

Mas, pode ocorrer que a derivada direcional $\frac{\partial f}{\partial v}$ exista em todos os pontos do domínio de f , segundo todos os vetores $v \in \mathbb{R}^n$, sem que se tenha necessariamente:

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

Exemplo 2.2. Seja $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função dada por

$$g(x, y) = \frac{x^2 y}{x^2 + y^2}, \quad \text{se } (x, y) \neq (0, 0), \quad \text{e} \quad g(0, 0) = 0.$$

Pode-se provar, a partir da definição, que existe $\frac{\partial g}{\partial v}(a)$ para todo $a \in \mathbb{R}^2$ e todo $v \in \mathbb{R}^2$. Em particular, na origem:

$$\bullet \frac{\partial g}{\partial v}(0,0) = \lim_{t \rightarrow 0} \frac{g(t\alpha, t\beta) - g(0,0)}{t} = \frac{\alpha^2 \beta}{\alpha^2 + \beta^2}, \quad \text{se } v = (\alpha, \beta) \neq (0,0).$$

e

$$\bullet \frac{\partial g}{\partial v}(0,0) = 0, \quad \text{se } v = (0,0).$$

Evidentemente, para $a = (0,0)$, não vale

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

Por exemplo, para $v = (1,1)$ e $w = (1,2)$:

$$\frac{\partial g}{\partial v}(0,0) = \frac{1}{2}, \quad \frac{\partial g}{\partial w}(0,0) = \frac{2}{5}, \quad \text{e} \quad \frac{\partial g}{\partial(v+w)}(0,0) = \frac{12}{13},$$

e, portanto, $\frac{\partial g}{\partial v}(0,0) + \frac{\partial g}{\partial w}(0,0) \neq \frac{\partial g}{\partial(v+w)}(0,0)$. $\square$

Observação 2.5. Na seção 3, mostraremos que $\frac{\partial f}{\partial v}$ depende linearmente de v se f é *diferenciável*, uma hipótese mais forte do que possuir derivadas direcionais.

A função g do exemplo anterior é contínua (ver exercício 8:27, capítulo 1), mas não é verdade, em geral, que a existência de todas as derivadas direcionais implique em continuidade.

Exemplo 2.3. Seja $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $h(x, y) = \frac{x^3 y}{x^6 + y^2}$, se $(x, y) \neq (0, 0)$, e $h(0, 0) = 0$.

Para $(a, b) \neq (0, 0)$ e $v = (\alpha, \beta) \in \mathbb{R}^2$, temos que, se $\lambda(t) = (a, b) + t(\alpha, \beta) = (a + t\alpha, b + t\beta)$, então:

$$(h \circ \lambda)(t) = \frac{(a + t\alpha)^3(b + t\beta)}{(a + t\alpha)^6 + (b + t\beta)^2}$$

e, portanto, a derivada $(h \circ \lambda)'(t)$ é dada por:

$$\frac{(3(a + t\alpha)^2\alpha(b + t\beta) + \beta(a + t\alpha)^3)((a + t\alpha)^6 + (b + t\beta)^2) - (a + t\alpha)^3(b + t\beta)(6\alpha(a + t\alpha)^5 + 2\beta(b + t\beta))}{((a + t\alpha)^6 + (b + t\beta)^2)^2}.$$

Logo,

$$\begin{aligned} \frac{\partial h}{\partial v}(a, b) = (h \circ \lambda)'(0) &= \frac{(3a^2b\alpha + \beta a^3)(a^6 + b^2) - a^3b(6\alpha a^5 + 2\beta b)}{(a^6 + b^2)^2} \\ &= \left(\frac{-3a^8b + 3a^2b^3}{(a^6 + b^2)^2} \right) \alpha + \left(\frac{a^9 - a^3b^2}{(a^6 + b^2)^2} \right) \beta, \end{aligned}$$

E para $(a, b) = (0, 0)$ e $v = (\alpha, \beta) \in \mathbb{R}^2$,

$$\frac{\partial h}{\partial v}(0, 0) = \lim_{t \rightarrow 0} \frac{h(t\alpha, t\beta)}{t} = \lim_{t \rightarrow 0} \frac{t^4 \alpha^3 \beta}{t(t^6 \alpha^6 + t^2 \beta^2)} = \lim_{t \rightarrow 0} \frac{t \alpha^3 \beta}{t^4 \alpha^6 + \beta^2} = 0, \quad \text{se } \beta \neq 0;$$

$$\text{e } \frac{\partial h}{\partial v}(0, 0) = \lim_{t \rightarrow 0} \frac{h(t\alpha, 0)}{t} = \lim_{t \rightarrow 0} 0 = 0, \quad \text{se } \beta = 0.$$

Assim, existem as derivadas direcionais $\frac{\partial h}{\partial v}(a)$, para todo $a \in \mathbb{R}^2$ e todo $v \in \mathbb{R}^2$, e dependem linearmente de v .

Em $\mathbb{R}^2 - \{(0, 0)\}$, a função h é contínua, mas h não é contínua na origem, pois $h(x, x^3) = \frac{1}{2}$ para todo $x \neq 0$. $\square$

• Outra propriedade desejável para um conceito adequado de derivada de uma função é que a composta de duas funções deriváveis seja também derivável.

Exemplo 2.4. Seja $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $\varphi(0, 0) = 0$ e $\varphi(x, y) = \frac{x^3 y}{x^4 + y^2}$, se $(x, y) \neq (0, 0)$.

Em $\mathbb{R}^2 - \{(0, 0)\}$, φ é contínua, e em $(0, 0)$, φ também é contínua, pois, para $(x, y) \neq (0, 0)$,

$$|\varphi(x, y)| = \left| x \frac{x^2}{\sqrt{x^4 + y^2}} \frac{y}{\sqrt{x^4 + y^2}} \right| \leq |x|,$$

e, portanto, $\lim_{(x, y) \rightarrow (0, 0)} \varphi(x, y) = 0$.

Além disso, para todo $v = (\alpha, \beta) \in \mathbb{R}^2$, $\beta \neq 0$,

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

para $v = (\alpha, 0) \in \mathbb{R}^2$.

Portanto, todas as derivadas direcionais existem na origem e dependem linearmente de v . De modo análogo ao exemplo anterior, podemos calcular as derivadas direcionais de φ num ponto $(a, b) \in \mathbb{R}^2 - \{(0, 0)\}$ e verificar que elas dependem linearmente de v .

Entretanto, se considerarmos o caminho derivável $\lambda : \mathbb{R} \rightarrow \mathbb{R}^2$, dado por $\lambda(t) = (t, t^2 \sin \frac{1}{t})$, se $t \neq 0$, $\lambda(0) = (0, 0)$, temos que $f \circ \lambda : \mathbb{R} \rightarrow \mathbb{R}$ não é derivável em $t = 0$, pois o limite

$$\lim_{t \rightarrow 0} \frac{\varphi(\lambda(t)) - \varphi(\lambda(0))}{t} = \lim_{t \rightarrow 0} \frac{\varphi(t, t^2 \sin \frac{1}{t})}{t} = \lim_{t \rightarrow 0} \frac{t^5 \sin \frac{1}{t}}{t^5 + t^5 \sin \frac{1}{t}} = \lim_{t \rightarrow 0} \frac{\sin \frac{1}{t}}{1 + \sin \frac{1}{t}},$$

não existe, uma vez que:

$$\begin{aligned} & \bullet \lim_{n \rightarrow \infty} \frac{\sin \frac{1}{t_n}}{1 + \sin \frac{1}{t_n}} = \lim_{n \rightarrow \infty} 0 = 0, \text{ quando } t_n = \frac{1}{n\pi}, \\ \text{e} & \bullet \lim_{n \rightarrow \infty} \frac{\sin \frac{1}{t_n}}{1 + \sin \frac{1}{t_n}} = \lim_{n \rightarrow \infty} \frac{1}{2} = \frac{1}{2}, \text{ quando } t_n = \frac{2}{(4n+1)\pi}. \quad \square \end{aligned}$$

• No entanto, a existência de derivadas direcionais permite demonstrar o Teorema do Valor Médio para funções reais de n variáveis sob a forma de igualdade, como no caso de uma só variável.

Teorema 2.1. (do Valor Médio)

Seja $f : U \rightarrow \mathbb{R}$ uma função definida no aberto $U \subset \mathbb{R}^n$. Se $[a, a+v] \subset U$, $f|_{[a, a+v]}$ é contínua e existe a derivada direcional $\frac{\partial f}{\partial v}(x)$ para todo $x \in (a, a+v)$, então existe $\theta_0 \in (0, 1)$ tal que

$$f(a+v) - f(a) = \frac{\partial f}{\partial v}(a + \theta_0 v)$$

Prova.

Seja $\lambda : [0, 1] \rightarrow U$ o caminho C^∞ dado por $\lambda(t) = a + tv$, $t \in [0, 1]$. Então a função $f \circ \lambda : [0, 1] \rightarrow \mathbb{R}$ é contínua em $[0, 1]$ e derivável em $(0, 1)$, pois, para $\theta \in (0, 1)$,

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

Assim, pelo Teorema do Valor Médio, para funções reais de uma variável real, existe $\theta_0 \in (0, 1)$ tal que $(f \circ \lambda)(1) - (f \circ \lambda)(0) = (f \circ \lambda)'(\theta_0)$, ou seja, existe $\theta_0 \in (0, 1)$ tal que

$$f(a+v) - f(a) = \frac{\partial f}{\partial v}(a + \theta_0 v). \quad \blacksquare$$

Observação 2.6. A existência de $\frac{\partial f}{\partial v}$ em todo ponto de $(a, a+v)$ garante a continuidade de $f|_{(a, a+v)}$.

De fato, como foi provado acima, $f \circ \lambda$ é derivável em $(0, 1)$ e, portanto, se $x_k = a + t_k v$, $t_k \in (0, 1)$, é uma sequência de pontos de $(a, a+v)$ que converge para o ponto $a + t_0 v \in (a, a+v)$, então

$$f(x_k) = f(a + t_k v) = f \circ \lambda(t_k) \longrightarrow f \circ \lambda(t_0) = f(a + t_0 v),$$

uma vez que $t_k = \frac{\|x_k - a\|}{\|v\|} \longrightarrow \frac{\|a + t_0 v - a\|}{\|v\|} = t_0$.

Corolário 2.1. *Seja $U \subset \mathbb{R}^n$ aberto e conexo. Se $f : U \longrightarrow \mathbb{R}$ possui derivadas direcionais em todo ponto $x \in U$ e $\frac{\partial f}{\partial v}(x) = 0$, para todo $x \in U$ e todo $v \in \mathbb{R}^n$, então f é constante.*

Prova.

Seja $a \in U$ fixo.

Afirmção: se $[a, b] \subset U$, então $f|_{[a, b]}$ é contínua.

- De fato, como $a, b \in U$ e U é aberto, existe $\varepsilon > 0$ tal que o segmento

$$(a - \varepsilon(b - a), a + (1 + \varepsilon)(b - a)) = \{a + t(b - a) \mid t \in (-\varepsilon, 1 + \varepsilon)\}$$

está contido em U .

Além disso, como existe $\frac{\partial f}{\partial(b-a)}(x)$ para todo $x \in U$, temos, pela observação anterior, que a restrição $f|_{(a-\varepsilon(b-a), a+(1+\varepsilon)(b-a))}$ é contínua.

Portanto, $f|_{[a, b]}$ é contínua.

- Resulta, então, do Teorema do Valor Médio, que se $[a, b] \subset U$, existe $\theta_0 \in (0, 1)$ tal que

$$f(b) - f(a) = f(a + (b - a)) - f(a) = \frac{\partial f}{\partial(b-a)}(a + \theta_0(b - a)) = 0,$$

ou seja, $f(b) = f(a)$.

Por outro lado, se $x \in U$ existe, pelo teorema 13.8 do Capítulo 1, uma poligonal contida em U com vértices $a_0 = a, a_1, \dots, a_k = x$.

Temos, então, sucessivamente, que

$$f(a) = f(a_0) = f(a_1) = \dots = f(a_k) = f(x),$$

ou seja, $f(x) = f(a)$ para todo $x \in U$. Logo f é constante. ■

Observação 2.7. Neste corolário, basta que as derivadas parciais $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, existam e sejam nulas em todos os pontos do aberto conexo $U \subset \mathbb{R}^n$, pois, pela observação 13.5 do Capítulo 1, dados $a, b \in U$, existe uma poligonal contida em U ligando os pontos a e b com lados paralelos a um dos eixos coordenados.

3 Funções diferenciáveis

A definição de função diferenciável que daremos abaixo é devida a *Maurice Fréchet* (França, 1878-1973) e *Otto Stolz* (Áustria, 1842-1905). Ela é uma extensão adequada do conceito de função derivável de uma só variável para funções de n variáveis.

Definição 3.1. Seja $f : U \rightarrow \mathbb{R}$ uma função definida no aberto $U \subset \mathbb{R}^n$. Dizemos que f é *diferenciável no ponto* $a \in U$ quando existem constantes $A_1, \dots, A_n \in \mathbb{R}$ tais que, para todo vetor $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, com $a + v \in U$, temos que:

$$f(a + v) = f(a) + A_1\alpha_1 + \dots + A_n\alpha_n + r(v),$$

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

Definição 3.2. Dizemos que $f : U \rightarrow \mathbb{R}$ é *diferenciável* quando f é diferenciável em todos os pontos de U .

Observação 3.1. Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável no ponto a . Então, se $v = te_i$, ou seja, $\alpha_j = 0, j \neq i, \alpha_i = t$, temos que

$$\frac{f(a + te_i) - f(a)}{t} = A_i + \frac{r(te_i)}{t} = A_i \pm \frac{r(te_i)}{\|te_i\|}, \quad i = 1, \dots, n.$$

Logo, como $\lim_{t \rightarrow 0} \frac{r(te_i)}{\|te_i\|} = 0$, para todo $i = 1, \dots, n$, obtemos que a derivada parcial $\frac{\partial f}{\partial x_i}(a)$ existe e é igual a A_i , para todo $i = 1, \dots, n$.

• Assim, $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$ se, e só se, as derivadas parciais $\frac{\partial f}{\partial x_i}(a), i = 1, \dots, n$, existem, e para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ tal que $a + v \in U$, temos

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

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

Observação 3.2. Se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$, então f é contínua no ponto a .

De fato, como $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|}$ implica que $\lim_{v \rightarrow 0} r(v) = \lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} \|v\| = 0$, temos que

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} (f(a) + A_1(x_1 - a_1) + \dots + A_n(x_n - a_n) + r(x - a)) = f(a),$$

uma vez que $v = x - a \rightarrow 0$ quando $x \rightarrow a$.

Observação 3.3. A condição $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0$ significa que $r(v)$ tende a zero mais rapidamente

do que v . Isto se exprime dizendo-se que $r(v)$ é um *infinitésimo de ordem superior a v* . Assim, f é diferenciável no ponto $a \in U$ quando $f(a + v) - f(a)$ é igual a um funcional linear

$$\sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \alpha_i + (\text{um resto infinitamente pequeno em relação a } v).$$

Observação 3.4. Fazendo $\rho(v) = \frac{r(v)}{\|v\|}$ se $v \neq 0$, $a + v \in U$, e $\rho(0) = 0$, temos que:

$f : U \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$ se, e só se, todas as derivadas parciais $\frac{\partial f}{\partial x_i}(a)$, $i = 1, \dots, n$, existem no ponto a e, para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ tal que $a + v \in U$ vale:

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

Ou seja, f é diferenciável no ponto $a \in U$ se, e só se, a função real

$$\rho : V_a = \{v \in \mathbb{R}^n \mid a + v \in U\} \rightarrow \mathbb{R}$$

é contínua no ponto $v = 0$. Note que o conjunto V_a é aberto em $\mathbb{R}^n$ e $0 \in V_a$.

Observação 3.5. Ser ou não ser diferenciável, independe da norma considerada em $\mathbb{R}^n$.

Observação 3.6. Para funções $f : I \rightarrow \mathbb{R}$ definidas num intervalo aberto $I \subset \mathbb{R}$, diferenciabilidade é o mesmo que derivabilidade, pois se $f(a + t) = f(a) + At + \rho(t) |t|$, ou seja,

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

então, $\lim_{t \rightarrow 0} \rho(t) = 0$ se, e só se, f é derivável no ponto a e $f'(a) = A$.

Observação 3.7. Se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$, então f possui derivada direcional no ponto a segundo qualquer vetor $v = (\alpha_1, \dots, \alpha_n)$ e

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

De fato, seja $v \in \mathbb{R}^n$. Então existe $\varepsilon > 0$ tal que $a + tv \in U$ para todo $t \in (-\varepsilon, \varepsilon)$, e

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

Como $\lim_{t \rightarrow 0} \rho(tv) = 0$, temos que

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \alpha_i + \lim_{t \rightarrow 0} (\pm \rho(tv) \|v\|) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \alpha_i.$$

Então $\frac{\partial f}{\partial v}(a)$ existe e depende linearmente de v , ou seja:

- $\frac{\partial f}{\partial(\alpha v)}(a) = \alpha \frac{\partial f}{\partial v}(a)$, para todo $\alpha \in \mathbb{R}$ e $v \in \mathbb{R}^n$,

e

$$\bullet \frac{\partial f}{\partial(v+w)}(a) = \frac{\partial f}{\partial v}(a) + \frac{\partial f}{\partial w}(a), \text{ para todos } v, w \in \mathbb{R}^n.$$

Teorema 3.1. (Regra da cadeia)

Sejam $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ abertos, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ tal que $f(U) \subset V$ e cada função coordenada $f_i : U \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$. Se $g : V \rightarrow \mathbb{R}$ é diferenciável no ponto $b = f(a)$, então a função composta $g \circ f : U \rightarrow \mathbb{R}$ é diferenciável no ponto a e suas derivadas parciais são:

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{k=1}^n \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_i}(a), \quad i = 1, \dots, m.$$

Prova.

Seja o aberto $U_0 = \{v \in \mathbb{R}^m \mid a + v \in U\}$ que contém o ponto $v = 0$.

Para cada $v = (\alpha_1, \dots, \alpha_m) \in U_0$ e $k = 1, \dots, n$, temos que

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

onde cada $\rho_k : U_0 \rightarrow \mathbb{R}$ é contínua no ponto 0 e $\rho_k(0) = 0$.

Seja a aplicação $\omega = (\beta_1, \dots, \beta_n) : U_0 \rightarrow \mathbb{R}^n$ contínua no ponto 0, com $\omega(0) = 0$, cujas funções coordenadas $\beta_k : U_0 \rightarrow \mathbb{R}$ são dadas por:

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

Considerando $\mathbb{R}^m$ com a norma da soma, por exemplo, temos que $\frac{|\alpha_i|}{\|v\|_s} \leq 1$ para todo $v \in \mathbb{R}^m - \{0\}$.

Logo, cada $\frac{|\beta_k(v)|}{\|v\|_s}$, $k = 1, \dots, n$, e, portanto, $\frac{\|\omega(v)\|_s}{\|v\|_s}$, é limitada em $U_1 - \{0\}$, onde U_1 é um aberto contido em U_0 tal que $0 \in U_1$ e $\rho_k|_{U_1}$ limitada para todo $k = 1, \dots, n$.

Seja $V_0 = \{w \in \mathbb{R}^n \mid w + b \in V\}$. Como V_0 é um aberto que contém o vetor 0, ω é contínua no ponto 0 e $\omega(0) = 0$, existe um aberto $U_2 \subset U_1$ tal que $0 \in U_2$ e $\omega(U_2) \subset V_0$.

Seja $v \in U_2$. Então $\omega(v) + b \in V$ e, como $g : V \rightarrow \mathbb{R}$ é diferenciável em $b = f(a)$, temos, por (I), que

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

onde $\sigma \circ \omega : U_2 \rightarrow \mathbb{R}$ é uma função contínua no ponto 0, com $\sigma \circ \omega(0) = 0$.

Logo, por (II),

$$(g \circ f)(a + v) = g \circ f(a) + \sum_{k=1}^n \frac{\partial g}{\partial y_k}(b) \left[\sum_{i=1}^m \frac{\partial f_k}{\partial x_i}(a) \alpha_i + \rho_k(v) \|v\| \right] + \sigma \circ \omega(v) \|\omega(v)\|,$$

ou seja,

$$(g \circ f)(a + v) = (g \circ f)(a) + \sum_{i=1}^m A_i \alpha_i + R(v),$$

$$\text{onde } A_i = \sum_{k=1}^n \frac{\partial g}{\partial y_k}(b) \frac{\partial f_k}{\partial x_i}(a) \text{ e } R(v) = \sum_{k=1}^n \frac{\partial g}{\partial y_k}(b) \rho_k(v) \|v\| + \sigma \circ \omega(v) \|\omega(v)\|.$$

Como,

$$\frac{R(v)}{\|v\|} = \sum_{k=1}^n \frac{\partial g}{\partial y_k}(b) \rho_k(v) + \sigma \circ \omega(v) \frac{\|\omega(v)\|}{\|v\|},$$

temos que $\lim_{v \rightarrow 0} \frac{R(v)}{\|v\|} = 0$, pois $\lim_{v \rightarrow 0} \rho_k(v) = 0$, $k = 1, \dots, n$, $\lim_{v \rightarrow 0} \sigma \circ \omega(v) = 0$ e $\frac{\|\omega(v)\|}{\|v\|}$ é limitado em $U_2 - \{0\}$.

Logo $g \circ f$ é diferenciável no ponto a e

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

para todo $i = 1, \dots, m$. ■

Corolário 3.1. Se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável no ponto $b \in U$ e $\lambda = (\lambda_1, \dots, \lambda_n) : (a - \varepsilon, a + \varepsilon) \rightarrow \mathbb{R}^n$ é um caminho diferenciável com $\lambda(a) = b$, então a função composta $f \circ \lambda : (a - \varepsilon, a + \varepsilon) \rightarrow \mathbb{R}$ é diferenciável no ponto a e

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

Observação 3.8. Se escrevemos $\lambda(t) = (x_1(t), \dots, x_n(t))$, então $\lambda'(t) = \left(\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right)$. Indicando com $\frac{df}{dt}$ a derivada da função composta $t \mapsto f \circ \lambda(t) = f(x_1(t), \dots, x_n(t))$, a regra da cadeia nos dá que:

$$\frac{df}{dt} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \quad (\text{notação clássica do Cálculo Diferencial.})$$

Corolário 3.2. Sejam $U \subset \mathbb{R}^n$ um conjunto aberto, $I \subset \mathbb{R}$ um intervalo aberto, $f : U \rightarrow \mathbb{R}$ uma função diferenciável no ponto $a \in U$, com $f(U) \subset I$, $g : I \rightarrow \mathbb{R}$ diferenciável no ponto $b = f(a)$.

Então $g \circ f : U \rightarrow \mathbb{R}$ é diferenciável no ponto a e

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

para todo $i = 1, \dots, n$.

Observação 3.9. Pela Regra da Cadeia, se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$, para calcularmos a derivada direcional $\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$ não é necessário nos restringir ao

caminho retilíneo $\lambda(t) = a + tv$. Ou seja, se $\lambda : (-\varepsilon, \varepsilon) \rightarrow U$ é um caminho diferenciável qualquer com $\lambda(0) = a$ e $\lambda'(0) = v$, ainda teremos

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

De fato, pela Regra da Cadeia,

$$(f \circ \lambda)'(0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \lambda'_i(0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \alpha_i = \frac{\partial f}{\partial v}(a).$$

Mas, o mesmo não é verdade se f possui derivadas direcionais em todos os pontos do domínio segundo qualquer vetor, mas não é diferenciável.

Por exemplo, considere a função $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $h(x, y) = \frac{x^3 y}{x^6 + y^2}$, $(x, y) \neq (0, 0)$, e $h(0, 0) = 0$, e seja $\lambda : \mathbb{R} \rightarrow \mathbb{R}^2$ o caminho diferenciável, $\lambda(t) = (t, t^2)$, com $\lambda(0) = (0, 0)$ e $\lambda'(0) = (1, 0)$. Então,

$$(h \circ \lambda)'(0) = \lim_{t \rightarrow 0} \frac{h(\lambda(t)) - h(0)}{t} = \lim_{t \rightarrow 0} \frac{t^5}{t^7 + t^5} = \lim_{t \rightarrow 0} \frac{1}{t^2 + 1} = 1 \neq \frac{\partial h}{\partial x}(0, 0) = 0.$$

(ver exemplo 2.3).

Observação 3.10. Nenhuma das funções definidas nos exemplos 2.1, 2.2, 2.3 e 2.4:

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = \frac{xy}{x^2 + y^2}$, $f(0, 0) = 0$;
- $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g(x, y) = \frac{x^2 y}{x^2 + y^2}$, $g(0, 0) = 0$;
- $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(x, y) = \frac{x^3 y}{x^6 + y^2}$, $h(0, 0) = 0$;
- $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\varphi(x, y) = \frac{x^3 y}{x^4 + y^2}$, $\varphi(0, 0) = 0$,

são diferenciáveis na origem de $\mathbb{R}^2$.

De fato:

- f porque não é contínua na origem nem possui derivada direcional segundo qualquer vetor na origem.
- g porque, embora seja contínua na origem e existe $\frac{\partial g}{\partial v}(0, 0)$, para todo $v \in \mathbb{R}^2$, as derivadas direcionais na origem não dependem linearmente de v .
- h porque não é contínua na origem, embora possua derivadas direcionais $\frac{\partial h}{\partial v}(p)$, para todo $v \in \mathbb{R}^2$ e todo $p \in \mathbb{R}^2$, que dependem linearmente de v .
- φ é contínua em $\mathbb{R}^2$, possui derivadas direcionais $\frac{\partial \varphi}{\partial v}$ segundo qualquer vetor $v \in \mathbb{R}^2$, em todos os pontos do plano, que dependem linearmente de v , mas contraria a Regra da Cadeia, pois $\varphi \circ \lambda : \mathbb{R} \rightarrow \mathbb{R}$ não é derivável na origem, onde $\lambda : \mathbb{R} \rightarrow \mathbb{R}^2$ é o caminho diferenciável dado por

$$\lambda(t) = \left(t, t^2 \operatorname{sen} \frac{1}{t}\right), t \neq 0, \text{ e } \lambda(0) = 0.$$

• Diretamente, podemos verificar que, embora cada uma das funções acima possua derivadas parciais na origem, elas não cumprem a condição:

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

onde $v = (\alpha, \beta)$.

Por exemplo, para $F = f$, temos que $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$ e $f(\alpha, \beta) = \frac{\alpha\beta}{\alpha^2 + \beta^2}$. Logo, o limite

$$\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = \lim_{\substack{\alpha \rightarrow 0 \\ \beta \rightarrow 0}} \frac{1}{\sqrt{\alpha^2 + \beta^2}} \frac{\alpha\beta}{\alpha^2 + \beta^2}$$

não existe, já que para as sequências $(\alpha_n = \frac{1}{n})$ e $(\beta_n = \frac{1}{n})$, que convergem para zero, a sequência $\left(\frac{1}{\sqrt{\alpha_n^2 + \beta_n^2}} \frac{\alpha_n \beta_n}{\alpha_n^2 + \beta_n^2} \right) = \left(\frac{n}{2\sqrt{2}} \right)$ não converge.

Observação 3.11. Seja $U \subset \mathbb{C}$ aberto. Dizemos que uma função complexa $f : U \rightarrow \mathbb{C}$ é *derivável* no ponto $z = x + iy \in U$, quando existe o limite

$$\lim_{H \rightarrow 0} \frac{f(z+H) - f(z)}{H} = A.$$

Neste caso, $A = f'(z)$ chama-se a *derivada* da função complexa f no ponto z .

A derivabilidade de f no ponto $z = x + iy$ é equivalente a dizer que:

$$f(z+H) = f(z) + A H + r(H),$$

$$\text{onde } \lim_{H \rightarrow 0} \frac{r(H)}{H} = 0.$$

Fazendo $A = a + ib$, $H = h + ik$ e $r = r_1 + ir_2$, f é derivável no ponto $z = x + iy$ se, e só se,

$$f(z+H) = f(z) + (ah - bk) + i(bh + ak) + r_1(H) + ir_2(H), \quad (I)$$

$$\text{onde } \lim_{H \rightarrow 0} \frac{r_1(H)}{|H|} = \lim_{H \rightarrow 0} \frac{r_2(H)}{|H|} = 0.$$

Sejam $u, v : U \rightarrow \mathbb{R}$ a parte real e a parte imaginária da função f , ou seja, $f(z) = u(z) + iv(z)$.

Em (I), separando a parte real e a parte imaginária, temos que:

- $u(x+h, y+k) = u(x, y) + ah - bk + r_1(h, k)$, onde $\lim_{h,k \rightarrow 0} \frac{r_1(h, k)}{\sqrt{h^2 + k^2}} = 0$,
- $v(x+h, y+k) = v(x, y) + bh + ak + r_2(h, k)$, onde $\lim_{h,k \rightarrow 0} \frac{r_2(h, k)}{\sqrt{h^2 + k^2}} = 0$.

Assim, se $f = u + iv$ é derivável no ponto $z = x + iy$, então u e v são diferenciáveis no ponto (x, y) e valem as identidades: $\frac{\partial u}{\partial x}(x, y) = \frac{\partial v}{\partial y}(x, y) (= a)$ e $\frac{\partial u}{\partial y}(x, y) = -\frac{\partial v}{\partial x}(x, y) (= -b)$, chamadas *equações de Cauchy-Riemann*.

Reciprocamente, se $u, v : U \rightarrow \mathbb{R}$ são funções diferenciáveis no ponto $z = (x, y)$ e satisfazem as equações de Cauchy-Riemann neste ponto, podemos provar, revertendo cada etapa do argumento anterior, que a função complexa $f = u + iv$ é derivável no ponto $z = x + iy$ e que:

$$f'(z) = \frac{\partial u}{\partial x}(z) - i \frac{\partial u}{\partial y}(z) = \frac{\partial v}{\partial y}(z) + i \frac{\partial v}{\partial x}(z).$$

Uma função complexa $f : U \rightarrow \mathbb{C}$ é **holomorfa** quando possui derivada $f'(z)$ em todos os pontos do aberto U .

Definição 3.3. Seja $U \subset \mathbb{R}^n$ aberto. Dizemos que uma função $f : U \rightarrow \mathbb{R}$ é de **classe C^1** quando f possui derivadas parciais $\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)$ em todos os pontos $x \in U$ e as funções $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}, i = 1, \dots, n$, são contínuas.

Mais geralmente, dizemos que uma função $f : U \rightarrow \mathbb{R}$ é de **classe C^k** , $k \geq 1$, quando ela possui derivadas parciais em todos os pontos de U e as funções $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}, i = 1, \dots, n$, são de classe C^{k-1} . Para completar a definição indutiva, dizemos que f é de **classe C^0** quando f é contínua.

Finalmente, dizemos que f é de **classe C^∞** quando f é de classe C^k para todo $k \geq 0$.

Então $C^0 \supset C^1 \supset C^2 \supset \dots \supset C^k \supset \dots \supset C^\infty$, sendo todas as inclusões estritas (ver **Curso de Análise, Vol. I** de E. Lima, pag. 278, ex. 21).

Teorema 3.2. Se uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ possui derivadas parciais em todos os pontos do aberto U e cada uma delas é contínua no ponto $c \in U$, então f é diferenciável no ponto c .

Prova.

Para simplificar a notação, vamos considerar apenas o caso $n = 2$.

Sejam $c = (a, b)$ e $\delta > 0$ tal que $B_M(c, \delta) = (a - \delta, a + \delta) \times (b - \delta, b + \delta) \subset U$.

Seja $v = (h, k)$ um vetor tal que $c + v \in B_M(c, \delta) \subset U$ e

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

Reescrevendo, temos:

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

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

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

De fato, existe $\varepsilon > 0$ tal que $(a + th, b + k), (a, b + tk) \subset U$ para todo $t \in (-\varepsilon, 1 + \varepsilon)$. Como as derivadas parciais da função f existem em todos os pontos de U , as funções reais

$f_1(t) = f(a + th, b + k)$ e $f_2(t) = f(a, b + tk)$ são deriváveis em $(-\varepsilon, 1 + \varepsilon)$ e

$$\begin{aligned} \bullet \quad f'_1(t_0) &= \lim_{t \rightarrow 0} \frac{f(a + (t_0 + t)h, b + k) - f(a + t_0h, b + k)}{t} \\ &= \lim_{t \rightarrow 0} h \frac{f((a + t_0h, b + k) + ht(1, 0)) - f(a + t_0h, b + k)}{th} \\ &= h \frac{\partial f}{\partial x}(a + t_0h, b + k) \\ \bullet \quad f'_2(t_0) &= \lim_{t \rightarrow 0} \frac{f(a, b + (t_0 + t)k) - f(a, b + t_0k)}{t} \\ &= \lim_{t \rightarrow 0} k \frac{f((a, b + t_0k) + kt(0, 1)) - f(a, b + t_0k)}{tk} \\ &= k \frac{\partial f}{\partial y}(a, b + t_0k). \end{aligned}$$

Logo,

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

Como $\frac{|h|}{\sqrt{h^2 + b^2}} \leq 1$, $\frac{|k|}{\sqrt{h^2 + b^2}} \leq 1$, $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ são contínuas no ponto $c = (a, b)$, temos que

$$\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0, \text{ ou seja, } f \text{ é diferenciável no ponto } c = (a, b). \blacksquare$$

Observação 3.12. Na realidade, para que f seja diferenciável no ponto (a, b) é suficiente que $\frac{\partial f}{\partial x}$ exista numa vizinhança deste ponto, que nele seja contínua e que $\frac{\partial f}{\partial y}(a, b)$ exista.

De fato, escrevendo

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

existe, pelo Teorema do Valor Médio para funções reais de uma variável real, $\theta \in (0, 1)$ tal que

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

Logo $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0$, pois $\frac{h}{\|v\|}$ e $\frac{k}{\|v\|}$ são limitadas, $\frac{\partial f}{\partial x}$ é contínua no ponto (a, b) e

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

• Para funções de n variáveis, a diferenciabilidade de f num ponto é assegurada quando $n - 1$ das suas derivadas parciais existem numa vizinhança do ponto, são contínuas neste ponto e a derivada parcial restante apenas exista neste ponto.

Corolário 3.3. Toda função de classe C^1 é diferenciável.

Mas a recíproca não é verdadeira.

Exemplo 3.1. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ a função dada por $f(x) = x^2 \sin \frac{1}{x}$, $x \neq 0$ e $f(0) = 0$. Então

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}, \text{ para } x \neq 0, \quad \text{e} \quad f'(0) = \lim_{x \rightarrow 0} \frac{x^2 \sin \frac{1}{x}}{x} = 0.$$

Logo f é diferenciável em $\mathbb{R}$, mas f não é de classe C^1 , pois f' não é contínua em $x = 0$. $\square$

Exemplo 3.2. Um polinômio em duas variáveis é uma função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \sum a_{ij} x^i y^j.$$

Então f é contínuo em $\mathbb{R}^2$ e possui derivadas parciais

$$\frac{\partial f}{\partial x} = \sum i a_{ij} x^{i-1} y^j \quad \text{e} \quad \frac{\partial f}{\partial y} = \sum j a_{ij} x^i y^{j-1}.$$

Como $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ são polinômios e, portanto, funções contínuas, temos que f é de classe C^1 .

Assim, todo polinômio é de classe C^1 .

Como $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ são polinômios, $\frac{\partial f}{\partial x} \in C^1$ e $\frac{\partial f}{\partial y} \in C^1$. Logo $f \in C^2$.

Podemos provar, usando o argumento acima, que se todo polinômio é de classe C^k , então todo polinômio é de classe C^{k+1} . Assim, concluímos, por indução, que todo polinômio é de classe C^∞ .

Do mesmo modo, podemos mostrar que todo polinômio $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de n variáveis

$$f(x) = \sum a_{i_1 i_2 \dots i_n} x_1^{i_1} \cdots x_n^{i_n},$$

é de classe C^∞ . $\square$

Observação 3.13. A soma $f + g$ e o produto fg de funções de classe C^k são funções de classe C^k .

Este resultado segue do fato análogo já provado para funções reais de uma variável real, ou pode ser provado por indução, primeiro para a soma e depois para o produto.

Corolário 3.4. Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$, tal que $f(U) \subset V$ e cada função coordenada $f_i : U \rightarrow \mathbb{R}$ é de classe C^k . Se $g : V \rightarrow \mathbb{R}$ é uma função de classe C^k , então a composta $g \circ f : U \rightarrow \mathbb{R}$ é de classe C^k .

Prova.

Para $k = 0$, o resultado é verdadeiro. Suponhamos, por indução, que o corolário vale para funções de classe C^{k-1} , $k \geq 1$, e que $g, f_i, i = 1, \dots, n$ são funções de classe C^k .

Então, pelo corolário 3.3, $g, f_i, i = 1, \dots, n$ são funções diferenciáveis e, pela Regra da Cadeia:

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

para todo $x \in U$ e todo $i = 1, \dots, m$, ou seja, vale a igualdade de funções:

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

Como $\frac{\partial g}{\partial y_j}$ e f são de classe C^{k-1} temos, pela hipótese de indução, que $\frac{\partial g}{\partial y_j} \circ f$ é de classe C^{k-1} para todo $j = 1, \dots, n$. Além disso, como $\frac{\partial f_j}{\partial x_i} \in C^{k-1}$, o produto $\left(\frac{\partial g}{\partial y_j} \circ f \right) \cdot \frac{\partial f_j}{\partial x_i}$ é de classe C^{k-1} , para todo $j = 1, \dots, n$, e portanto, a soma $\sum_{j=1}^n \left(\frac{\partial g}{\partial y_j} \circ f \right) \cdot \frac{\partial f_j}{\partial x_i}$ é de classe C^{k-1} .

Logo $\frac{\partial(g \circ f)}{\partial x_i} \in C^{k-1}$ para todo $i = 1, \dots, m$, ou seja, $g \circ f \in C^k$. ■

Observação 3.14. Seja $g : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^k , com $g(x) \neq 0$ para todo $x \in U$.

Então a função $\frac{1}{g}$ é de classe C^k , pois $\frac{1}{g} = \rho \circ g$, onde $\rho : \mathbb{R} - \{0\} \rightarrow \mathbb{R}$, dada por $\rho(x) = \frac{1}{x}$, é de classe C^∞ .

Exemplo 3.3. O produto interno $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x, y) = \sum_{i=1}^n x_i y_i$, é uma função de classe C^∞ , pois f é um polinômio de $2n$ variáveis (de grau 2).

Também, a função $g : \mathbb{R}^n \rightarrow \mathbb{R}$, $g(x) = \|x\|^2 = \sum_{i=1}^n x_i^2$, por ser um polinômio de n variáveis, é de classe C^∞ .

Então a norma $h : \mathbb{R}^n - \{0\} \rightarrow \mathbb{R}$, $h(x) = \|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ é de classe C^∞ , pois $h = \rho \circ g$, onde $\rho : (0, \infty) \rightarrow \mathbb{R}$ é a função C^∞ dada por $\rho(x) = \sqrt{x}$.

Na origem, a função norma h não possui derivadas parciais, pois:

$$\bullet \lim_{t \rightarrow 0^+} \frac{h(0 + te_i) - h(0)}{t} = \lim_{t \rightarrow 0^+} \frac{|t|}{t} = 1, \quad \text{e} \quad \bullet \lim_{t \rightarrow 0^-} \frac{h(0 + te_i) - h(0)}{t} = \lim_{t \rightarrow 0^-} \frac{|t|}{t} = -1.$$

• Pode ocorrer que normas $\| \cdot \|$ que não provém de um produto interno não sejam diferenciáveis em pontos $x \neq 0$.

Por exemplo, se $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ é a norma da soma $\varphi(x, y) = |x| + |y|$, então não existe $\frac{\partial \varphi}{\partial x}$ nos pontos $(0, y)$ e não existe $\frac{\partial \varphi}{\partial y}$ nos pontos $(x, 0)$.

De fato, $\lim_{t \rightarrow 0^\pm} \frac{\varphi(t, y) - \varphi(0, y)}{t} = \lim_{t \rightarrow 0^\pm} \frac{|t|}{t} = \pm 1$, e $\lim_{t \rightarrow 0^\pm} \frac{\varphi(x, t) - \varphi(x, 0)}{t} = \lim_{t \rightarrow 0^\pm} \frac{|t|}{t} = \pm 1$. □

4 A diferencial de uma função

Definição 4.1. Sejam $U \subset \mathbb{R}^n$ um aberto e $f : U \rightarrow \mathbb{R}$ uma função diferenciável no ponto a .

A *diferencial de f no ponto a* é o funcional linear $df(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ dado por

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

onde $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

Então $\left(\frac{\partial f}{\partial x_1}(a) \quad \dots \quad \frac{\partial f}{\partial x_n}(a) \right)$ é a matriz $1 \times n$ do funcional linear $df(a)$ em relação à base canônica $\{e_1, \dots, e_n\}$ de $\mathbb{R}^n$.

Quando f é diferenciável em todo ponto de U , podemos definir a aplicação

$$df : U \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}) = (\mathbb{R}^n)^*$$

que associa a cada $x \in U$ o funcional $df(x)$, cuja matriz é $\left(\frac{\partial f}{\partial x_1}(x) \quad \dots \quad \frac{\partial f}{\partial x_n}(x) \right)$.

Identificando o funcional $df(x)$ com sua matriz, temos que: *df é uma aplicação contínua* $\iff$ *cada uma de suas funções coordenadas $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ é contínua* $\iff$ *f é C^1 .*

Exemplo 4.1. Todo funcional linear $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável e $d\varphi(x) = \varphi$, ou seja, $d\varphi(x)v = \varphi(v)$ para quaisquer $x, v \in \mathbb{R}^n$.

De fato, como $\varphi(x) = a_1x_1 + \dots + a_nx_n$, temos $\frac{\partial \varphi}{\partial x_i}(x) = a_i$ para todo $x \in \mathbb{R}^n$ e todo $i = 1, \dots, n$. Logo,

$$d\varphi(x)v = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(x) \alpha_i = \sum_{i=1}^n a_i \alpha_i = \varphi(v). \quad \square$$

Notação.

Seja $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $\pi_i(x) = x_i$, a projeção sobre a i -ésima coordenada, $i = 1, \dots, n$. Então $\{\pi_1, \dots, \pi_n\}$ é a base de $(\mathbb{R}^n)^*$ dual da base canônica.

Fazendo $\pi_i = x_i$, temos, pelo exemplo acima, que

$$dx_i(a)v = d\pi_i(a)v = \pi_i(v) = \alpha_i,$$

para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

Logo, podemos escrever:

$$df(a)v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i(a)(v), \quad \text{ou seja,} \quad df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i,$$

se f é diferenciável em todo ponto $a \in U$.

Com a identificação feita acima, temos que $\{dx_1, \dots, dx_n\}$ é a base de $(\mathbb{R}^n)^*$ dual da base canônica.

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

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

Teorema 4.1. *Sejam $f, g : U \rightarrow \mathbb{R}$ funções diferenciáveis no ponto $a \in U$. Então:*

- (a) $f + g : U \rightarrow \mathbb{R}$ é diferenciável no ponto a e $d(f + g)(a) = df(a) + dg(a)$.
- (b) $f \cdot g : U \rightarrow \mathbb{R}$ é diferenciável no ponto a e $d(f \cdot g)(a) = f(a) dg(a) + g(a) df(a)$.
- (c) Se $g(a) \neq 0$, $\frac{f}{g}$ é diferenciável no ponto a e $d\left(\frac{f}{g}\right)(a) = \frac{g(a) df(a) - f(a) dg(a)}{g(a)^2}$.

Prova.

Como as funções $s, m : \mathbb{R}^2 \rightarrow \mathbb{R}$, $q : \mathbb{R} \times (\mathbb{R} - \{0\}) \rightarrow \mathbb{R}$ dadas por $s(x, y) = x + y$, $m(x, y) = xy$ e $q(x, y) = \frac{x}{y}$ são diferenciáveis, por serem de classe C^∞ , e a função $F : U \rightarrow \mathbb{R}^2$, $F(x) = (f(x), g(x))$, tem coordenadas diferenciáveis no ponto a , temos, pela Regra da Cadeia, que as funções $s \circ F = f + g$, $m \circ F = f \cdot g$ e $q \circ F = \frac{f}{g}$ são diferenciáveis no ponto a e, além disso:

$$\begin{aligned} \frac{\partial(f+g)}{\partial x_i}(a) &= \frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a) \\ \frac{\partial(f \cdot g)}{\partial x_i}(a) &= g(a) \frac{\partial f}{\partial x_i}(a) + f(a) \frac{\partial g}{\partial x_i}(a) \\ \frac{\partial(f/g)}{\partial x_i}(a) &= \frac{g(a) \frac{\partial f}{\partial x_i}(a) - f(a) \frac{\partial g}{\partial x_i}(a)}{g(a)^2}. \end{aligned}$$

Assim,

- $d(f + g)(a) = \sum_{i=1}^n \frac{\partial(f+g)}{\partial x_i}(a) dx_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i + \sum_{i=1}^n \frac{\partial g}{\partial x_i}(a) dx_i = df(a) + dg(a)$;
- $d(f \cdot g)(a) = \sum_{i=1}^n \frac{\partial(f \cdot g)}{\partial x_i}(a) dx_i = g(a) \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i + f(a) \sum_{i=1}^n \frac{\partial g}{\partial x_i}(a) dx_i = g(a) df(a) + f(a) dg(a)$;
- $d(f/g)(a) = \sum_{i=1}^n \frac{\partial(f/g)}{\partial x_i}(a) dx_i = \frac{g(a) \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i - f(a) \sum_{i=1}^n \frac{\partial g}{\partial x_i}(a) dx_i}{g(a)^2} = \frac{g(a) df(a) - f(a) dg(a)}{g(a)^2}$.



Teorema 4.2. (do Valor Médio)

Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável em todos os pontos do segmento aberto $(a, a + v)$ e contínua no segmento fechado $[a, a + v] \subset U$. Então existe $\theta \in (0, 1)$ tal que

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

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

Corolário 4.1. Seja $U \subset \mathbb{R}^n$ aberto conexo. Se $f : U \rightarrow \mathbb{R}$ é diferenciável e $df(x) = 0$ (isto é, $\frac{\partial f}{\partial x_i}(x) = 0$, $i = 1, \dots, n$) para todo $x \in U$, então f é constante.

Corolário 4.2. Sejam $U \subset \mathbb{R}^n$ aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função diferenciável. Se $\|df(x)\| \leq M$ para todo $x \in U$, então

$$|f(x) - f(y)| \leq M \|x - y\|,$$

para quaisquer $x, y \in U$.

Prova.

Neste corolário, estamos assumindo que

$$\|df(x)\| = \sup \{ |df(x)v| \mid v \in \mathbb{R}^n, \|v\| = 1 \} = \sup \left\{ \left| \frac{\partial f}{\partial v}(x) \right| \mid v \in \mathbb{R}^n, \|v\| = 1 \right\}.$$

Logo, se $x, y \in U$, o segmento fechado $[x, x + (y - x)] \subset U$, uma vez que U é convexo.

Assim, pelo Teorema do Valor Médio, existe $\theta \in (0, 1)$ tal que

$$f(y) - f(x) = df(x + \theta(y - x))(y - x),$$

e, portanto,

$$|f(y) - f(x)| = |df(x + \theta(y - x))(y - x)| \leq M \|y - x\|. \blacksquare$$

Observação 4.1. Se tomarmos em $\mathbb{R}^n$ a norma euclidiana, ou a norma da soma, ou a norma do máximo, então $\|df(x)\|$ assume, respectivamente, os valores:

$$\sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) \right)^2}, \quad \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x) \right|, \quad \text{ou} \quad \max_{1 \leq i \leq n} \left\{ \left| \frac{\partial f}{\partial x_i}(x) \right| \right\}.$$

De fato, se $\| \cdot \|$ é a norma euclidiana, por exemplo, temos que:

$$|df(x)v| = \left| \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \alpha_i \right| \leq \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) \right)^2} \sqrt{\sum_{i=1}^n \alpha_i^2} \leq \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) \right)^2},$$

para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ com $\|v\| = \sqrt{\sum_{i=1}^n \alpha_i^2} = 1$.

Logo $\|df(x)\| \leq \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2}$. Por outro lado, se $\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2 \neq 0$, podemos tomar o vetor

$$v = \frac{\left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right)}{\sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2}}.$$

Então, como $\|v\| = 1$, temos que:

$$\|df(x)\| \geq |df(x)v| = \frac{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2}{\sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2}} = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2},$$

ou seja, $\|df(x)\| \geq \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2}$. Assim, $\|df(x)\| = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x)\right)^2}$.

Observação 4.2. Se V não é convexo, uma função $g : V \rightarrow \mathbb{R}$ pode ser diferenciável, com diferencial dg limitada em V , sem ser Lipschitziana.

Por exemplo, sejam $U = \mathbb{R}^2 - X$, onde $X = \{(x, 0) \mid x \geq 0\}$, e $V = \{(x, y) \in U \mid \sqrt{x^2 + y^2} < 2\}$.

Seja $g = f|_V$, onde $f : U \rightarrow \mathbb{R}$ é a função definida por $f(x, y) = x^2$ se $x > 0$ e $y > 0$ e $f(x, y) = 0$ se $x \leq 0$ ou $y \leq 0$.

Então $\frac{\partial f}{\partial y}(x, y) = 0$ para todo $(x, y) \in U$; $\frac{\partial f}{\partial x}(x, y) = 2x$ se $x > 0$, $y > 0$; $\frac{\partial f}{\partial x}(x, y) = 0$ se $(x, y) \in U - \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y > 0\}$, pois $f \equiv 0$ neste aberto e, também, $\frac{\partial f}{\partial x}(0, y) = 0$ para $y > 0$, uma vez que

$$\bullet \lim_{t \rightarrow 0^+} \frac{f(t, y) - f(0, y)}{t} = \lim_{t \rightarrow 0^+} \frac{t^2}{t} = 0, \quad \text{e} \quad \bullet \lim_{t \rightarrow 0^-} \frac{f(t, y) - f(0, y)}{t} = \lim_{t \rightarrow 0^-} \frac{0}{t} = 0.$$

Logo f é diferenciável, pois $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ são contínuas em U , ou seja, f é de classe C^1 em U . Além disso, como $|x| < 2$ para todo $(x, y) \in V$,

$$\|df(x, y)\| = \sqrt{\left(\frac{\partial f}{\partial x}(x, y)\right)^2 + \left(\frac{\partial f}{\partial y}(x, y)\right)^2} \leq 4,$$

para todo $(x, y) \in V$.

Mas, f não é uniformemente contínua em V , pois, para as sequências $z_n = \left(1, \frac{1}{n}\right)$ e $w_n = \left(1, -\frac{1}{n}\right)$ de pontos de V , temos que:

$$z_n - w_n = \left(0, \frac{2}{n}\right) \rightarrow (0, 0) \quad \text{e} \quad f(z_n) - f(w_n) = 1 \rightarrow 1.$$

Em particular, f não é Lipschitziana em V .

Observação 4.3. Como consequência do corolário 4.2, temos que se $U \subset \mathbb{R}^n$ é aberto e convexo e $f : U \rightarrow \mathbb{R}$ é uma função diferenciável com derivadas parciais limitadas em U , então f é uniformemente contínua em U . Em particular, f é a restrição de uma função uniformemente contínua $g : \bar{U} \rightarrow \mathbb{R}$.

5 O gradiente de uma função diferenciável

O produto interno canônico induz um isomorfismo entre $\mathbb{R}^n$ e seu dual $(\mathbb{R}^n)^*$ dado por:

$$\begin{aligned} \mathbb{R}^n &\longrightarrow (\mathbb{R}^n)^* \\ v &\longmapsto v^* : \mathbb{R}^n \longrightarrow \mathbb{R} \\ x &\longmapsto \langle v, x \rangle, \end{aligned}$$

pois dado $\varphi \in (\mathbb{R}^n)^*$, $\varphi = v^*$, onde $v = (\varphi(e_1), \dots, \varphi(e_n))$, uma vez que

$$\varphi(x_1, \dots, x_n) = \varphi(e_1)x_1 + \dots + \varphi(e_n)x_n.$$

Além disso, como $v^*(e_i) = \alpha_i$, $i = 1, \dots, n$, $\begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix}$ é a matriz $1 \times n$ do funcional v^* em relação à base canônica.

Definição 5.1. Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável no aberto $U \subset \mathbb{R}^n$. O *gradiente de f no ponto $a \in U$* é o vetor $\text{grad } f(a)$ que corresponde ao funcional $df(a)$ segundo o isomorfismo acima, ou seja,

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

para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

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

Observação 5.1. As coordenadas de $\text{grad } f(a)$ em relação à base canônica são iguais às coordenadas de $df(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i$ em relação à base $\{dx_1, \dots, dx_n\}$ de $(\mathbb{R}^n)^*$, dual da base canônica.

• Veremos agora as *três propriedades* mais importantes do gradiente de uma função diferenciável $f : U \rightarrow \mathbb{R}$. Para isso, seja $a \in U$ tal que $\text{grad } f(a) \neq 0$.

Primeira propriedade. O gradiente aponta para uma direção segundo a qual a função f é crescente.

De fato, se $w = \text{grad } f(a)$, então

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

Assim, se $\lambda : (-\varepsilon, \varepsilon) \rightarrow U$ é um caminho diferenciável tal que $\lambda(0) = a$ e $\lambda'(0) = \text{grad } f(a)$, então

$$(f \circ \lambda)'(0) = df(\lambda(0))\lambda'(0) > 0.$$

Então, se f e λ são de classe C^1 , existe $\varepsilon > 0$ tal que $(f \circ \lambda)'(t) > 0$ para todo $t \in (-\varepsilon, \varepsilon)$, e, portanto, $f \circ \lambda$ é crescente. Isto é, ***f cresce na direção do gradiente***.

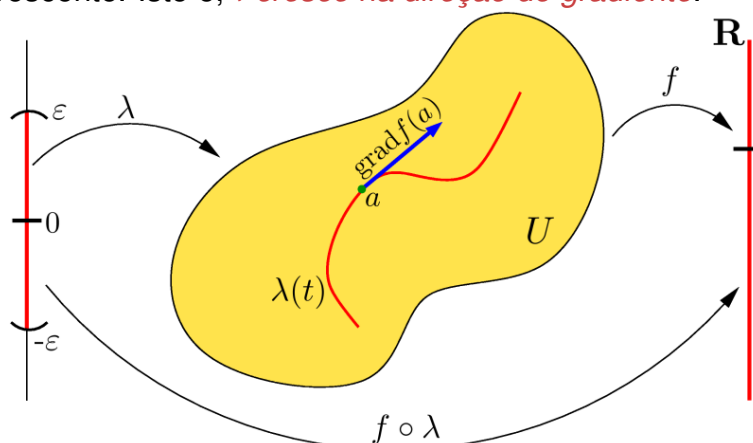


Fig. 4: Gradiente de f no ponto a

Segunda propriedade. *Dentre todas as direções ao longo das quais a função f cresce, a direção do gradiente é a de crescimento mais rápido.*

De fato, não se tem $df(a)v = \langle \text{grad } f(a), v \rangle > 0$ apenas quando $v = \text{grad } f(a)$, pois $\langle \text{grad } f(a), v \rangle > 0$ para todo v que faz um ângulo agudo com $\text{grad } f(a)$. Então f cresce ao longo destas direções, mas $\text{grad } f(a)$ é a direção segundo a qual o crescimento de f é o mais rápido.

Ou seja, se v é um vetor tal que $\|v\| = \|\text{grad } f(a)\|$, então

$$\frac{\partial f}{\partial v}(a) \leq \frac{\partial f}{\partial (\text{grad } f(a))}(a),$$

pois, pela desigualdade de Cauchy-Schwarz,

$$\frac{\partial f}{\partial v}(a) = \langle \text{grad } f(a), v \rangle \leq \|\text{grad } f(a)\| \|v\| = \|\text{grad } f(a)\|^2 = \frac{\partial f}{\partial (\text{grad } f(a))}(a).$$

Observe, ainda, que a igualdade ocorre se, e só se, $v = \text{grad } f(a)$.

Terceira propriedade. *O gradiente de f no ponto a é perpendicular à "superfície" de nível de f que passa por esse ponto.*

Dado $c \in \mathbb{R}$, chamamos $f^{-1}(c) = \{x \in U \mid f(x) = c\}$ ***conjunto de nível de f*** e se $f(x) = c$, isto é, $x \in f^{-1}(c)$, dizemos que ***x está no nível c*** ou que ***x tem nível c*** .

O ***Teorema da Função Implícita***, que provaremos depois, garante que $f^{-1}(c)$ é uma su-

perfície (se $n \geq 3$), ou uma curva (se $n = 2$), quando $\text{grad } f(x) \neq 0$ para todo $x \in f^{-1}(c)$.

Dizer que $w = \text{grad } f(a)$ é perpendicular ao conjunto de nível $f^{-1}(c)$, onde $f(a) = c$, significa que w é perpendicular ao vetor velocidade $\lambda'(0)$ de qualquer caminho diferenciável em $t = 0$, com $\lambda(0) = a$ e $\lambda(t) \in f^{-1}(c)$ para todo $t \in (-\varepsilon, \varepsilon)$. De fato, como $f(\lambda(t)) = c$ para todo $t \in (-\varepsilon, \varepsilon)$,

$$0 = (f \circ \lambda)'(0) = df(\lambda(0)) \lambda'(0) = \langle \text{grad } f(a), \lambda'(0) \rangle.$$

Exemplo 5.1. Sejam $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por: $f(x, y) = ax + by$, $a^2 + b^2 \neq 0$; $g(x, y) = x^2 + y^2$ e $h(x, y) = x^2 - y^2$.

- As curvas de nível de f são as retas $ax + by = c$ para qualquer $c \in \mathbb{R}$ e $\text{grad } f(x, y) = (a, b)$ para todo $(x, y) \in \mathbb{R}^2$. Assim, (a, b) é o vetor normal às retas $ax + by = c$, e $\{(x, y) \in \mathbb{R}^2 \mid ax + by > c\}$ é o semi-plano para o qual o vetor (a, b) aponta.

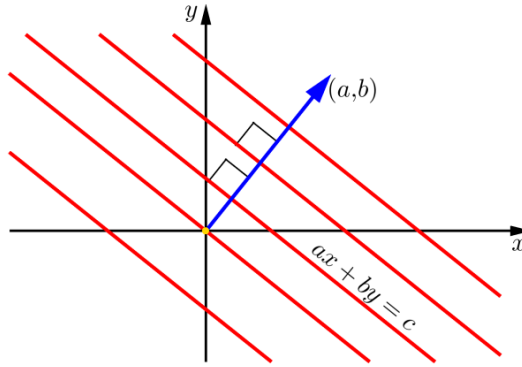


Fig. 5: Gradiante de f

- Seja $c \in \mathbb{R}$ e seja $g^{-1}(c) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = c\}$ a curva de nível c da função g . Então: $g^{-1}(c) = \emptyset$ se $c < 0$, $g^{-1}(0) = \{(0, 0)\}$, $g^{-1}(c)$ é o círculo de centro na origem e raio $\sqrt{c}$, e $\text{grad } f(x, y) = (2x, 2y)$ é um vetor paralelo ao raio e, portanto, perpendicular ao vetor tangente ao círculo naquele ponto.

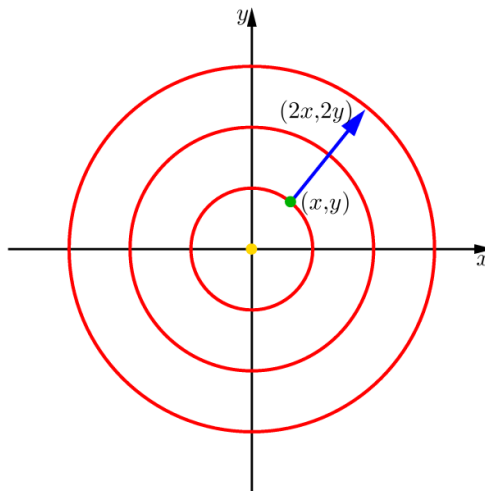


Fig. 6: Gradiante de g

- As curvas de nível c da função h são:

$$h^{-1}(0) = \{(x, y) \in \mathbb{R}^2 \mid x^2 = y^2\} = \{(x, y) \in \mathbb{R}^2 \mid x = \pm y\}$$

que consiste de duas retas, $x = y$ e $x = -y$, perpendiculares que se cortam na origem; ou

$$h^{-1}(c) = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = c\}$$

que é uma hipérbole cuja reta focal é o eixo x , se $c > 0$, e uma hipérbole cuja reta focal é o eixo y , se $c < 0$.

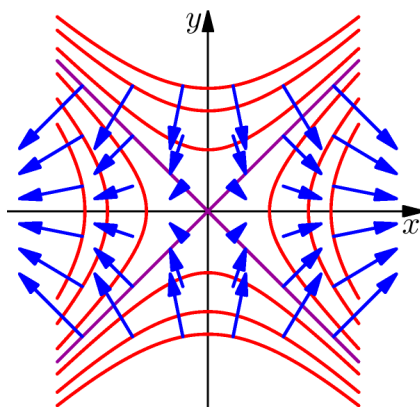


Fig. 7: Gradiente de h .

O gradiente de h , $\text{grad } h(x, y) = (2x, -2y)$, é perpendicular às curvas de nível e indica a direção de crescimento de h . $\square$

- Nos pontos onde o gradiente se anula ocorre uma quebra de regularidade na disposição das curvas de nível. Um ponto onde o gradiente de uma função é o vetor nulo é chamado *singular* ou *crítico*.

Exemplo 5.2. Considere, agora, as funções definidas no espaço $\mathbb{R}^3$ tridimensional:

$$f(x, y, z) = ax + by + cz; \quad g(x, y, z) = x^2 + y^2 + z^2 \quad \text{e} \quad h(x, y, z) = x^2 + y^2 - z^2.$$

As superfícies de nível de f são planos de equação $ax + by + cz = d$, $d \in \mathbb{R}$, todos perpendiculares ao vetor (a, b, c) , que é o gradiente de f em qualquer ponto.

A superfície de nível c da função g é o conjunto vazio, se $c < 0$; consiste apenas da origem, se $c = 0$ e é a esfera de centro na origem e raio $\sqrt{c}$, se $c > 0$, sendo $\text{grad } g(x, y, z) = 2(x, y, z)$ perpendicular à superfície de nível c que passa pelo ponto $(x, y, z) \neq (0, 0, 0)$.

A superfície de nível c da função h é o cone de revolução $z^2 = x^2 + y^2$ de vértice na origem e eixo z , o hiperbolóide de revolução de uma folha $x^2 + y^2 - z^2 = c$ de eixo z , se $c > 0$, e o hiperbolóide de revolução de duas folhas $x^2 + y^2 - z^2 = c$ de eixo z , se $c < 0$, sendo $\text{grad } h(x, y, z) = 2(x, y, -z)$ perpendicular à superfície de nível que passa por (x, y, z) . $\square$

6 A regra de Leibniz

Teorema 6.1. (Regra de Leibniz – derivação sob o sinal de integral)

Sejam $U \subset \mathbb{R}^n$ aberto e $f : U \times [a, b] \rightarrow \mathbb{R}$ uma função com as seguintes propriedades:

(1) Para todo $x \in U$, a função $t \mapsto f(x, t)$ é integrável em $[a, b]$.

(2) A i -ésima derivada parcial $\frac{\partial f}{\partial x_i}(x, t)$ existe para todo $(x, t) \in U \times [a, b]$ e a função $\frac{\partial f}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ é contínua.

Então a função $\varphi : U \rightarrow \mathbb{R}$, dada por $\varphi(x) = \int_a^b f(x, t) dt$, possui i -ésima derivada parcial em todo ponto $x \in U$, sendo

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

Ou seja, pode-se derivar sob o sinal de integral, desde que o integrando resultante seja uma função contínua.

Prova.

Dado $x_0 \in U$, existe $\delta_0 > 0$ tal que $[x_0, x_0 + se_i] \subset U$, para todo $s \in \mathbb{R}$ com $|s| < \delta_0$. Então, pelo Teorema do Valor Médio, existe $\theta \in (0, 1)$ tal que:

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

Como $\frac{\partial f}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ é contínua, temos, pelo teorema 11.4 do capítulo 1, que dado $\varepsilon > 0$, existe $0 < \delta < \delta_0$ tal que:

$$|s| < \delta \implies \left| \frac{\partial f}{\partial x_i}(x_0 + s\theta e_i, t) - \frac{\partial f}{\partial x_i}(x_0, t) \right| < \frac{\varepsilon}{2(b-a)},$$

para todo $t \in [a, b]$. Então, se $0 < |s| < \delta$,

$$\left| \frac{\varphi(x_0 + se_i) - \varphi(x_0)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t) dt \right| < \varepsilon.$$

Provamos, então, que φ possui i -ésima derivada parcial no ponto x_0 e

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

Corolário 6.1. Se $f : U \times [a, b] \rightarrow \mathbb{R}$ é contínua e possui as n derivadas parciais $\frac{\partial f}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ contínuas, então $\varphi : U \rightarrow \mathbb{R}$, dada por $\varphi(x) = \int_a^b f(x, t) dt$, é de classe C^1 .

Prova.

Pelo teorema anterior, φ possui as n derivadas parciais e $\frac{\partial \varphi}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x, t) dt$ para todo $x \in U$, $i = 1, \dots, n$. Além disso, como $\frac{\partial f}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ é contínua, para todo $i = 1, \dots, n$, temos, pela aplicação do teorema 11.4 do capítulo 1, que $\frac{\partial \varphi}{\partial x_i} : U \rightarrow \mathbb{R}$ é contínua para todo $i = 1, \dots, n$. ■

Observação 6.1. Se $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ é uma função contínua, temos, pela aplicação do teorema 11.4 do capítulo 1, que a função $\xi : [a, b] \rightarrow \mathbb{R}$, $\xi(s) = \int_c^d f(s, t) dt$, é contínua e, portanto, integrável.

A integral $\int_a^b \xi(s) ds$ se escreve como:

$$\int_a^b \left[\int_c^d f(s, t) dt \right] ds \quad \text{ou} \quad \int_a^b ds \int_c^d f(s, t) dt.$$

Teorema 6.2. (da Inversão da Ordem nas Integrais Repetidas)

Se $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ é uma função contínua, então

$$\int_a^b ds \int_c^d f(s, t) dt = \int_c^d dt \int_a^b f(s, t) ds.$$

Prova.

Seja $g : [a, b] \times [c, d] \rightarrow \mathbb{R}$ definida por $g(x, t) = \int_a^x f(s, t) ds$.

Para cada $x \in [a, b]$ fixo, a função $t \mapsto \int_a^x f(s, t) ds$ é contínua e, portanto, integrável. Além disso, $\frac{\partial g}{\partial x}(x, t) = f(x, t)$ para todo $(x, t) \in [a, b] \times [c, d]$, pois o integrando $s \mapsto f(s, t)$ é contínuo para todo $t \in [c, d]$.

Como $\frac{\partial g}{\partial x} = f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ é contínua, temos, pela Regra de Leibniz, que a função $\varphi : [a, b] \rightarrow \mathbb{R}$, dada por

$$\varphi(x) = \int_c^d g(x, t) dt = \int_c^d \left(\int_a^x f(s, t) ds \right) dt,$$

é derivável e $\varphi'(x) = \int_c^d \frac{\partial g}{\partial x}(x, t) dt = \int_c^d f(x, t) dt$.

Como $\varphi' : [a, b] \rightarrow \mathbb{R}$ é integrável (por ser contínua), temos, pelo Teorema Fundamental do Cálculo, que

$$\varphi(b) - \varphi(a) = \int_a^b \varphi'(s) ds = \int_a^b \left(\int_c^d f(s, t) dt \right) ds.$$

Sendo $\varphi(a) = 0$ e $\varphi(b) = \int_c^d \left(\int_a^b f(s, t) ds \right) dt$, obtemos

$$\int_c^d \left(\int_a^b f(s, t) ds \right) dt = \int_a^b \left(\int_c^d f(s, t) dt \right) ds. \blacksquare$$

Corolário 6.2. *Seja $f : U \times [a, b] \rightarrow \mathbb{R}$ uma função contínua, com derivadas parciais contínuas $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \times [a, b] \rightarrow \mathbb{R}$, e seja $g : U \rightarrow [a, b]$ uma função de classe C^1 , onde $U \subset \mathbb{R}^n$ é aberto. Então a função $\varphi : U \rightarrow \mathbb{R}$, definida por $\varphi(x) = \int_a^{g(x)} f(x, t) dt$, é de classe C^1 e suas derivadas parciais são:*

$$\frac{\partial \varphi}{\partial x_i}(x) = \int_a^{g(x)} \frac{\partial f}{\partial x_i}(x, t) dt + \frac{\partial g}{\partial x_i}(x) f(x, g(x)),$$

para todo $x \in U$.

Prova.

Seja $\xi : U \times [a, b] \rightarrow \mathbb{R}$ a função dada por $\xi(x, u) = \int_a^u f(x, t) dt$. Então, como a função $t \mapsto f(x, t)$ é contínua, $\frac{\partial \xi}{\partial u}(x, u) = f(x, u)$ para todo $(x, u) \in U \times [a, b]$.

Além disso, pela Regra de Leibniz, $\frac{\partial \xi}{\partial x_i}(x, u) = \int_a^u \frac{\partial f}{\partial x_i}(x, t) dt$.

Afirmção: $\frac{\partial \xi}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ é contínua, para $i = 1, \dots, n$.

De fato, como $\frac{\partial f}{\partial x_i} : U \times [a, b] \rightarrow \mathbb{R}$ é contínua, temos, pelo teorema 11.4 do capítulo 1, que dados $x_0 \in U$, $u_0 \in [a, b]$ e $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\|x - x_0\| < \delta \implies \left| \frac{\partial f}{\partial x_i}(x, t) - \frac{\partial f}{\partial x_i}(x_0, t) \right| < \varepsilon',$$

para todo $t \in [a, b]$, onde $\varepsilon' = \frac{\varepsilon}{2}$ se $u_0 = a$ e $\varepsilon' = \frac{\varepsilon}{2(u_0 - a)}$ se $u_0 \neq a$.

Sendo $t \mapsto \frac{\partial f}{\partial x_i}(x_0, t)$ contínua no compacto $[a, b]$, existe $M > 0$ tal que $\left| \frac{\partial f}{\partial x_i}(x_0, t) \right| \leq M$ para todo $t \in [a, b]$. Assim, $\left| \frac{\partial f}{\partial x_i}(x, t) \right| \leq N = \varepsilon' + M$, para todo $t \in [a, b]$ e $x \in B(x_0, \delta)$.

Então, se $|u - u_0| < \frac{\varepsilon}{2N}$ e $\|x - x_0\| < \delta$,

$$\begin{aligned} \left| \frac{\partial \xi}{\partial x_i}(x, u) - \frac{\partial \xi}{\partial x_i}(x_0, u_0) \right| &= \left| \int_a^u \frac{\partial f}{\partial x_i}(x, t) dt - \int_a^{u_0} \frac{\partial f}{\partial x_i}(x_0, t) dt \right| \\ &\leq \left| \int_a^{u_0} \frac{\partial f}{\partial x_i}(x, t) dt - \int_a^{u_0} \frac{\partial f}{\partial x_i}(x_0, t) dt \right| + \left| \int_{u_0}^u \frac{\partial f}{\partial x_i}(x, t) dt \right| \\ &\leq \varepsilon' |u_0 - a| + N |u_0 - u| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Logo ξ é de classe C^1 , pois $\frac{\partial \xi}{\partial u} = f$ e $\frac{\partial \xi}{\partial x_i}$, $i = 1, \dots, n$ são contínuas.

Sendo g e ξ são de classe C^1 e, portanto, diferenciáveis, temos, pela Regra da Cadeia, que a função composta $\varphi(x) = \xi(x, g(x))$ é diferenciável e, para todo $i = 1, \dots, n$,

$$\frac{\partial \varphi}{\partial x_i}(x) = \frac{\partial \xi}{\partial x_i}(x, g(x)) + \frac{\partial \xi}{\partial u}(x, g(x)) \frac{\partial g}{\partial x_i}(x) = \int_a^{g(x)} \frac{\partial f}{\partial x_i}(x, t) dt + \frac{\partial g}{\partial x_i}(x) f(x, g(x)).$$

Logo $\frac{\partial \varphi}{\partial x_i}$ é contínua para todo $i = 1, \dots, n$, ou seja, φ é de classe C^1 . ■

Observação 6.2. De modo análogo, podemos provar que se $f : U \times [a, b] \rightarrow \mathbb{R}$ satisfaz as hipóteses do corolário acima e $g, h : U \rightarrow [a, b]$ são de classe C^1 , então as funções

$$\psi(x) = \int_{g(x)}^b f(x, t) dt, \quad \text{e} \quad \lambda(x) = \int_{g(x)}^{h(x)} f(x, t) dt,$$

são de classe C^1 e

- $\frac{\partial \psi}{\partial x_i}(x) = \int_{g(x)}^b \frac{\partial f}{\partial x_i}(x, t) dt - \frac{\partial g}{\partial x_i}(x) f(x, g(x));$
- $\frac{\partial \lambda}{\partial x_i}(x) = \int_{g(x)}^{h(x)} \frac{\partial f}{\partial x_i}(x, t) dt + \frac{\partial h}{\partial x_i}(x) f(x, h(x)) - \frac{\partial g}{\partial x_i}(x) f(x, g(x)),$

uma vez que, $\int_a^b - \int_a^{g(x)} = \int_{g(x)}^b$, e $\int_a^b - \int_a^{g(x)} - \int_{h(x)}^b = \int_{g(x)}^{h(x)}$.

Observação 6.3. Seja $f : I \rightarrow \mathbb{R}$ uma função contínua definida no intervalo I , com $0 \in I$.

Seja $F_0 = f$ e $F_n : I \rightarrow \mathbb{R}$, $n \geq 1$, definida por

$$F_n(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t) dt.$$

Então F_n é de classe C^n , $F_n(0) = F'_n(0) = \dots = F_n^{(n-1)}(0) = 0$ e $F_n^{(n)}(x) = f(x)$ para todo $x \in I$.

De fato, para $n = 1$, F_1 é de classe C^1 , $F_1(0) = 0$ e $F'_1(x) = f(x)$ para todo $x \in I$.

Suponhamos o resultado válido para $n-1$, $n-1 \geq 1$. Sejam as funções $G : I \times I \rightarrow \mathbb{R}$ e $g : I \rightarrow I$ dadas por

$$G(x, t) = \frac{(x-t)^{n-1}}{(n-1)!} f(t), \quad \text{e} \quad g(x) = x.$$

Então $F_n(0) = 0$ e, pelo corolário acima, F_n é de classe C^1 e

$$F'_n(x) = \int_0^x \frac{(x-t)^{n-2}}{(n-2)!} f(t) dt + G(x, x) g'(x) = \int_0^x \frac{(x-t)^{n-2}}{(n-2)!} f(t) dt = F_{n-1}(x),$$

pois $G(x, x) = 0$.

Como, por indução, F_{n-1} é de classe C^{n-1} e $F_{n-1}(0) = \dots = F_{n-1}^{(n-2)}(0) = 0$ e $F_{n-1}^{(n-1)}(x) = f(x)$, temos que F_n é de classe C^n , $F_n(0) = F'_n(0) = \dots = F_n^{(n-1)}(0) = 0$ e $F_n^{(n)}(x) = f(x)$ para todo $x \in I$.

7 O Teorema de Schwarz

Definição 7.1. Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável no aberto $U \subset \mathbb{R}^n$. Se as derivadas parciais $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$, $i = 1, \dots, n$, são diferenciáveis num ponto $a \in U$, dizemos que f é *duas vezes diferenciável no ponto a* . Neste caso, existem as *derivadas parciais de segunda ordem*

$$\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) = \frac{\partial^2 f}{\partial x_j \partial x_i} (a),$$

para todo $i, j = 1, \dots, n$.

Se $f : U \rightarrow \mathbb{R}$ é duas vezes diferenciável em U , ficam definidas n^2 funções

$$\frac{\partial^2 f}{\partial x_j \partial x_i} : U \rightarrow \mathbb{R}, \quad 1 \leq i, j \leq n.$$

Se todas estas funções são diferenciáveis num ponto $a \in U$, dizemos que f é *três vezes diferenciável nesse ponto*. E assim por diante.

Observação 7.1. Já sabemos que se $f \in C^1$, então f é diferenciável.

Suponhamos, por indução, que se uma função é de classe C^k , então ela é k -vezes diferenciável.

Seja $f \in C^{k+1}$. Então suas derivadas parciais $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, são de classe C^k .

Logo, por indução, $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, são k -vezes diferenciáveis, e, portanto, f é $(k+1)$ -vezes diferenciável.

Cabe, então, determinar sob quais hipóteses a ordem em que são tomadas as derivadas parciais repetidas não influi no resultado final.

Teorema 7.1. (de Schwarz)

Se $f : U \rightarrow \mathbb{R}$ é duas vezes diferenciável num ponto $c \in U \subset \mathbb{R}^n$, então

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

para quaisquer $1 \leq i, j \leq n$.

Prova.

Vamos supor, para simplificar a notação, que $U \subset \mathbb{R}^2$ e $c = (a, b)$. Devemos, então, provar que $\frac{\partial^2 f}{\partial x \partial y} (a, b) = \frac{\partial^2 f}{\partial y \partial x} (a, b)$.

Seja $\varepsilon > 0$ tal que $(a - \varepsilon, a + \varepsilon) \times (b - \varepsilon, b + \varepsilon) \subset U$. Para todo $t \in (-\varepsilon, \varepsilon)$ e $x \in (a - \varepsilon, a + \varepsilon)$, sejam:

- $\varphi(t) = f(a+t, b+t) - f(a+t, b) - f(a, b+t) + f(a, b)$.
- $\xi(x) = f(x, b+t) - f(x, b)$.

Então $\varphi(t) = \xi(a+t) - \xi(a)$. Pelo Teorema do Valor Médio para funções de uma variável real, existe $\theta \in (0, 1)$ tal que $\varphi(t) = \xi'(a+\theta t)t$, ou seja,

$$\varphi(t) = \left(\frac{\partial f}{\partial x}(a+\theta t, b+t) - \frac{\partial f}{\partial x}(a+\theta t, b) \right) t.$$

Como a função $\frac{\partial f}{\partial x} : U \rightarrow \mathbb{R}$ é diferenciável no ponto $c = (a, b)$ temos que:

- $\frac{\partial f}{\partial x}(a+\theta t, b+t) = \frac{\partial f}{\partial x}(a, b) + \frac{\partial^2 f}{\partial x^2}(a, b)\theta t + \frac{\partial^2 f}{\partial y \partial x}(a, b)t + \rho_1 t$, com $\lim_{t \rightarrow 0} \rho_1 = 0$.

e

- $\frac{\partial f}{\partial x}(a+\theta t, b) = \frac{\partial f}{\partial x}(a, b) + \frac{\partial^2 f}{\partial x^2}(a, b)\theta t + \rho_2 t$, com $\lim_{t \rightarrow 0} \rho_2 = 0$.

Logo $\varphi(t) = \frac{\partial^2 f}{\partial y \partial x}(a, b)t^2 + (\rho_1 - \rho_2)t^2$, e, portanto,

$$\lim_{t \rightarrow 0} \frac{\varphi(t)}{t^2} = \frac{\partial^2 f}{\partial y \partial x}(a, b). \quad (I)$$

Seja, agora, $\eta(y) = f(a+t, y) - f(a, y)$. Então $\varphi(t) = \eta(b+t) - \eta(b)$. Pelo teorema do Valor Médio, existe $\theta \in (0, 1)$ tal que $\varphi(t) = \eta'(b+\theta t)t$, ou seja,

$$\varphi(t) = \left(\frac{\partial f}{\partial y}(a+t, b+\theta t) - \frac{\partial f}{\partial y}(a, b+\theta t) \right) t.$$

Como a função $\frac{\partial f}{\partial y} : U \rightarrow \mathbb{R}$ é diferenciável no ponto $c = (a, b)$, temos que:

- $\frac{\partial f}{\partial y}(a+t, b+\theta t) = \frac{\partial f}{\partial y}(a, b) + \frac{\partial^2 f}{\partial x \partial y}(a, b)t + \frac{\partial^2 f}{\partial y^2}(a, b)\theta t + \rho_3 t$, com $\lim_{t \rightarrow 0} \rho_3 = 0$,

e

- $\frac{\partial f}{\partial y}(a, b+\theta t) = \frac{\partial f}{\partial y}(a, b) + \frac{\partial^2 f}{\partial y^2}(a, b)\theta t + \rho_4 t$, com $\lim_{t \rightarrow 0} \rho_4 = 0$.

Logo $\varphi(t) = \left[\frac{\partial^2 f}{\partial x \partial y}(a, b) + (\rho_3 - \rho_4) \right] t^2$, e, portanto,

$$\lim_{t \rightarrow 0} \frac{\varphi(t)}{t^2} = \frac{\partial^2 f}{\partial x \partial y}(a, b). \quad (II)$$

Assim, por (I) e (II), $\frac{\partial^2 f}{\partial y \partial x}(a, b) = \frac{\partial^2 f}{\partial x \partial y}(a, b)$. ■

Corolário 7.1. Se $f : U \rightarrow \mathbb{R}$ é de classe C^2 no aberto $U \subset \mathbb{R}^n$, então $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ para todo $x \in U$ e para todo $1 \leq i, j \leq n$.

Exemplo 7.1. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função dada por $f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$, se $(x, y) \neq (0, 0)$, e $f(0, 0) = 0$.

A função f é de classe C^∞ em $\mathbb{R}^2 - \{(0, 0)\}$. Além disso, temos que:

- $\frac{\partial f}{\partial x}(x, 0) = \lim_{t \rightarrow 0} \frac{f(x+t, 0) - f(x, 0)}{t} = 0, \quad x \in \mathbb{R};$
- $\frac{\partial^2 f}{\partial x^2}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial x}(t, 0) - \frac{\partial f}{\partial x}(0, 0)}{t} = 0;$
- $\frac{\partial f}{\partial y}(0, y) = \lim_{t \rightarrow 0} \frac{f(0, y+t) - f(0, y)}{t} = 0, \quad y \in \mathbb{R};$
- $\frac{\partial^2 f}{\partial y^2}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial y}(0, t) - \frac{\partial f}{\partial y}(0, 0)}{t} = 0;$
- $\frac{\partial f}{\partial x}(0, y) = \lim_{t \rightarrow 0} \frac{f(t, y) - f(0, y)}{t} = \lim_{t \rightarrow 0} \frac{ty(t^2 - y^2)}{t(t^2 + y^2)} = -y, \quad y \in \mathbb{R};$
- $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, t) - \frac{\partial f}{\partial x}(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{-t}{t} = -1;$
- $\frac{\partial f}{\partial y}(x, 0) = \lim_{t \rightarrow 0} \frac{f(x, t) - f(x, 0)}{t} = \lim_{t \rightarrow 0} \frac{xt(x^2 - t^2)}{t(x^2 + t^2)} = x, \quad x \in \mathbb{R};$
- $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial y}(t, 0) - \frac{\partial f}{\partial y}(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{t}{t} = 1.$

Logo f possui derivadas parciais de segunda ordem em todos os pontos do plano, mas

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Pode-se verificar também que $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ são contínuas em $\mathbb{R}^2$, ou seja, f é de classe C^1 em $\mathbb{R}^2$, mas $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ não são diferenciáveis na origem. Logo f é diferenciável na origem, mas não é duas-vezes diferenciável na origem.

Além disso, apesar das derivadas de segunda ordem $\frac{\partial^2 f}{\partial x \partial y}$ e $\frac{\partial^2 f}{\partial y \partial x}$ existirem em todos os pontos do plano, elas não são contínuas na origem.

De fato, como para $(x, y) \neq (0, 0)$,

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{(y(x^2 - y^2) + xy \cdot 2x)(x^2 + y^2) - 2x \cdot xy(x^2 - y^2)}{(x^2 + y^2)^2} \\ &= \frac{((3x^2y - y^3))(x^2 + y^2) - 2x^2y(x^2 - y^2)}{(x^2 + y^2)^2} \\ &= \frac{x^4y - y^5 + 4x^2y^3}{(x^2 + y^2)^2}; \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{(x^4 - 5y^4 + 12x^2y^2)(x^2 + y^2)^2 - (x^4y - y^5 + 4x^2y^3)2(x^2 + y^2)2y}{(x^2 + y^2)^4} \\ &= \frac{(x^4 - 5y^4 + 12x^2y^2)(x^2 + y^2) - 4y(x^4y - y^5 + 4x^2y^3)}{(x^2 + y^2)^3}, \end{aligned}$$

temos que $\frac{\partial^2 f}{\partial y \partial x}(t, t) = \frac{8t^4 \cdot 2t^2 - 16t^6}{8t^6} = 0$ e, portanto, $\lim_{t \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(t, t) = 0 \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$.

Como f é de classe C^∞ em $\mathbb{R}^2 - \{(0, 0)\}$, e, portanto, duas vezes diferenciável em todos os pontos $(x, y) \neq (0, 0)$, temos que $\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y)$ para todo $(x, y) \neq (0, 0)$.

Logo $\lim_{t \rightarrow 0} \frac{\partial^2 f}{\partial x \partial y}(t, t) = 0 \neq \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$. $\square$

Daremos, agora, outra versão do Teorema de Schwarz que decorre da Regra de Leibniz.

Teorema 7.2. *Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função tal que existem $\frac{\partial f}{\partial x_i}$ e $\frac{\partial^2 f}{\partial x_i \partial x_j}$ em todos os pontos de U , e as funções $\frac{\partial f}{\partial x_j}, \frac{\partial^2 f}{\partial x_i \partial x_j} : U \rightarrow \mathbb{R}$ são contínuas. Então, a derivada $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existe em todos os pontos de U e $\frac{\partial^2 f}{\partial x_j \partial x_i} \equiv \frac{\partial^2 f}{\partial x_i \partial x_j}$.*

Prova.

Vamos supor $n = 2$ para simplificar a notação.

Dado $(x_0, y_0) \in U$, existe $\varepsilon > 0$ tal que $I \times J \subset U$, onde $I = (x_0 - \varepsilon, x_0 + \varepsilon)$ e $J = (y_0 - \varepsilon, y_0 + \varepsilon)$. Seja $b \in J$. Pelo Teorema Fundamental do Cálculo, temos que

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

para todo $(x, y) \in I \times J$, uma vez que $\frac{\partial f}{\partial y}$ é contínua, e, portanto, integrável.

Como $\frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x \partial y} : I \times J \rightarrow \mathbb{R}$ são contínuas, por hipótese, temos, pela Regra de Leibniz, que:

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

Logo, como o integrando $\frac{\partial^2 f}{\partial x \partial y}$ é contínuo, temos, também, que a função $\int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t) dt$ é derivável em relação a y e

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

Assim, $\frac{\partial f}{\partial x}$ possui derivada em relação a y e $\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y)$ para todo $(x, y) \in I \times J$. $\blacksquare$

Observação 7.2. *Seja $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função três vezes diferenciável. Então as seis derivadas mistas de terceira ordem satisfazem:*

$$\frac{\partial^3 f}{\partial x \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x \partial x} \quad \text{e} \quad \frac{\partial^3 f}{\partial y \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial y}.$$

De fato, pelo Teorema de Schwarz,

$$\frac{\partial^3 f}{\partial x \partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial^2 f}{\partial x \partial y} \right) = \frac{\partial}{\partial x} \left(\frac{\partial^2 f}{\partial y \partial x} \right) = \frac{\partial^3 f}{\partial x \partial y \partial x},$$

e, fazendo $g = \frac{\partial f}{\partial x}$, temos que

$$\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial}{\partial x} \left(\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) \right) = \frac{\partial^2 g}{\partial x \partial y} = \frac{\partial^2 g}{\partial y \partial x} = \frac{\partial}{\partial y} \left(\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) \right) = \frac{\partial^3 f}{\partial y \partial x \partial x},$$

uma vez que f e g são duas vezes diferenciáveis.

Analogamente, podemos provar as outras três igualdades acima.

No caso geral, se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função p -vezes diferenciável no aberto U , então para toda sequência de inteiros não-negativos $i_1, \dots, i_n$, com $i_1 + \dots + i_n = \alpha \leq p$, a derivada de ordem α , $\frac{\partial^\alpha}{\partial x_1^{i_1} \dots \partial x_n^{i_n}}$, que consiste em derivar i_1 vezes em relação à variável $x_1, \dots, i_n$ vezes em relação à variável x_n , não depende da ordem em que essas derivações foram efetuadas.

Para demonstrar o caso geral, basta sabermos que podemos trocar a ordem de duas derivadas sucessivas e que qualquer mudança de ordem numa sequência finita pode ser obtida por transposições sucessivas entre dois termos consecutivos da sequência.

8 Fórmula de Taylor; pontos críticos.

Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função p -vezes diferenciável no ponto a . Para cada vetor $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, escrevemos:

$$\begin{aligned} df(a)v &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \alpha_i; \\ d^2f(a)v^2 &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \alpha_i \alpha_j; \\ &\vdots \\ d^pf(a)v^p &= \sum_{i_1, \dots, i_p=1}^n \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(a) \alpha_{i_1} \dots \alpha_{i_p}; \end{aligned}$$

Para cada $p > 0$, a forma $d^pf(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ chama-se *p -ésima diferencial da função f no ponto a* .

Observação 8.1. $df^p(a)(tv)^p = t^p d^pf(a)v^p$, ou seja, $df^p(a)$ é um polinômio homogêneo de grau p nas coordenadas de v .

Observação 8.2. Usando a notação acima, a Regra da Cadeia enuncia-se do seguinte modo:

Seja $f = (f_1, \dots, f_n) : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ uma aplicação tal que $f_i : U \rightarrow \mathbb{R}$ é diferenciável em a

para todo $i = 1, \dots, n$, e seja $g : V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável em $f(a) = b$, com $f(U) \subset V$. Então $g \circ f : U \rightarrow \mathbb{R}$ é diferenciável em a e, para todo $v \in \mathbb{R}^n$,

$$d(g \circ f)(a)v = dg(f(a)) \cdot (df_1(a)v, \dots, df_n(a)v) = dg(f(a)) df(a)v,$$

De fato,

$$\begin{aligned} d(g \circ f)(a)v &= \sum_{i=1}^m \frac{\partial(g \circ f)}{\partial x_i}(a) \alpha_i = \sum_{i=1}^m \left(\sum_{k=1}^n \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_i}(a) \right) \alpha_i \\ &= \sum_{k=1}^n \sum_{i=1}^m \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_i}(a) \alpha_i = \sum_{k=1}^m \frac{\partial g}{\partial y_k}(f(a)) df_k(a)v \\ &= dg(f(a))(df_1(a)v, \dots, df_n(a)v). \end{aligned}$$

Teorema 8.1. (Fórmula de Taylor com resto de Lagrange)

Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^p , $(p+1)$ -vezes diferenciável no segmento aberto $(a, a+v)$, com $[a, a+v] \subset U$. Então existe $\theta \in (0, 1)$ tal que:

$$r_p(v) = \frac{1}{(p+1)!} df^{(p+1)}(a + \theta v) v^{p+1},$$

onde $r_p(v)$ é dado pela igualdade:

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

Prova.

Seja $\varepsilon > 0$ tal que $a + tv \in U$ para todo $t \in (-\varepsilon, 1 + \varepsilon)$, e seja $\lambda : (-\varepsilon, 1 + \varepsilon) \rightarrow \mathbb{R}^n$ o caminho C^∞ dado por $\lambda(t) = a + tv$. Então a função $\varphi = f \circ \lambda : (-\varepsilon, 1 + \varepsilon) \rightarrow \mathbb{R}$ é de classe C^p em $(-\varepsilon, 1 + \varepsilon)$ e é $(p+1)$ -vezes diferenciável em $(0, 1)$.

Logo, pela Fórmula de Taylor com resto de Lagrange para uma função real de uma variável real, existe $\theta \in (0, 1)$, tal que

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{\varphi''(0)}{2!} + \dots + \frac{\varphi^{(p)}(0)}{p!} + r_p,$$

$$\text{onde } r_p = \frac{\varphi^{(p+1)}(\theta)}{(p+1)!}. \quad (I)$$

Afirmção: $\varphi^{(i)}(t) = d^{(i)}f(a + tv)v^i$, $1 \leq i \leq p+1$, $t \in (0, 1)$.

De fato,

$$\varphi'(t) = \frac{\partial f}{\partial v}(a + tv) = df(a + tv)v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + tv) \alpha_i.$$

Suponhamos, por indução, o resultado válido para uma função p -vezes diferenciável.

Seja $f : U \rightarrow \mathbb{R}$ uma função $(p+1)$ -vezes diferenciável em $(a, a+v)$. Então $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ é p -vezes diferenciável, para todo $i = 1, \dots, n$.

Portanto, pela hipótese de indução, $\lambda_j^{(i)}(t) = d^i \left(\frac{\partial f}{\partial x_j} \right) (a + tv) v^i$, $i = 1, \dots, p$, onde $\lambda_j(t) = \frac{\partial f}{\partial x_j}(a + tv)$. Assim,

$$\begin{aligned} \varphi^{(k+1)}(t) &= \sum_{j=1}^n \lambda_j^{(k)}(t) \alpha_j = \sum_{j=1}^n \left(d^k \left(\frac{\partial f}{\partial x_j} \right) (a + tv) v^k \right) \alpha_j \\ &= \sum_{j=1}^n \left(\sum_{j_1, \dots, j_k=1}^n \frac{\partial^k \left(\frac{\partial f}{\partial x_j} \right)}{\partial x_{j_1} \dots \partial x_{j_k}} (a + tv) \alpha_{j_1} \dots \alpha_{j_k} \right) \alpha_j \\ &= \sum_{j, j_1, \dots, j_k=1}^n \frac{\partial^{k+1} f}{\partial x_{j_1} \dots \partial x_{j_k} \partial x_j} (a + tv) \alpha_{j_1} \dots \alpha_{j_k} \alpha_j \\ &= d^{k+1} f(a + tv) v^{k+1} \end{aligned}$$

para todo $k = 1, \dots, p$ e todo $v \in \mathbb{R}^n$.

• Como $\varphi(1) = f(a + v)$, $\varphi(0) = f(a)$, $\varphi^{(i)}(0) = d^i f(a) v^i$ e $\varphi^{(p+1)}(\theta) = d^{p+1} f(a + \theta v) v^{p+1}$, temos, por (I), que a fórmula de Taylor com resto de Lagrange também é válida para funções reais de n -variáveis. ■

Teorema 8.2. (Fórmula de Taylor com resto integral)

Se $f : U \rightarrow \mathbb{R}$ é uma função de classe C^{p+1} e $[a, a + v] \subset U$, então

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

Prova.

Como $\varphi = f \circ \lambda$ é de classe C^{p+1} em $(-\varepsilon, 1 + \varepsilon)$, temos, pela Fórmula de Taylor com resto integral para funções reais de uma variável real, que

$$\varphi(1) = \varphi(0) + \varphi'(0) + \dots + \frac{\varphi^{(p)}(0)}{p!} + r_p,$$

onde $r_p = \frac{1}{p!} \int_0^1 (1-t)^p \varphi^{(p+1)}(t) dt$.

Logo,

$$f(a + v) = f(a) + df(a) v + \dots + \frac{1}{p!} d^{(p)}(a) v^p + r_p(v),$$

onde

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

Antes de provarmos a Fórmula de Taylor Infinitesimal, faremos algumas considerações de

caráter geral.

Definição 8.1. Seja $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ o produto cartesiano de k —cópias do espaço $\mathbb{R}^n$ e seja $L : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ uma transformação k —linear. Dizemos que L é *simétrica* se

$$L(v_1, \dots, \widehat{v}_i, \dots, \widehat{v}_j, \dots, v_k) = L(v_1, \dots, \widehat{v}_j, \dots, \widehat{v}_i, \dots, v_k),$$

para quaisquer $v_1, \dots, v_k \in \mathbb{R}^n$ e todo par $i, j = 1, \dots, n$, com $i < j$.

Então, se $v_j = (\alpha_1^j, \dots, \alpha_n^j)$, $j = 1, \dots, k$, temos

$$L(v_1, \dots, v_k) = \sum a_{i_1, \dots, i_k} \alpha_{i_1}^1 \dots \alpha_{i_k}^k,$$

onde $a_{i_1, \dots, i_k} = L(e_{i_1}, \dots, e_{i_k})$ independe da ordem dos índices $i_1, \dots, i_k = 1, \dots, n$.

Observação 8.3. Se $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ é uma função p —vezes diferenciável no ponto a , a transformação k —linear $d^k f(a) : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ definida por:

$$d^k f(a)(v_1, \dots, v_k) = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f(a)}{\partial x_{i_1} \dots \partial x_{i_k}} \alpha_{i_1}^1 \dots \alpha_{i_k}^k,$$

chama-se *k —ésima diferencial* da função f no ponto a , para $k = 1, \dots, p$.

Por Schwarz, temos que $d^k f(a)$ é simétrica, $1 \leq k \leq p$.

Observe que $d^k f(a)v^k = d^k f(a)(v, \dots, v)$ é a forma associada à aplicação k —linear $d^k f(a)$.

Definição 8.2. Dizemos que uma função $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ é *k —homogênea* quando $f(tx) = t^k f(x)$ para todo $x \in \mathbb{R}^n$ e $t \in \mathbb{R}$.

Exemplo 8.1. Se $L : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ é k —linear, então $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ definida por $g(x) = L(x, \dots, x)$ é k —homogênea, ou melhor, g é um polinômio homogêneo de grau k e, portanto, g é C^∞ . $\square$

Observação 8.4. Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ uma função k —homogênea de classe C^k , $k \geq 1$.

Afirmção 1: $\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}$ é uma função $(k-j)$ —homogênea para todo $1 \leq j \leq k$ e para quaisquer $i_1, \dots, i_j = 1, \dots, n$.

Como $f(tx) = t^k f(x)$, temos, pela Regra da Cadeia, que

$$\frac{\partial f}{\partial x_i}(tx) t = t^k \frac{\partial f}{\partial x_i}(x),$$

para todo $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, $i = 1, \dots, n$.

Logo, se $t \neq 0$, $\frac{\partial f}{\partial x_i}(tx) = t^{k-1} \frac{\partial f}{\partial x_i}(x)$ para todo $x \in \mathbb{R}^n$. Como $f \in C^k$, $k \geq 1$, temos que

$$\frac{\partial f}{\partial x_i}(tx) = t^{k-1} \frac{\partial f}{\partial x_i}(x) \text{ para todo } x \in \mathbb{R}^n, t \in \mathbb{R}.$$

Suponhamos, por indução, que o resultado é válido para funções $k - 1$ homogêneas, $k - 1 \geq 1$.

Sendo $f \in C^k$, temos que $\frac{\partial f}{\partial x_i} : \mathbb{R}^n \longrightarrow \mathbb{R}$ são de classe C^{k-1} e $(k - 1)$ -homogêneas, para todo $i = 1, \dots, n$.

Logo, pela hipótese de indução, para cada $i = 1, \dots, n$, temos que:

$$\frac{\partial^j \left(\frac{\partial f}{\partial x_i} \right)}{\partial x_{i_1} \dots \partial x_{i_j}} (tx) = t^{k-1-j} \frac{\partial^j \left(\frac{\partial f}{\partial x_i} \right)}{\partial x_{i_1} \dots \partial x_{i_j}} (x),$$

para quaisquer $i_1, \dots, i_j = 1, \dots, n$ e para todo $j = 1, \dots, k - 1$. Ou seja,

$$\frac{\partial^{j+1} f}{\partial x_{i_1} \dots \partial x_{i_j} \partial x_i} (tx) = t^{k-1-j} \frac{\partial^{j+1} f}{\partial x_{i_1} \dots \partial x_{i_j} \partial x_i} (x),$$

para todo $j + 1 = 2, \dots, k$, e para quaisquer $i_1, \dots, i_j, i = 1, \dots, n$.

• Logo, se $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ é uma função k -homogênea de classe C^k , então

$$d^j f(tx)(v_1, \dots, v_j) = t^{k-j} d^j f(x)(v_1, \dots, v_j)$$

para todo $j = 1, \dots, k$.

Assim, $d^k f(tx)(v_1, \dots, v_k) = d^k f(x)(v_1, \dots, v_k)$ para todo $t \in \mathbb{R}$ e todo $x \in \mathbb{R}^n$. Em particular, $d^k f(x) = d^k f(0)$ independe do ponto $x \in \mathbb{R}^n$.

Como $\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x) = \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(0)$ para todo $x \in \mathbb{R}$ e para quaisquer $i_1, \dots, i_k = 1, \dots, n$, temos que todas as derivadas parciais de ordem k de f são constantes.

Logo f é de classe C^∞ e $d^j f(x) = 0$ para todo $j > k$ e para todo $x \in \mathbb{R}^n$.

Afirmção 2: $d^k f(0)x^k = k! f(x)$ e $d^j f(0)x^j = 0$, se $j \neq k$. (II)

De fato, seja $\varphi(t) = f(tx) = t^k f(x)$. Então, como foi provado no Teorema 8.1, temos:

$$\varphi^{(i)}(t) = d^i f(tx)x^i, \quad \text{para todo } i \in \mathbb{N}.$$

Mas, por outro lado, $\varphi^{(i)}(t) = \frac{k!}{(k-i)!} t^{k-i} f(x)$, para todo $1 \leq i \leq k$, e $\varphi^{(j)}(t) = 0$ para $j > k$.

Logo $d^i f(0)x^i = 0$ para $i \neq k$ e $d^k f(0)x^k = k! f(x)$.

Então $f(x) = L(x, \dots, x)$, onde $L = \frac{1}{k!} d^k f(0)$ é uma transformação k -linear simétrica.

Como $d^k f(x) = d^k f(0)$ para todo $x \in \mathbb{R}^n$, temos que $d^k f(x) = k! L$ para todo $x \in \mathbb{R}^n$.

• Se f não é de classe C^k , f não é necessariamente a forma associada a uma transformação k -linear simétrica.

Exemplo 8.2. Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ a função definida por $f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2}$, $(x, y) \neq (0, 0)$, e $f(0, 0) = 0$.

Então $f(tx, ty) = t^2 f(x, y)$ para todo $t \in \mathbb{R}$ e todo $(x, y) \in \mathbb{R}^2$, ou seja, f é uma função 2-homogênea. Mas, f não é a forma quadrática de uma transformação bilinear. Isso ocorre porque f é de classe C^1 , mas f não é duas vezes diferenciável na origem (verifique!). $\square$

Afirmção 3: $d^j f(x)(v_1, \dots, v_j) = \frac{1}{(k-j)!} d^k f(0)(x, \dots, x, v_1, \dots, v_j)$ para todo $1 \leq j \leq k$.

Sejam $1 \leq j \leq k$ e $g(x) = \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x)$, onde $i_1, \dots, i_j \in \{1, \dots, n\}$. Como

$$\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(tx) = t^{k-j} \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x),$$

temos que g é $(k-j)$ -homogênea e, portanto, por (II), $d^{(k-j)} g(0)x^{k-j} = (k-j)! g(x)$, ou seja,

$$d^{k-j} \left(\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}} \right) (0)x^{k-j} = (k-j)! \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x),$$

para todo $x \in \mathbb{R}^n$ e quaisquer $i_1, \dots, i_j = 1, \dots, n$.

Logo, sendo $v_\ell = (\alpha_1^\ell, \dots, \alpha_n^\ell)$, $\ell = 1, \dots, j$, temos que:

$$\begin{aligned} d^j f(x)(v_1, \dots, v_j) &= \sum_{i_1, \dots, i_j=1}^n \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x) \alpha_{i_1}^1 \dots \alpha_{i_j}^j \\ &= \frac{1}{(k-j)!} \sum_{i_1, \dots, i_j=1}^n \left(\sum_{\ell_1, \dots, \ell_{k-j}=1}^n \frac{\partial^k f(0)}{\partial x_{\ell_1} \dots \partial x_{\ell_{k-j}} \partial x_{i_1} \dots \partial x_{i_j}} x_{\ell_1} \dots x_{\ell_{k-j}} \right) \alpha_{i_1}^1 \dots \alpha_{i_j}^j \\ &= \frac{1}{(k-j)!} d^k f(0)(x, \dots, x, v_1, \dots, v_j). \end{aligned}$$

• Em particular, seja $T : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ uma transformação k -linear e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $f(x) = T(x, \dots, x)$. Então, como f é k -homogênea e de classe C^∞ , temos, por (II), que

$$f(x) = T(x, \dots, x) = \frac{1}{k!} d^k f(0)(x, \dots, x),$$

ou seja,

$$d^k f(0)(x, \dots, x) = k! T(x, \dots, x). \quad (\text{III})$$

• Dada uma transformação k -linear $T : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$, a transformação k -linear $T_S = \sum_{\sigma \in \mathcal{P}} T_\sigma$, onde $\mathcal{P}$ é o conjunto de todas as permutações de $\{1, \dots, k\}$ e $T_\sigma(v_1, \dots, v_k) = T(v_{\sigma(1)}, \dots, v_{\sigma(k)})$, é chamada **simetrização da transformação T** .

Observe que T_S é k -linear **simétrica** e $T_S(x, \dots, x) = k! T(x, \dots, x)$.

Então, por (III),

$$d^k f(0)(x, \dots, x) = T_S(x, \dots, x). \quad (\text{IV})$$

Afirmção 4: $d^k f(x) = d^k f(0) = T_S$. Em particular $d^k f(x) = d^k f(0) = k! T$, se T é simétrica.

De fato, por (IV), basta mostrar que se $U : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ é uma transformação k -linear simétrica tal que $g(x) = U(x, \dots, x) = 0$ para todo $x \in \mathbb{R}^n$, então $U \equiv 0$.

Vamos fazer a prova deste fato usando indução em $k \in \mathbb{N}$.

Se $k = 1$, a afirmação é evidente.

Suponhamos o resultado válido para transformações $(k - 1)$ -lineares, $k - 1 \geq 1$.

Seja $U : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ uma transformação k -linear simétrica tal que $U(x, \dots, x) = 0$ para todo $x \in \mathbb{R}^n$.

Sejam $v, w \in \mathbb{R}^n$ e $t \in \mathbb{R}$. Então,

$$\begin{aligned} 0 &= U(v + tw, v + tw, \dots, v + tw) = t^{k-1} \binom{k}{k-1} U(v, w, \dots, w) \\ &\quad + t^{k-2} \binom{k}{k-2} U(v, v, w, \dots, w) + \dots + t \binom{k}{1} U(v, \dots, v, w), \end{aligned}$$

para todo $t \in \mathbb{R}$.

Logo $U(v, w, \dots, w) = 0$ para quaisquer $v, w \in \mathbb{R}^n$.

Seja $v \in \mathbb{R}^n$ e defina $U_1 : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ por $U_1(v_1, \dots, v_{k-1}) = U(v, v_1, \dots, v_{k-1})$. Então U_1 é uma transformação $(k - 1)$ -linear simétrica tal que $U_1(w, \dots, w) = U(v, w, \dots, w) = 0$ para todo $w \in \mathbb{R}^n$.

Logo, pela hipótese de indução, $U_1 \equiv 0$, ou seja, $U_1(v_1, \dots, v_{k-1}) = 0$ para quaisquer $k - 1$ vetores $v_1, \dots, v_{k-1} \in \mathbb{R}^n$. Então $U(v, v_1, \dots, v_{k-1}) = 0$ para quaisquer $v, v_1, \dots, v_{k-1} \in \mathbb{R}^n$. Assim $U \equiv 0$.

• Resumindo, se $T : \mathbb{R}^n \times \dots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ é uma transformação k -linear e $f(x) = T(x, \dots, x)$, então para todo $x \in \mathbb{R}^n$:

◦ $d^k f(x) = d^k f(0) = T_S$;

◦ $d^j f(x) = 0$, se $j > k$.

◦ $d^j f(x)(v_1, \dots, v_j) = \frac{1}{(k-j)!} T_S(x, \dots, x, v_1, \dots, v_j)$, se $1 \leq j \leq k$, quaisquer que sejam $v_1, \dots, v_j \in \mathbb{R}^n$.

◦ $d^j f(0) = 0$, se $1 \leq j < k$. ◻

• Passamos, agora, a analisar a *Fórmula de Taylor Infinitesimal*.

Se $f : U \longrightarrow \mathbb{R}$ é p -vezes diferenciável no ponto $a \in U$, então $\lim_{v \rightarrow 0} \frac{r_p(v)}{\|v\|^p} = 0$, onde $r_p : U_0 = \{v \in \mathbb{R}^n; a + v \in U\} \longrightarrow \mathbb{R}$ é dada por:

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

De fato, seja $g : U_0 \rightarrow \mathbb{R}$ dada por $g(v) = f(a + v)$. Então g é p -vezes diferenciável na origem, pois a função $v \mapsto a + v$ é de classe C^∞ e f é p -vezes diferenciável em a .

Afirmção: $d^jg(0) = d^jf(a)$, $1 \leq j \leq p$.

Basta mostrar, por indução, que $\frac{\partial^k g}{\partial x_{i_1} \dots \partial x_{i_k}}(0) = \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a)$, para todo $1 \leq k \leq p$ e para quaisquer $i_1, \dots, i_k \in \{1, \dots, n\}$.

Para $j = 1$, temos, pela Regra da Cadeia (ver observação 8.2), que $dg(0)v = df(a)v$ para todo $v \in \mathbb{R}^n$, ou seja, $\frac{\partial g}{\partial x_i}(0) = \frac{\partial f}{\partial x_i}(a)$ para todo $i = 1, \dots, n$.

Suponhamos que o resultado seja válido para funções $(p - 1)$ -vezes diferenciáveis no ponto $a \in U$, $p - 1 \geq 1$. Seja f uma função p -vezes diferenciável no ponto a . Então $\frac{\partial f}{\partial x_i}$ é $(p - 1)$ -vezes diferenciável no ponto a , para todo $i = 1, \dots, n$.

Pela hipótese de indução, a função h dada por $h(v) = \frac{\partial f}{\partial x_i}(a + v)$, $v \in U_0$, é $(p - 1)$ -vezes diferenciável na origem e

$$\frac{\partial^k h}{\partial x_{i_1} \dots \partial x_{i_k}}(0) = \frac{\partial^k \left(\frac{\partial f}{\partial x_i} \right)}{\partial x_{i_1} \dots \partial x_{i_k}}(a), \quad (V)$$

para $1 \leq k \leq p - 1$, quaisquer que sejam $i_1, \dots, i_k \in \{1, \dots, n\}$.

Logo, como $p \geq 2$, temos que f é diferenciável numa vizinhança do ponto a e, portanto, $\frac{\partial g}{\partial x_i}(v) = \frac{\partial f}{\partial x_i}(a + v)$ para todo $i = 1, \dots, n$, e todo v numa vizinhança da origem.

Assim, $h(v) = \frac{\partial g}{\partial x_i}(v)$ e, por (V),

$$\frac{\partial^k \left(\frac{\partial g}{\partial x_i} \right)}{\partial x_{i_1} \dots \partial x_{i_k}}(0) = \frac{\partial^{k+1} f}{\partial x_{i_1} \dots \partial x_{i_k} \partial x_i}(a),$$

ou seja,

$$\frac{\partial^{k+1} g}{\partial x_{i_1} \dots \partial x_{i_k} \partial x_i}(0) = \frac{\partial^{k+1} f}{\partial x_{i_1} \dots \partial x_{i_k} \partial x_i}(a),$$

para todo $k + 1 = 2, \dots, p$ e quaisquer $i_1, \dots, i_k, i \in \{1, \dots, n\}$.

• Sendo $H_k : \mathbb{R}^n \rightarrow \mathbb{R}$, $H_k(v) = d^kf(a)v^k$, $1 \leq k \leq p$, temos, pelo provado na observação 8.4, que $d^jH_k(0) = 0$ se $j \in \{1, \dots, p\}$ e $j \neq k$, e $d^kH_k(0) = k! d^kf(a)$.

Logo $r_p(0) = 0$ e $d^jr_p(0) = d^jf(a) - d^jf(a) = 0$ para todo $j = 1, \dots, p$.

Lema 8.1. *Seja $r : U_0 \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função p -vezes diferenciável no ponto $0 \in U_0$. Então $r(0) = dr(0) = \dots = d^pr(0) = 0$ se, e somente se, $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^p} = 0$.*

Prova.

($\Rightarrow$) Para $p = 0$, estamos supondo r contínua no ponto 0.

Para $p = 1$, r é diferenciável na origem e $r(0) = dr(0) = 0$. Logo, como

$$r(v) = r(0) + dr(0)v + \rho(v)\|v\|,$$

com $\lim_{v \rightarrow 0} \rho(v) = 0$, temos que $\rho(v) = \frac{r(v)}{\|v\|}$, e, portanto, $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0$.

Suponhamos que o resultado é válido para funções $(p - 1)$ -vezes diferenciáveis na origem, $p - 1 \geq 1$.

Seja $r : U_0 \rightarrow \mathbb{R}$ uma função p -vezes diferenciável na origem com $r(0) = dr(0) = \dots = d^p r(0) = 0$.

Então, para todo $1 \leq i \leq n$, $\varphi_i = \frac{\partial r}{\partial x_i} : U_0 \rightarrow \mathbb{R}$ é $(p - 1)$ -vezes diferenciável na origem e

$\varphi_i(0) = d\varphi_i(0) = \dots = d^{p-1}\varphi_i(0)$. Logo, pela hipótese de indução, $\lim_{v \rightarrow 0} \frac{\frac{\partial r}{\partial x_i}(v)}{\|v\|^{p-1}} = 0$.

Como $p \geq 2$, r é diferenciável numa vizinhança $V_0 \subset U_0$ da origem e, portanto, pelo teorema do valor médio, para todo $v \in U_0$, existe $\theta_v \in (0, 1)$ tal que

$$\frac{r(v)}{\|v\|^p} = \frac{\sum_{i=1}^n \frac{\partial r}{\partial x_i}(\theta_v v) \alpha_i}{\|v\|^p} = \sum_{i=1}^n \left[\left(\frac{\frac{\partial r}{\partial x_i}(\theta_v v)}{\|\theta_v v\|^{p-1}} \right) \frac{\alpha_i}{\|v\|} |\theta_v|^{p-1} \right].$$

Considerando $\mathbb{R}^n$ com a norma do máximo, temos que $\left| \frac{\alpha_i}{\|v\|} \right| \leq 1$, para todo $i = 1, \dots, n$.

Logo $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^p} = 0$, uma vez que $\lim_{v \rightarrow 0} \frac{\frac{\partial r}{\partial x_i}(\theta_v v)}{\|\theta_v v\|^{p-1}} = 0$, para todo $i = 1, \dots, n$.

($\Leftarrow$) Para $p = 0$, $\lim_{v \rightarrow 0} r(v) = 0$, e, portanto, $r(0) = 0$, pois estamos supondo r contínua na origem.

Para $p = 1$, $\lim_{v \rightarrow 0} r(v) = \lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} \|v\| = 0$. Então $r(0) = 0$, pois r é contínua na origem, uma vez que r é diferenciável neste ponto. Além disso, como f é diferenciável na origem,

$$r(v) = r(0) + dr(0)v + \bar{r}(v) = dr(0)v + \bar{r}(v),$$

onde $\lim_{v \rightarrow 0} \frac{\bar{r}(v)}{\|v\|} = 0$. Logo, para todo $v \in \mathbb{R}^n - \{0\}$ e para todo $t \in \mathbb{R} - \{0\}$, $\frac{r(tv)}{t} = dr(0)v + \frac{\bar{r}(tv)}{t}$.

Como

$$\lim_{t \rightarrow 0} \frac{r(tv)}{\|tv\|} = \lim_{t \rightarrow 0} \frac{\bar{r}(tv)}{\|tv\|} = 0,$$

temos que

$$dr(0)v = \lim_{t \rightarrow 0} \frac{r(tv)}{t} - \lim_{t \rightarrow 0} \frac{\bar{r}(tv)}{t} = \lim_{t \rightarrow 0} \pm \|v\| \left(\frac{r(tv)}{\|tv\|} - \frac{\bar{r}(tv)}{\|tv\|} \right) = 0,$$

para todo $v \in \mathbb{R}^n - \{0\}$. Logo $dr(0) = 0$.

Suponhamos que o resultado é válido para funções p -vezes diferenciáveis no ponto 0, $p \geq 1$.

Seja $r : U_0 \rightarrow \mathbb{R}$ uma função $(p+1)$ -vezes diferenciável na origem com $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^{p+1}} = 0$. Como

$$\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^p} = \lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^{p+1}} \|v\| = 0, \text{ temos, pela hipótese de indução, que}$$

$$r(0) = dr(0) = \dots = d^p r(0) = 0.$$

Mostraremos, agora, que $d^{p+1} r(0) = 0$.

De fato, pelo provado na primeira parte do lema, temos que

$$\lim_{v \rightarrow 0} \frac{r(v) - \frac{1}{(p+1)!} d^{p+1} r(0) v^{p+1}}{\|v\|^{p+1}} = 0,$$

já que $d^j \varphi(0) = 0$, $j = 1, \dots, p$, e $d^{p+1} \varphi(0) = (p+1)! d^{p+1} r(0)$, onde $\varphi(v) = d^{p+1} r(0) v^{p+1}$.

Então, para todo $v \in \mathbb{R}^n - \{0\}$,

$$\lim_{t \rightarrow 0^+} \frac{\left(r(tv) - \frac{1}{(p+1)!} d^{p+1} r(0) (tv)^{p+1} \right)}{\|tv\|^{p+1}} = 0,$$

e, portanto,

$$\frac{1}{(p+1)!} \frac{d^{p+1} r(0) v^{p+1}}{\|v\|^{p+1}} = \lim_{t \rightarrow 0^+} \frac{r(tv)}{\|tv\|^{p+1}} = 0.$$

Ou seja, $d^{p+1} r(0) v^{p+1} = 0$ para todo $v \in \mathbb{R}^n$. Então $d^{p+1} r(0) = 0$. ■

Observação 8.5. (Unicidade da Fórmula de Taylor)

Seja $f : U \rightarrow \mathbb{R}$ uma função p -vezes diferenciável no ponto $a \in U$ e, para cada $i = 1, \dots, p$, seja $\varphi_i : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ uma função i -linear. Se

$$f(a+v) = f(a) + \varphi_1 v + \varphi_2 v^2 + \dots + \varphi_p v^p + r_p(v),$$

com $\lim_{v \rightarrow 0} \frac{r_p(v)}{\|v\|^p} = 0$, então $\varphi_i v^i = \frac{1}{i!} d^i f(a) v^i$, para todo $i = 1, \dots, p$ e todo $v \in \mathbb{R}^n$.

De fato, como r_p é p -vezes diferenciável no ponto 0 e $\lim_{v \rightarrow 0} \frac{r_p(v)}{\|v\|^p} = 0$, temos, pelo lema acima, que $r_p(0) = dr_p(0) = \dots = d^p r_p(0) = 0$. Mas, pela observação 8.4, $d^i r_p(0) = d^i f(a) - \varphi_i^S$, para todo $i = 1, \dots, p$, onde φ_i^S é a simetrização de φ_i . Logo $\varphi_i^S = d^i f(a)$, ou seja,

$$\varphi_i v^i = \frac{1}{i!} \varphi_i^S v^i = \frac{1}{i!} d^i f(a) v^i,$$

para todo $i = 1, \dots, p$.

Definição 8.3. Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes diferenciável no ponto $a \in U$. A **forma Hessiana** $Hf(a)$, de f no ponto a é a forma quadrática da transformação bilinear simétrica $d^2 f(a)$, ou seja,

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

onde $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

- Pelo teorema de Schwarz, a matriz $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)$, chamada *matriz Hessiana de f no ponto a*, é simétrica.

Definição 8.4. Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável. Um ponto $a \in U$ é um *ponto crítico* de f (ou um *ponto singular*) quando $df(a) = 0$, ou seja, $\frac{\partial f}{\partial x_1}(a) = \dots = \frac{\partial f}{\partial x_n}(a) = 0$.

Definição 8.5. Dizemos que a função f tem um *máximo* (respectivamente, um *mínimo*) *local* no ponto $a \in U$ quando existe $\delta > 0$ tal que

$$\|v\| < \delta \implies f(a+v) \leq f(a) \text{ (respectivamente, } f(a) \leq f(a+v) \text{)}.$$

Observação 8.6. Se $f : U \rightarrow \mathbb{R}$ é diferenciável no ponto $a \in U$ e a é um ponto de máximo local (ou de mínimo local), então a é um ponto crítico de f .

De fato, neste caso o ponto 0 é um ponto de máximo (ou de mínimo) local para as funções reais de uma variável real dadas por: $\varphi_i(t) = f(a + te_i)$, $i = 1, \dots, n$. Logo $\frac{\partial f}{\partial x_i}(a) = \varphi_i'(0) = 0$, para todo $i = 1, \dots, n$.

Então $df(a) = 0$, ou seja, a é um ponto crítico de f .

Definição 8.6. Dizemos que um ponto crítico a de f é *não-degenerado* quando a matriz Hessiana de f no ponto a é invertível, ou seja, $\det \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right) \neq 0$.

Teorema 8.3. Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes diferenciável. Todo ponto crítico não-degenerado $a \in U$ é um ponto crítico isolado.

Este teorema é consequência do seguinte resultado.

Teorema 8.4. Seja $F = (f_1, \dots, f_n) : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função onde cada função coordenada $f_i : U \rightarrow \mathbb{R}$, $i = 1, \dots, n$, é diferenciável no ponto $a \in U$. Se a matriz $H = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{n \times n}$ tem determinante diferente de zero, então existe $\delta > 0$ tal que

$$0 < \|x - a\| < \delta \implies F(x) \neq F(a).$$

A matriz H , referida no teorema acima, é chamada a *matriz Jacobiana de f no ponto a*.

Lema 8.2. Seja $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma transformação linear invertível. Então existe $c > 0$ tal que $\|H(x)\| \geq c\|x\|$ para todo $x \in \mathbb{R}^n$.

Prova.

Seja $\frac{1}{c} = \|H^{-1}\| = \sup \{ \|H^{-1}(x)\| \mid \|x\| = 1 \} > 0$. Então, para todo $x \in \mathbb{R}^n$:

$$\|x\| = \|H^{-1}(H(x))\| \leq \|H^{-1}\| \|H(x)\| = \frac{\|H(x)\|}{c},$$

ou seja, $\|H(x)\| \geq c\|x\|$. ■

Prova.

(Demonstração do teorema 8.4)

Como a função $f_i : U \rightarrow \mathbb{R}$ é diferenciável no ponto a , para cada $i = 1, \dots, n$, temos:

$$f_i(x) = f_i(a) + \sum_{j=1}^n h_{ij}(x_j - a_j) + \rho_i(x)\|x - a\|,$$

onde $\lim_{x \rightarrow a} \rho_i(x) = 0$ e $h_{ij} = \frac{\partial f_i}{\partial x_j}(a)$.

Fazendo $\rho(x) = (\rho_1(x), \dots, \rho_n(x))$, temos que:

$$F(x) = F(a) + H(x - a) + \rho(x)\|x - a\|,$$

onde $\lim_{x \rightarrow a} \rho(x) = 0$.

Pelo lema 8.2, existe $c = \frac{1}{\|H^{-1}\|} > 0$ tal que $\|H(x)\| \geq c\|x\|$ para todo $x \in \mathbb{R}^n$.

Como $\lim_{x \rightarrow a} \rho(x) = 0$, existe $\delta > 0$ tal que $0 < \|x - a\| < \delta \implies \|\rho(x)\| < \frac{c}{2}$.

Logo, se $0 < \|x - a\| < \delta$, obtemos:

$$\begin{aligned} \|F(x) - F(a)\| &= \|H(x - a) + \rho(x)\|x - a\|\| \geq \|H(x - a)\| - \|\rho(x)\|\|x - a\| \\ &\geq c\|x - a\| - \frac{c}{2}\|x - a\| = \frac{c}{2}\|x - a\|, \end{aligned}$$

ou seja, $\|F(x) - F(a)\| \geq \frac{c}{2}\|x - a\|$.

Então $F(x) \neq F(a)$ para todo $x \in U$ tal que $0 < \|x - a\| < \delta$. ■

Prova.

(Demonstração do teorema 8.3)

Seja $F : U \rightarrow \mathbb{R}^n$ dada por $F(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)$. Então F tem funções coordenadas

$f_i = \frac{\partial f}{\partial x_i}$ diferenciáveis no ponto a e a matriz $\left(\frac{\partial f_i}{\partial x_j}(a) \right) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \right)$ é a matriz Hessiana de f no ponto a . Logo, pelo teorema 8.4, existe $\delta > 0$ tal que $0 < \|x - a\| < \delta \implies F(x) \neq F(a) = 0$, ou seja, $\text{grad } f(x) \neq 0$. Provamos, assim, que se $0 < \|x - a\| < \delta$, então x não é um ponto crítico de f . ■

Corolário 8.1. *O conjunto dos pontos críticos não-degenerados de uma função duas vezes diferenciável é enumerável.*

Prova.

Basta lembrar que todo conjunto discreto é enumerável. ■

Corolário 8.2. *Se todos os pontos críticos de uma função $f : U \rightarrow \mathbb{R}$, duas vezes diferenciável, são não-degenerados, então em cada compacto $K \subset U$ há apenas um número finito deles.*

Prova.

Como f é de classe C^1 , o conjunto C dos pontos críticos é um subconjunto fechado de U , pois $C = F^{-1}(0)$, onde F é a função contínua dada por $F(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)$. Logo o conjunto dos pontos críticos de f contidos num compacto $K \subset U$ é fechado em K e é, portanto, compacto. Como $C \cap K$ é compacto e discreto, temos que $C \cap K$ é finito. ■

Definição 8.7. Seja $H : \mathbb{R}^n \rightarrow \mathbb{R}$ a forma quadrática dada por $Hv^2 = \sum_{i,j=1}^n h_{ij} \alpha_i \alpha_j$, onde $h_{ij} = h_{ji}$, $i, j = 1, \dots, n$, e $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

Dizemos que H é **positiva** (respectivamente **negativa**) se $Hv^2 > 0$ (respectivamente $Hv^2 < 0$) para todo $v \in \mathbb{R}^n - \{0\}$.

Se uma forma quadrática é positiva ou negativa, dizemos que ela é **definida**. E dizemos que uma forma quadrática H é **indefinida** quando existem $v, w \in \mathbb{R}^n$ tais que $Hv^2 > 0$ e $Hw^2 < 0$.

Exemplo 8.3. Se $\langle \cdot, \cdot \rangle$ é um produto interno de $\mathbb{R}^n$, a forma quadrática $Hv^2 = \langle v, v \rangle$ é positiva, e a forma quadrática $Hv^2 = -\langle v, v \rangle$ é negativa.

E, para todo $i = 1, \dots, n-1$, a forma quadrática

$$Hv^2 = \alpha_1^2 + \dots + \alpha_i^2 - \alpha_{i+1}^2 - \dots - \alpha_n^2,$$

é indefinida. □

Observação 8.7.

- H é positiva se, e somente se, todos os autovalores da matriz simétrica (h_{ij}) são positivos.
- H é negativa se, e somente se, todos os autovalores da matriz simétrica (h_{ij}) são negativos.

Em particular, se H é definida então $\det(h_{ij}) \neq 0$, ou seja, a matriz (h_{ij}) é invertível.

Podemos também provar isto, observando que se $Hv^2 \neq 0$ para todo $v \in \mathbb{R}^n - \{0\}$ então $Hv^2 = \langle H_0 v, v \rangle \neq 0$ para todo $v \in \mathbb{R}^n - \{0\}$, onde $H_0 = (h_{ij})$. Logo $H_0 v \neq 0$ para todo $v \in \mathbb{R}^n - \{0\}$ e, portanto, H_0 é invertível.

- H é indefinida se, e somente se, $H_0 = (h_{ij})$ possui um autovalor positivo e outro negativo.

Observação 8.8. Se f é duas vezes diferenciável no ponto a , $df(a) = 0$ e $Hf(a)$ é positiva ou negativa, então a é um ponto crítico não-degenerado.

Teorema 8.5. Sejam $f : U \rightarrow \mathbb{R}$ uma função duas vezes diferenciável no ponto crítico $a \in U$ e H a forma quadrática Hessiana de f no ponto a . Então:

- (1) Se H é positiva, a é ponto de mínimo local não-degenerado;
- (2) Se H é negativa, a é ponto de máximo local não-degenerado;
- (3) Se H é indefinida, a não é ponto de mínimo local nem de máximo local de f .

Prova.

Seja $\delta_0 > 0$ tal que $B_{\delta_0}(a) \subset U$. Então $a + v \in U$ se $0 < \|v\| < \delta_0$.

Para todo $v \in \mathbb{R}^n$, com $0 < \|v\| < \delta_0$, temos

$$f(a + v) = f(a) + \frac{1}{2} H v^2 + r(v) = f(a) + \left[\frac{1}{2} H \left(\frac{v}{\|v\|} \right)^2 + \frac{r(v)}{\|r(v)\|^2} \right] \|v\|^2. \quad (*)$$

Como a função $\varphi_0 : \mathbb{R}^n \rightarrow \mathbb{R}$, $\varphi_0(v) = H v^2$ é contínua e $S^{n-1} = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$ é compacto, temos que se H é positiva, existe $c > 0$ tal que $\varphi_0(u) \geq c$ para todo $u \in S^{n-1}$.

Logo $H \left(\frac{v}{\|v\|} \right)^2 \geq c$ para todo $v \in \mathbb{R}^n - \{0\}$.

Além disso, temos que $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|^2} = 0$, pois f é duas vezes diferenciável no ponto a . Logo existe

$0 < \delta < \delta_0$, tal que $0 < \|v\| < \delta \implies \left| \frac{r(v)}{\|v\|^2} \right| < \frac{c}{4}$.

Assim, $f(a + v) - f(a) \geq \left(\frac{c}{2} - \frac{c}{4} \right) \|v\|^2 = \frac{c}{4} \|v\|^2 > 0$ para todo $0 < \|v\| < \delta$, ou seja, $f(a + v) > f(a)$ para todo $0 < \|v\| < \delta$. Então a é um ponto de mínimo local para f .

A afirmação (2) prova-se de modo análogo.

Se H é indefinida, existem $v, w \in \mathbb{R}^n - \{0\}$ tais que $H v^2 > 0$ e $H w^2 < 0$. Então, para todo $t \neq 0$, temos que $H (tv)^2 = t^2 H v^2 > 0$ e $H (tw)^2 = t^2 H w^2 < 0$. Logo, por (*),

$$\frac{f(a + tv) - f(a)}{t^2} = H v^2 + \frac{r(tv)}{t^2} \quad \text{e} \quad \frac{f(a + tw) - f(a)}{t^2} = H w^2 + \frac{r(tw)}{t^2}.$$

Como $\lim_{t \rightarrow 0} \frac{r(tv)}{t^2} = \lim_{t \rightarrow 0} \frac{r(tw)}{t^2} = 0$, segue-se que

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t^2} = H v^2 > 0 \quad \text{e} \quad \lim_{t \rightarrow 0} \frac{f(a + tw) - f(a)}{t^2} = H w^2 < 0.$$

Logo existe $\delta > 0$ tal que $0 < |t| < \delta \implies f(a + tv) - f(a) > 0$ e $f(a + tw) - f(a) < 0$.

Portanto, a não é ponto de máximo local nem de mínimo local para f . ■

Exemplo 8.4. Seja $f : \mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ a função definida por $f(x, y) = \langle x, x \rangle - \langle y, y \rangle$, onde $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^n$. Então $\frac{\partial f}{\partial x_i} = 2x_i$ e $\frac{\partial f}{\partial y_j} = -2y_j$. Logo $\text{grad } f(x, y) = 2(x, -y)$ e, portanto, a origem é o único ponto crítico de f .

A matriz Hessiana de f em qualquer ponto de $\mathbb{R}^{m+n}$ é a matriz diagonal cujas m primeiras entradas na diagonal principal são iguais a 2 e as n últimas são iguais a -2 .

Então a matriz Hessiana é positiva se $n = 0$, negativa se $m = 0$, e indefinida se $mn \neq 0$. Assim, a origem é ponto de mínimo se $n = 0$ e de máximo se $m = 0$.

Para $mn \neq 0$, f não admite mínimo nem máximo na origem, que se chama um *ponto de sela*, devido à forma do gráfico da função $f(x, y) = x^2 - y^2$. $\square$

Observação 8.9. Como vimos na demonstração do teorema 8.5, se $\text{grad } f(a) = 0$ e $Hv^2 > 0$ para algum $v \in \mathbb{R}^n$, então existe $\delta > 0$ tal que $0 < |t| < \delta \implies f(a + tv) > f(a)$. Então se a é um ponto de máximo local de f , a forma Hessiana de f no ponto a é não-positiva, isto é, $Hv^2 \leq 0$ para todo $v \in \mathbb{R}^n$. De modo análogo, se a é um ponto de mínimo local de f , então a forma Hessiana de f no ponto a é não-negativa, ou seja, $Hv^2 \geq 0$ para todo $v \in \mathbb{R}^n$.

Mas a recíproca destas afirmações são falsas, ou seja, quando a forma hessiana de f num ponto crítico é ≤ 0 (ou ≥ 0) não se pode afirmar que a função tem um máximo (ou um mínimo) neste ponto.

Por exemplo, sejam as funções $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$f(x, y) = x^2 \quad \text{e} \quad g(x, y) = x^2 + y^3.$$

Então $\text{grad } f(x, y) = (2x, 0)$, $\text{grad } g(x, y) = (2x, 3y^2)$, e as hessianas de f e g no ponto crítico $(0, 0)$ coincidem e são não-negativas, pois $Hf(0, 0)v^2 = Hg(0, 0)v^2 = 2\alpha^2$ para todo $v = (\alpha, \beta) \in \mathbb{R}^2$. Mas a origem é um ponto de mínimo para f e não é um mínimo local para g .

9 O teorema da função implícita

Começamos observando o seguinte exemplo:

Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x, y) = x^2 + y^2$. Então $S^1 = f^{-1}(1) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.

A equação $x^2 + y^2 = 1$ não define y como função de x , nem x como função de y , globalmente. Mas, se tomarmos $U_1 = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$; $U_2 = \{(x, y) \in \mathbb{R}^2 \mid y < 0\}$; $U_3 = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ e $U_4 = \{(x, y) \in \mathbb{R}^2 \mid x < 0\}$, temos que:

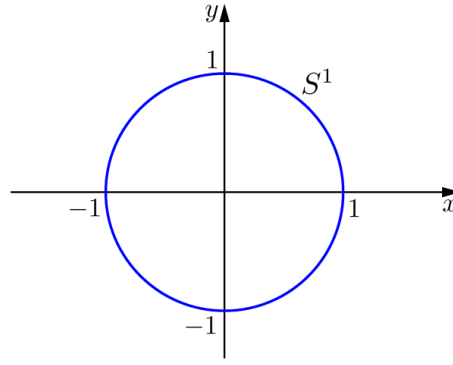


Fig. 8: O círculo unitário $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$

- $(x, y) \in S^1 \cap U_1 \iff y = \sqrt{1 - x^2} \text{ e } x \in (-1, 1)$;
- $(x, y) \in S^1 \cap U_2 \iff y = -\sqrt{1 - x^2} \text{ e } x \in (-1, 1)$;
- $(x, y) \in S^1 \cap U_3 \iff x = \sqrt{1 - y^2} \text{ e } y \in (-1, 1)$;
- $(x, y) \in S^1 \cap U_4 \iff x = -\sqrt{1 - y^2} \text{ e } y \in (-1, 1)$.

Como $S^1 = (U_1 \cap S^1) \cup (U_2 \cap S^1) \cup (U_3 \cap S^1) \cup (U_4 \cap S^1)$, temos que

$$S^1 = \text{Graf } \xi_1 \cup \text{Graf } \xi_2 \cup \text{Graf } \xi_3 \cup \text{Graf } \xi_4,$$

onde $\xi_i : (-1, 1) \rightarrow \mathbb{R}$, $i = 1, 2, 3, 4$, são as funções de classe C^∞ dadas por:

$$\xi_1(x) = \sqrt{1 - x^2}, \quad \xi_2(x) = -\sqrt{1 - x^2}, \quad \xi_3(y) = \sqrt{1 - y^2}, \quad \text{e } \xi_4(y) = -\sqrt{1 - y^2},$$

Logo todo ponto $(x_0, y_0) \in S^1$ pertence a um aberto V de $\mathbb{R}^2$ tal que $V \cap S^1$ é o gráfico de uma função de classe C^∞ definida num aberto de $\mathbb{R}$.

Definição 9.1. Dizemos que um conjunto $C \subset \mathbb{R}^2$ é uma *curva de classe C^k* ($0 \leq k \leq \infty$) quando C é localmente o gráfico de uma função de classe C^k . Ou seja, para todo ponto $p \in C$ existe um aberto $V \subset \mathbb{R}^2$ tal que $p \in V$ e $V \cap C$ é o gráfico de uma função ξ de classe C^k definida num aberto de $\mathbb{R}$.

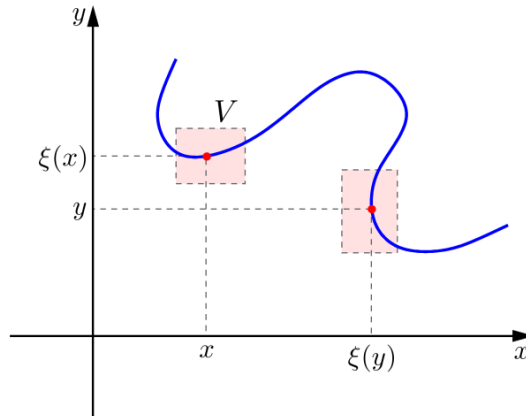


Fig. 9: Uma curva de classe C^k é, localmente, o gráfico de uma função de classe C^k

Exemplo 9.1. O círculo S^1 é uma curva de classe C^∞ . $\square$

Exemplo 9.2. O conjunto $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 0\}$ não é uma curva nem de classe C^0 , pois, para todo aberto V contendo a origem, $C \cap V$ não é o gráfico de uma função $y = \xi(x)$ nem $x = \xi(y)$, uma vez que $C \cap V$ contém sempre dois segmentos de reta de inclinação ± 1 que se cortam na origem. $\square$

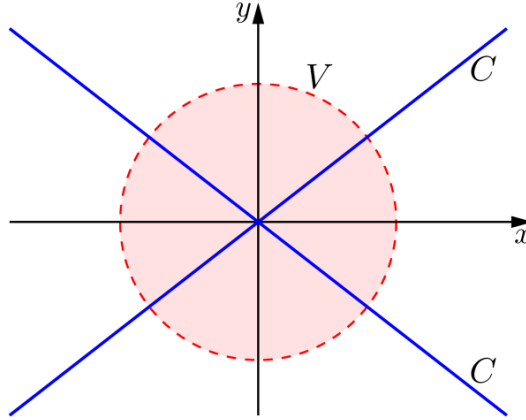


Fig. 10: O conjunto C não é uma curva nem de classe C^0 .

Exemplo 9.3. O conjunto $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1\}$ é uma curva desconexa de classe C^∞ , pois $C = (V_1 \cap C) \cup (V_2 \cap C)$, onde $V_1 = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ e $V_2 = \{(x, y) \in \mathbb{R}^2 \mid x < 0\}$ são abertos de $\mathbb{R}^2$ tais que:

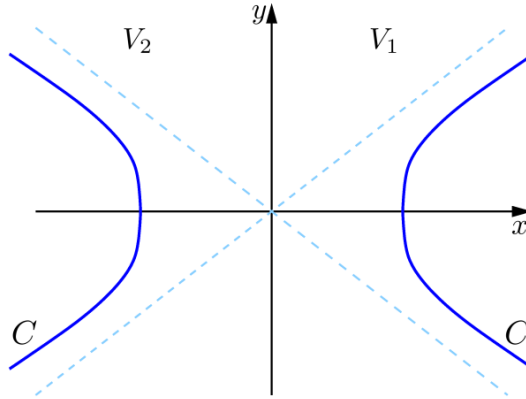


Fig. 11: O conjunto C é uma curva desconexa de classe C^∞ .

- $V_1 \cap C$ é o gráfico da função C^∞ $\xi_1 : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $\xi_1(y) = \sqrt{1 + y^2}$,
- e
- $V_2 \cap C$ é o gráfico da função C^∞ $\xi_2 : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $\xi_2(y) = -\sqrt{1 + y^2}$. $\square$

Analisaremos, agora, um exemplo de um subconjunto de $\mathbb{R}^{n+1}$ que é dado localmente como o gráfico de uma função definida num aberto de $\mathbb{R}^n$.

Exemplo 9.4. Seja $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ dada por $f(x) = \langle x, x \rangle$ e seja

$$f^{-1}(1) = S^n = \{x \in \mathbb{R}^{n+1} \mid \langle x, x \rangle = 1\}$$

a esfera unitária n -dimensional.

Indiquemos por $U \subset \mathbb{R}^n$ a bola aberta de raio 1 e centro na origem.

Para cada $i = 1, \dots, n+1$, sejam $V_i = \{x \in \mathbb{R}^{n+1} \mid x_i > 0\}$ e $W_i = \{x \in \mathbb{R}^{n+1} \mid x_i < 0\}$.

Escrevendo $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, temos:

$$\begin{aligned} x \in S^n \cap V_i &\iff \|x^*\| < 1 \quad \text{e} \quad x_i = \sqrt{1 - \langle x^*, x^* \rangle} \\ x \in S^n \cap W_i &\iff \|x^*\| < 1 \quad \text{e} \quad x_i = -\sqrt{1 - \langle x^*, x^* \rangle}. \end{aligned}$$

Logo, se $\xi : U \rightarrow \mathbb{R}$ é a função C^∞ dada por $\xi(u) = \sqrt{1 - \langle u, u \rangle}$, $S^n \cap V_i$ é o gráfico da função $x_i = \xi(x^*)$ e $S^n \cap W_i$ é o gráfico da função $x_i = -\xi(x^*)$, para cada $i = 1, \dots, n+1$.

Como $S^n = \left(\bigcup_{i=1}^{n+1} V_i \cap S^n \right) \cup \left(\bigcup_{i=1}^{n+1} W_i \cap S^n \right)$, todo ponto $p \in S^n$ pertence a um aberto Z de $\mathbb{R}^{n+1}$ tal que $Z \cap S^n$ é o gráfico de uma função de classe C^∞ definida num aberto de $\mathbb{R}^n$. $\square$

Definição 9.2. Um conjunto $M \subset \mathbb{R}^{n+1}$ chama-se uma *hipersuperfície* (ou *hiperfície*) de classe C^k , $0 \leq k \leq \infty$, de $\mathbb{R}^{n+1}$ quando M é localmente o gráfico de uma função de classe C^k de n variáveis. Ou seja, todo ponto $p \in M$ pertence a um aberto $V \subset \mathbb{R}^{n+1}$ tal que $V \cap M$ é o gráfico de uma função de classe C^k definida num aberto de $\mathbb{R}^n$ (existem um aberto $U \subset \mathbb{R}^n$, uma função $\xi : U \rightarrow \mathbb{R}$ de classe C^k e um inteiro $i \in \{1, \dots, n+1\}$ tais que $x_i = \xi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$ e $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) \in U$).

Quando $n = 1$, dizemos que $M \subset \mathbb{R}^2$ é uma *curva*, e quando $n = 2$, dizemos que $M \subset \mathbb{R}^3$ é uma *superfície*.

Observação 9.1. Podemos também considerar as *hipersuperfícies diferenciáveis* (caso intermediário entre C^0 e C^1) que são localmente gráficos de funções diferenciáveis.

Exemplo 9.5. S^n é uma hipersuperfície de classe C^∞ de $\mathbb{R}^{n+1}$. $\square$

Seja $M \subset \mathbb{R}^{n+1}$ e seja $p \in M$. Definimos $T_p M$ como sendo o conjunto de todos os vetores velocidade $\lambda'(0)$, onde $\lambda : (-\varepsilon, \varepsilon) \rightarrow M \subset \mathbb{R}^{n+1}$ é um caminho diferenciável em $t = 0$ e $\lambda(0) = p$.

Quando M é uma hipersuperfície diferenciável, o conjunto $T_p M$ chama-se o *espaço tangente a M no ponto p* .

Teorema 9.1. Se $M \subset \mathbb{R}^{n+1}$ é uma hipersuperfície diferenciável, então $T_p M$ é um subespaço vetorial de dimensão n do espaço euclidiano $\mathbb{R}^{n+1}$, para cada $p \in M$.

Prova.

Dado $p = (a_1, \dots, a_{n+1}) \in M$, existem abertos $V \subset \mathbb{R}^{n+1}$, $U \subset \mathbb{R}^n$, com $p \in V$, um inteiro $i \in \{1, \dots, n+1\}$ e uma função $\xi : U \rightarrow \mathbb{R}$ diferenciável tais que $x \in V \cap M \iff x_i = \xi(x^*)$, onde $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) \in U$.

Afirmção: $T_p M = \left\{ v = (\alpha_1, \dots, \alpha_{n+1}) \in \mathbb{R}^{n+1} \mid \alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \alpha_j \right\}$,

onde $p^* = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_{n+1})$.

De fato, seja $v \in T_p M$. Então existe um caminho diferenciável em $t = 0$, $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$, com $\lambda(0) = p$ e $\lambda'(0) = v$. Como V é aberto, $p \in V$ e λ é contínuo em $t = 0$, existe $0 < \varepsilon_0 \leq \varepsilon$ tal que $\lambda(t) \in M \cap V$ para todo $t \in (-\varepsilon_0, \varepsilon_0)$.

Logo $\lambda_i(t) = \xi(\lambda_1(t), \dots, \lambda_{i-1}(t), \lambda_{i+1}(t), \dots, \lambda_{n+1}(t))$ para todo $t \in (-\varepsilon_0, \varepsilon_0)$.

Pela Regra da Cadeia,

$$\lambda'_i(0) = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \lambda'_j(0),$$

ou seja, $\alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \alpha_j$.

Sejam agora $v = (\alpha_1, \dots, \alpha_{n+1}) \in \mathbb{R}^{n+1}$ tal que $\alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \alpha_j$ e $\varepsilon > 0$ tal que $p^* + tv^* \in U$ para todo $t \in (-\varepsilon, \varepsilon)$, onde $v^* = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{n+1})$.

Podemos, assim, definir o caminho $\lambda : (-\varepsilon, \varepsilon) \rightarrow M \cap V$ pondo $\lambda_j(t) = \alpha_j + t\alpha_j$, $j \neq i$, e

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

Logo λ é diferenciável em $t = 0$, $\lambda(0) = p$ e $\lambda'(0) = v$. Então $v \in T_p M$, provando, assim, a afirmação.

Assim, $T_p M$ é um subespaço vetorial de dimensão n de $\mathbb{R}^{n+1}$ gerado pelos vetores linearmente independentes

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

onde $c_j = \left(\frac{\partial \xi}{\partial x_j} \right)(p^*)$.

Outra maneira de interpretar a afirmação acima é dizer que ela caracteriza $T_p M$ como o núcleo do funcional linear não-nulo $\varphi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, dado por

$$\varphi(v) = \alpha_i - \sum_{j \neq i} c_j \alpha_j,$$

onde $v = (\alpha_1, \dots, \alpha_{n+1})$ e $c_j = \frac{\partial \xi}{\partial x_j}(p^*)$. Ou ainda, $T_p M$ é o gráfico do funcional linear

$d\xi(p^*) : \mathbb{R}^n \longrightarrow \mathbb{R}$, dado por:

$$v^* = (\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_{n+1}) \longmapsto d\xi(p^*)v^* = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \alpha_j.$$

■

Exemplo 9.6. Seja $S^n = \{x \in \mathbb{R}^{n+1} \mid \langle x, x \rangle = 1\}$. Já sabemos que S^n é uma hipersuperfície de classe C^∞ .

Afirmção: $T_p S^n = \{v \in \mathbb{R}^{n+1} \mid \langle v, p \rangle = 0\} = [p]^\perp$, para todo $p \in S^n$.

De fato, seja $\lambda : (-\varepsilon, \varepsilon) \longrightarrow S^n$ uma curva diferenciável em $t = 0$ com $\lambda(0) = p$ e $\lambda'(0) = v$.

Então, como $\langle \lambda(t), \lambda(t) \rangle = 1$ para todo $t \in (-\varepsilon, \varepsilon)$, temos que $2\langle \lambda'(0), \lambda(0) \rangle = 0$, ou seja, $\langle v, p \rangle = 0$. Logo $T_p S^n \subset [p]^\perp$ e, portanto, $T_p S^n = [p]^\perp$, pois $\dim T_p S^n = \dim [p]^\perp = n$. $\square$

Para hipersuperfícies $M \subset \mathbb{R}^{n+1}$ de classe C^0 , $T_p M$ pode não ser um espaço vetorial de dimensão n .

Exemplo 9.7. Seja $X = \{(x, y, z) \in \mathbb{R}^3 \mid z = \sqrt{x^2 + y^2}\}$ o cone de vértice na origem e eixo $-z$. Então, para $p = (0, 0, 0)$, $T_p M = \{(0, 0, 0)\}$.

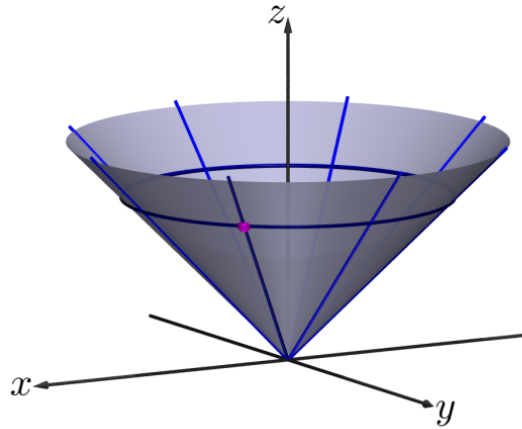


Fig. 12: Cone X de vértice na origem.

De fato, seja $\lambda : (-\varepsilon, \varepsilon) \longrightarrow X$ uma curva diferenciável em $t = 0$ com $\lambda(0) = (0, 0, 0)$ e $\lambda'(0) = (v_1, v_2, v_3)$. Então, se $\lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t))$, $\lambda_3(t) = \sqrt{(\lambda_1(t))^2 + (\lambda_2(t))^2}$, $v_1 = \lambda'_1(0) = \lim_{t \rightarrow 0} \frac{\lambda_1(t)}{t}$ e $v_2 = \lambda'_2(0) = \lim_{t \rightarrow 0} \frac{\lambda_2(t)}{t}$.

Logo,

$$v_3 = \lim_{t \rightarrow 0^+} \frac{1}{t} \sqrt{(\lambda_1(t))^2 + (\lambda_2(t))^2} = \lim_{t \rightarrow 0^+} \sqrt{\frac{(\lambda_1(t))^2 + (\lambda_2(t))^2}{t^2}} = \sqrt{v_1^2 + v_2^2},$$

e

$$v_3 = \lim_{t \rightarrow 0^-} \frac{1}{t} \sqrt{(\lambda_1(t))^2 + (\lambda_2(t))^2} = \lim_{t \rightarrow 0^-} -\sqrt{\frac{(\lambda_1(t))^2 + (\lambda_2(t))^2}{t^2}} = -\sqrt{v_1^2 + v_2^2}.$$

Portanto, $\sqrt{v_1^2 + v_2^2} = 0$, ou seja, $v_1 = v_2 = v_3 = 0$. $\square$

Exemplo 9.8. Seja Y a superfície de classe C^0 dada por $Y = \{(x, y, z) \in \mathbb{R}^3 \mid z = |x|\}$. Então, para $p = (0, 0, 0)$, $T_p Y = \{(0, \beta, 0) \mid \beta \in \mathbb{R}\}$ é um espaço vetorial de dimensão 1 ($\neq 2$) em $\mathbb{R}^3$.

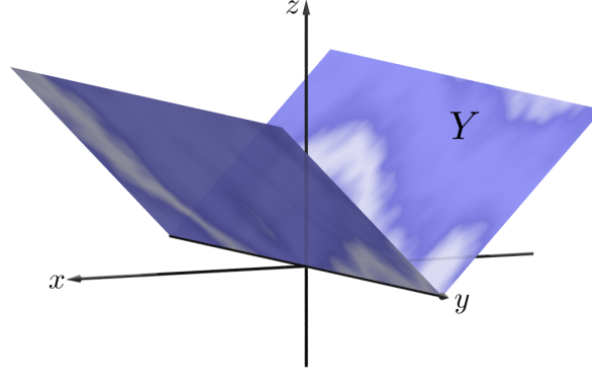


Fig. 13: Superfície Y .

De fato, seja $\lambda : (-\varepsilon, \varepsilon) \rightarrow Y$, $\lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t))$, uma curva diferenciável em $t = 0$ com $\lambda(0) = (0, 0, 0)$ e $\lambda'(0) = (v_1, v_2, v_3) = v$.

Então $\lambda_3(t) = |\lambda_1(t)|$ e $v_1 = \lambda'_1(0) = \lim_{t \rightarrow 0} \frac{\lambda_1(t)}{t}$.

Suponhamos que $v_1 > 0$. Então existe $0 < \varepsilon_0 < \varepsilon$ tal que $\lambda_1(t) > 0$ para $t \in (0, \varepsilon_0)$ e $\lambda_1(t) < 0$ para $t \in (-\varepsilon_0, 0)$. Assim,

$$v_3 = \lambda'_3(0) = \lim_{t \rightarrow 0^\pm} \frac{\lambda_3(t)}{t} = \lim_{t \rightarrow 0^\pm} \frac{|\lambda_1(t)|}{t} = \lim_{t \rightarrow 0^\pm} \pm \frac{\lambda_1(t)}{t} = \pm v_1.$$

Logo $v_1 = 0$, uma contradição. De modo análogo, podemos provar que v_1 não pode ser negativo.

Então $v_1 = 0$ e, portanto, $v_3 = 0$, ou seja, $v \in \{(0, \beta, 0) \in \mathbb{R}^3 \mid \beta \in \mathbb{R}\}$.

Reciprocamente, seja $v = (0, \beta, 0)$, $\beta \in \mathbb{R}$. Então a curva $\lambda : \mathbb{R} \rightarrow Y$, dada por $\lambda(t) = (0, \beta t, 0)$, é de classe C^∞ , $\lambda(0) = (0, 0, 0)$ e $\lambda'(0) = (0, \beta, 0)$. Logo $(0, \beta, 0) \in T_p Y$ para todo $\beta \in \mathbb{R}$.

Assim, $T_p Y = \{(0, \beta, 0) \in \mathbb{R}^3 \mid \beta \in \mathbb{R}\}$. $\square$

Definição 9.3. Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no aberto U . Dizemos que $c \in \mathbb{R}$ é um **valor regular de f** quando não existem pontos críticos de f no nível c , ou seja, $\text{grad } f(x) \neq 0$ para todo $x \in f^{-1}(c)$. Quando c é um valor regular de f , diz-se que o **nível c é regular**. Quando existem pontos críticos $x \in U$ tais que $f(x) = c$, dizemos que c é um **nível crítico de f** .

Observação 9.2. Se $f^{-1}(c) = \emptyset$, então c é um valor regular.

Exemplo 9.9. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função de classe C^∞ dada por $f(x, y) = x^2 + y^2$.

Como $\text{grad } f(x, y) = (2x, 2y)$ para todo $(x, y) \in \mathbb{R}^2$, temos que $\text{grad } f(x, y) = (0, 0)$ se, e só se, $(x, y) = (0, 0)$. Logo $f^{-1}(c)$ é um nível regular para todo $c \in \mathbb{R} - \{0\}$, pois $f(0, 0) = 0$. $\square$

Teorema 9.2. (Teorema Global da Função Implícita)

Sejam $f : U \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ uma função de classe C^k , $k \geq 1$, definida no aberto U , e $c \in f(U)$ um valor regular de f . Então $M = f^{-1}(c)$ é uma hipersuperfície de classe C^k e

$$T_p M = \ker df(p) = \{v \in \mathbb{R}^{n+1} \mid df(p)(v) = 0\} = \{v \in \mathbb{R}^{n+1} \mid \langle v, \text{grad } f(p) \rangle = 0\},$$

para todo $p \in M$.

Exemplo 9.10. Seja $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ a função de classe C^∞ dada por $f(x) = \langle x, x \rangle$. Como $\text{grad } f(x) = 2x$, pois $\frac{\partial f}{\partial x_i}(x) = 2x_i$, para todo $i = 1, \dots, n+1$, $\text{grad } f(x) = 0$ se, e somente se, $x = 0$, ou seja, se, e só se, $f(x) = 0$. Assim, $f^{-1}(c)$ é um nível regular para todo $c \in \mathbb{R} - \{0\}$, sendo $f^{-1}(c) = \emptyset$, se $c < 0$, e $f^{-1}(c) = S_{\sqrt{c}}^n(0)$, se $c > 0$. Logo, pelo teorema acima, $S_{\sqrt{c}}^n$ é uma hipersuperfície de classe C^∞ e

$$T_p S_{\sqrt{c}}^n(0) = \{v \in \mathbb{R}^{n+1} \mid \langle v, 2p \rangle = 0\} = [p]^\perp,$$

para todo $p \in S_{\sqrt{c}}^n(0)$. $\square$

Exemplo 9.11. Seja $\det : \mathbb{R}^{n^2} = \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ a função de classe C^∞ que associa a cada matriz $n \times n$, $X = (x_{ij})$, o seu determinante.

Como a expansão de $\det X$ pelas entradas da i -ésima linha é

$$\det X = \sum_{j=1}^n (-1)^{i+j} x_{ij} X_{[i,j]},$$

onde $X_{[i,j]}$ é o determinante da matriz $(n-1) \times (n-1)$ que se obtém da matriz X omitindo a i -ésima linha e a j -ésima coluna, temos que

$$\frac{\partial \det}{\partial x_{ij}}(X) = (-1)^{i+j} X_{[i,j]},$$

para todo $X \in \mathbb{R}^{n^2}$ e todos $i, j = 1, \dots, n$.

Em particular, no ponto $X = I$, temos $\frac{\partial \det}{\partial x_{ij}}(I) = \delta_{ij}$, $i, j = 1, \dots, n$, ou seja, o gradiente da função determinante no ponto I é a matriz identidade.

Seja $U = \{X \in \mathbb{R}^{n^2} \mid \det X \neq 0\}$ o conjunto aberto formado pelas matrizes $n \times n$ invertíveis. Então a restrição $\det : U \rightarrow \mathbb{R}$ é uma função C^∞ sem pontos críticos. De fato, se $\frac{\partial \det}{\partial x_{ij}}(X) = 0$ para todo $i, j = 1, \dots, n$, então

$$\det X = \sum_{j=1}^n (-1)^{i+j} x_{ij} X_{[i,j]} = 0,$$

e, portanto, $X \notin U$. Logo todo $c \in \mathbb{R}$ é um valor regular para a função $\det : U \rightarrow \mathbb{R}$.

Em particular,

$$M = \det^{-1}(1) = (\text{conjunto das matrizes } n \times n \text{ que têm determinante igual a } 1)$$

é uma hipersuperfície de classe C^∞ em $\mathbb{R}^{n^2}$. M é um grupo relativamente à multiplicação de matrizes, conhecido como o *grupo unimodular* de $\mathbb{R}^n$.

O espaço tangente $T_I(M)$ de M no ponto I é o subespaço de dimensão $n^2 - 1$ de $\mathbb{R}^{n^2}$ formado pelas matrizes $n \times n$ de traço nulo, pois $\text{grad}(\det(I)) = I$ e, portanto,

$$T_I M = \left\{ X \in \mathbb{R}^{n^2} \mid \langle X, I \rangle = \sum_{i,j=1}^n x_{ij} \delta_{ij} = \sum_{i=1}^n x_{ii} = \text{traço } X = 0 \right\}. \quad \square$$

Observação 9.3. Toda hipersuperfície $M \subset \mathbb{R}^{n+1}$, sendo localmente o gráfico de uma função $x_i = \xi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) = \xi(x^*)$, de n variáveis, é também localmente a imagem inversa $f^{-1}(0)$ do valor regular 0 da função $f(x) = x_i - \xi(x^*)$, definida no aberto $V \subset \mathbb{R}^{n+1}$ tal que $V \cap M$ é o gráfico de ξ , pois $\frac{\partial f}{\partial x_i}(x) = 1$ para todo $x \in V$ e $f^{-1}(0) = \{x \in V \mid x_i = \xi(x^*)\} = V \cap M$.

Para isso, estamos supondo $V = \prod_{j=1}^{n+1} I_j$, onde cada I_j é um intervalo aberto, e $U = \prod_{j \neq i} I_j$ é o domínio da função ξ .

Mas não é verdade que toda hipersuperfície $M \subset \mathbb{R}^{n+1}$ seja globalmente a imagem inversa de um valor regular, pois se $M = f^{-1}(c)$, a aplicação $\varphi = \text{grad } f : M \rightarrow \mathbb{R}^{n+1}$ fornece *um campo contínuo de vetores normais não-nulos ao longo de M* , uma vez que $\varphi(p) = \text{grad } f(p) \perp v$ para todo $v \in T_p M$. As hipersuperfícies que admitem um campo contínuo de vetores normais não-nulos $\varphi : M \rightarrow \mathbb{R}^{n+1}$ chamam-se *hipersuperfícies orientáveis*. Mas nem toda hipersuperfície em $\mathbb{R}^{n+1}$ é orientável, como a faixa de Möbius em $\mathbb{R}^3$ (ver §14, Cap. V).

Portanto, existem hipersuperfícies em $\mathbb{R}^{n+1}$ que não são globalmente a imagem inversa de um valor regular.

Lema 9.1. *Sejam $X \subset \mathbb{R}^m$, $K \subset \mathbb{R}^k$ compacto, $f : X \times K \rightarrow \mathbb{R}^p$ contínua e $c \in \mathbb{R}^p$. Se $f^{-1}(c)$ é o gráfico de uma aplicação $\xi : X \rightarrow K$ (isto é, para todo $x \in X$ existe um único $y = \xi(x) \in K$ tal que $f(x, \xi(x)) = c$) então ξ é contínua.*

Prova.

Dado $x_0 \in X$, seja $y_0 = \xi(x_0) \in K$ e seja $\{x_n\}$ uma sequência de pontos de X tal que $x_n \rightarrow x_0$.

Queremos provar que $\lim_{n \rightarrow \infty} \xi(x_n) = y_0$.

Como a sequência $\{\xi(x_n)\}$ é limitada, pois $\xi(x_n) \in K$ para todo $n \in \mathbb{N}$, basta mostrar que toda subsequência $\{\xi(x_{n'})\}_{n' \in \mathbb{N}'}$ convergente em $\mathbb{R}^k$ tem limite y_0 .

Seja $N' \subset \mathbb{N}$ tal que $\lim_{n \in N'} \xi(x_n) = y$. Então $y \in K$, pois K é compacto. Além disso, como f é contínua e $f(x_n, \xi(x_n)) = c$ para todo $n \in \mathbb{N}$, temos $c = \lim_{n \in N'} f(x_n, \xi(x_n)) = f(x_0, y)$.

Logo $f(x_0, y) = f(x_0, y_0)$ e, portanto, pela unicidade, $y = y_0$. ■

Observação 9.4. Supondo K apenas limitado, o lema acima nem sempre é válido. Por exemplo, seja $f : \mathbb{R} \times [0, 1) \rightarrow \mathbb{R}$ a função contínua definida por $f(x, y) = (x^2 + y^2)(ye^{|x|} - 1)$. Então, para cada $x \in \mathbb{R}$, existe um único $y \in [0, 1)$ tal que $f(x, y) = 0$, pois se $x = 0$, então $y = 0$, uma vez que $1 \notin [0, 1)$, e se $x \neq 0$, $y = e^{-|x|} \in [0, 1)$.

Logo $f^{-1}(0)$ é o gráfico da função $\xi : \mathbb{R} \rightarrow [0, 1)$ dada por $\xi(0) = 0$ e $\xi(x) = e^{-|x|}$, se $x \in \mathbb{R} - \{0\}$, que não é contínua em $x = 0$.

No teorema abaixo, representaremos os pontos de $\mathbb{R}^{n+1}$ por pares (x, y) , onde $x \in \mathbb{R}^n$ e $y \in \mathbb{R}$.

Teorema 9.3. (Teorema da Função Implícita)

Seja $f : U \rightarrow \mathbb{R}$ uma função de classe C^k , $k \geq 1$, definida num aberto $U \subset \mathbb{R}^{n+1}$. Seja $p = (x_0, y_0) \in U$ tal que $f(p) = c$ e $\frac{\partial f}{\partial y}(p) \neq 0$.

Então existem uma bola aberta $B = B_\delta(x_0) \subset \mathbb{R}^n$ e um intervalo aberto $J = (y_0 - \varepsilon, y_0 + \varepsilon)$ tais que $B \times J \subset U$ e $f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi : B \rightarrow J$ de classe C^k (isto é, para todo $x \in B$ existe um único $y = \xi(x) \in J$ tal que $f(x, y) = c$).

Para cada $x \in B$, tem-se:

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

A função $y = \xi(x)$ diz-se **definida implicitamente no aberto $U \times J$ pela equação $f(x, y) = c$** .

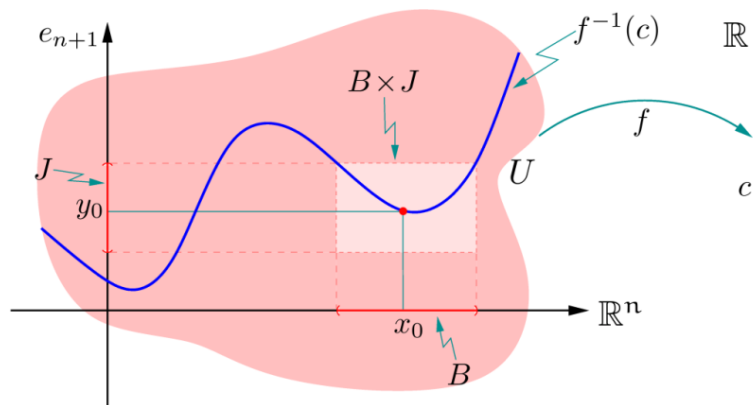


Fig. 14: Função $y = \xi(x)$ definida implicitamente no aberto $U \times J$.

Prova.

Suponhamos que $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $\frac{\partial f}{\partial y} : U \rightarrow \mathbb{R}$ é contínua, existem $\delta' > 0$ e $\varepsilon > 0$, tais que $B' \times \bar{J} \subset U$ e $\frac{\partial f}{\partial y}(x, y) > 0$ para todo $(x, y) \in B' \times \bar{J}$, onde $B' = B_{\delta'}(x_0)$ e $J = (y_0 - \varepsilon, y_0 + \varepsilon)$. Então, para todo $x \in B'$, a função $y \mapsto f(x, y)$ é estritamente crescente no intervalo $\bar{J} = [y_0 - \varepsilon, y_0 + \varepsilon]$. Como $f(x_0, y_0) = c$, temos que $f(x_0, y_0 - \varepsilon) < c$ e $f(x_0, y_0 + \varepsilon) > c$.

Pela continuidade de f , existe $0 < \delta < \delta'$ tal que $f(x, y_0 - \varepsilon) < c$ e $f(x, y_0 + \varepsilon) > c$ para todo $x \in B = B_\delta(x_0)$. Então, pelo Teorema do Valor Intermediário, existe, para cada $x \in B$, um único $y = \xi(x) \in \bar{J}$ tal que $f(x, y) = c$. Logo $y = \xi(x) \in J$ e $f^{-1}(c) \cap (B \times \bar{J}) = f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi : B \rightarrow J$ a qual, pelo lema anterior, é contínua.

Mostraremos agora que, em todo ponto $x \in B$, existem as derivadas parciais de ξ .

Seja $x \in B$ e tome $k = k(t) = \xi(x + te_i) - \xi(x)$. Então,

$$\xi(x + te_i) = \xi(x) + k \quad \text{e} \quad f(x + te_i, \xi(x) + k) = f(x, \xi(x)) = c,$$

para todo $t \in (-\delta_0, \delta_0)$, onde δ_0 foi escolhido de modo que $x + te_i \in B$ para todo $t \in (-\delta_0, \delta_0)$.

Pelo Teorema do Valor Médio, para todo $t \in (-\delta_0, \delta_0)$, existe $\theta = \theta(t) \in (0, 1)$ tal que:

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

Logo,

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

Pela continuidade de ξ , $\lim_{t \rightarrow 0} k(t) = 0$. Então, pela continuidade das derivadas parciais de f , a derivada parcial $\frac{\partial \xi}{\partial x_i}(x)$ existe e é igual a

$$\frac{\partial \xi}{\partial x_i}(x) = -\frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))} \quad (I)$$

para todo $i = 1, \dots, n$.

Como f é de classe C^1 e ξ é contínua, temos, por (I), que $\frac{\partial \xi}{\partial x_i}$ é contínua para todo $i = 1, \dots, n$, ou seja, ξ é de classe C^1 .

Suponhamos, por indução, que se f é de classe C^{k-1} , então ξ é de classe C^{k-1} , $k - 1 \geq 1$.

Seja $f \in C^k$. Então ξ é de classe C^{k-1} e as derivadas parciais de f são de classe C^{k-1} .

Assim, por (I), $\frac{\partial \xi}{\partial x_i}$ é de classe C^{k-1} para todo $i = 1, \dots, n$, ou seja, ξ é de classe C^k . ■

Observação 9.5. No teorema da função implícita, não há nada especial a respeito da última variável. Ou seja, vale o seguinte resultado:

Seja $f : U \rightarrow \mathbb{R}$ uma função de classe C^k definida no aberto $U \subset \mathbb{R}^{n+1}$. Se um ponto $p = (x_1^0, \dots, x_{n+1}^0) \in U$ é tal que $f(p) = c$ e $\frac{\partial f}{\partial x_i}(p) \neq 0$ para algum $i = 1, \dots, n+1$, então existe $\varepsilon > 0$

tal que $V = \prod_{k=1}^{n+1} (x_k^0 - \varepsilon, x_k^0 + \varepsilon) \subset U$ e uma função $\xi : B = \prod_{\substack{k=1 \\ k \neq i}}^{n+1} (x_k^0 - \varepsilon, x_k^0 + \varepsilon) \rightarrow (x_i^0 - \varepsilon, x_i^0 + \varepsilon)$

de classe C^k cujo gráfico é $f^{-1}(c) \cap V$, ou seja, o conjunto $f^{-1}(c) \cap V$ é dado por:

$$\{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) \in B \text{ e } \xi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) = x_i\}.$$

Além disso,

$$\frac{\partial \xi}{\partial x_j}(x^*) = - \frac{\frac{\partial f}{\partial x_j}(x_1, \dots, x_{i-1}, \xi(x^*), x_{i+1}, \dots, x_{n+1})}{\frac{\partial f}{\partial x_i}(x_1, \dots, x_{i-1}, \xi(x^*), x_{i+1}, \dots, x_{n+1})},$$

para todo $x \in B$ e todo $j = 1, \dots, n+1$, $j \neq i$, onde $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$.

Corolário 9.1. Seja $f : U \rightarrow \mathbb{R}$ uma função de classe C^k , $k \geq 1$, no aberto $U \subset \mathbb{R}^{n+1}$. Se $\xi : W \rightarrow \mathbb{R}$ é contínua no aberto $W \subset \mathbb{R}^n$ com $(x, \xi(x)) \in U$, $\frac{\partial f}{\partial y}(x, \xi(x)) \neq 0$ e $f(x, \xi(x)) = c$ para todo $x \in W$, então ξ é de classe C^k .

Observação 9.6. No corolário acima, não basta supor que c é um valor regular de f . Por exemplo, seja a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^∞ , dada por $f(x, y) = x - y^3$. Então, como $\text{grad } f(x, y) = (1, -3y^2)$, todo $c \in \mathbb{R}$ é valor regular de f , mas a função contínua $\xi : \mathbb{R} \rightarrow \mathbb{R}$, dada por $\xi(x) = \sqrt[3]{x}$, satisfaz $f(x, \xi(x)) = 0$ para todo $x \in \mathbb{R}$ e não é diferenciável na origem.

Observe que $\frac{\partial f}{\partial y}(x, 0) = 0$ para todo $x \in \mathbb{R}$.

Prova.

(do Teorema Global da Função Implícita)

Seja $p \in f^{-1}(c)$. Como $\text{grad } f(p) \neq 0$, existe $i \in \{1, \dots, n+1\}$ tal que $\frac{\partial f}{\partial x_i}(p) \neq 0$. Logo, pelo teorema da função implícita, existe um aberto $V \subset \mathbb{R}^{n+1}$ tal que $p \in V$ e $V \cap f^{-1}(c)$ é o gráfico de uma função de classe C^k definida num aberto de $\mathbb{R}^n$. Então $M = f^{-1}(c)$ é uma hipersuperfície de classe C^k .

Seja $v \in T_p M$. Então existe uma curva $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$ diferenciável em $t = 0$ tal que $\lambda(0) = p$ e $\lambda'(0) = v$. Logo $df(p)v = (f \circ \lambda)'(0) = 0$, pois $f(\lambda(t)) = c$ para todo $t \in (-\varepsilon, \varepsilon)$. Assim, $\langle \text{grad } f(p), v \rangle = 0$ para todo $v \in T_p M$, ou seja, $T_p M \subset [\text{grad } f(p)]^\perp$ e, portanto, $T_p M = [\text{grad } f(p)]^\perp$, pois $\dim T_p M = \dim [\text{grad } f(p)]^\perp = n$. ■

10 Multiplicador de Lagrange

Seja $M \subset \mathbb{R}^{n+1}$ uma hipersuperfície de classe C^k , $k \geq 1$, contida num aberto $U \subset \mathbb{R}^{n+1}$, e $f : U \rightarrow \mathbb{R}$ uma função de classe C^k .

Os pontos críticos de $f : U \rightarrow \mathbb{R}$ são, como já definimos anteriormente, os pontos $x \in U$ tais que $\text{grad } f(x) = 0$, ou seja, $\frac{\partial f}{\partial v}(x) = 0$ para todo $v \in \mathbb{R}^{n+1}$. Isto equivale a dizer que $(f \circ \lambda)'(0) = 0$ para todo caminho $\lambda : (-\varepsilon, \varepsilon) \rightarrow U$ diferenciável em $t = 0$ tal que $\lambda(0) = x$.

Por analogia, daremos a seguinte definição:

Definição 10.1. Dizemos que $p \in M$ é um ponto crítico de $f|_M$ se $(f \circ \lambda)'(0) = 0$ para todo caminho $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$ diferenciável em $t = 0$ com $\lambda(0) = p$. Isto significa que $\frac{\partial f}{\partial v}(p) = 0$ para todo $v \in T_p M$, ou seja, $p \in M$ é um ponto crítico de $f|_M$ se, e só se, $\langle \text{grad } f(p), v \rangle = 0$ para todo $v \in T_p M$, ou ainda, se, e somente se, o vetor $\text{grad } f(p)$ é normal à hipersuperfície M no ponto p .

Observação 10.1. Se $p \in M$ é um ponto de máximo ou de mínimo local de $f|_M$, então p é um ponto crítico de $f|_M$, pois para toda curva $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$ diferenciável em $t = 0$ com $\lambda(0) = p$, 0 é ponto de máximo ou de mínimo local da função real $f \circ \lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ e, portanto, $df(p)v = (f \circ \lambda)'(0) = 0$.

Observação 10.2. Todo ponto crítico de f em U que pertence a M é um ponto crítico de $f|_M$, pois, neste caso, $\text{grad } f(p) = 0$ e, portanto, $\langle \text{grad } f(p), v \rangle = 0$ para todo $v \in \mathbb{R}^{n+1}$.

Mas pode existir um ponto crítico de $f|_M$ que não é ponto crítico de f em U , isto é, no qual $\text{grad } f$ não se anula.

Exemplo 10.1. Sejam $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função de classe C^∞ dada por $f(x, y) = y$, e $M = S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Então f não possui ponto crítico, pois $\text{grad } f(x, y) = (0, 1) \neq (0, 0)$ para todo $(x, y) \in \mathbb{R}^2$. Mas $(0, -1)$ e $(0, 1)$ são pontos críticos de $f|_M$, pois $(0, -1)$ é o ponto de mínimo e $(0, 1)$ é o ponto de máximo de $f|_M$.

Em geral, se a hipersuperfície $M \subset \mathbb{R}^{n+1}$ é compacta, então $f|_M$ admite pelo menos dois pontos críticos: os pontos onde $f|_M$ assume seus valores máximo e mínimo. $\square$

Teorema 10.1. (do Multiplicador de Lagrange)

Sejam $\varphi : U \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ uma função de classe C^k , $M = \varphi^{-1}(c)$, onde $c \in \varphi(U)$ é um valor regular de φ , e $f : U \rightarrow \mathbb{R}$ uma função de classe C^k . Um ponto $p \in M$ é ponto crítico de $f|_M$ se, e só se, existe um número real $\lambda \in \mathbb{R}$ tal que $\text{grad } f(p) = \lambda \text{ grad } \varphi(p)$.

Prova.

Para todo ponto $p \in M$, temos $T_p M = [\text{grad } \varphi(p)]^\perp$, pois M é uma hipersuperfície de nível de φ . Além disso, p é ponto crítico de $f|_M$ se, e só se, $\text{grad } f(p) \perp T_p M$.

Como $T_p M \subset \mathbb{R}^{n+1}$ é um subespaço vetorial de dimensão n , temos que $p \in M$ é ponto crítico de $f|_M$ se, e só se, $\text{grad } f(p)$ é um múltiplo de $\text{grad } \varphi(p)$. ■

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

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

nas $n + 2$ incógnitas $\lambda, x_1, \dots, x_{n+1}$, onde $p = (x_1, \dots, x_{n+1})$. O número λ chama-se o *multiplicador de Lagrange*.

Observação 10.3. A condição $\text{grad } f(p) = \lambda \text{grad } \varphi(p)$ significa que a hipersuperfície M é tangente à hipersuperfície de nível de f que passa pelo ponto crítico p da função $f|_M$. No caso em que se podem esboçar as superfícies de nível da função f , esta observação auxilia a localizar os pontos críticos (ver exemplo abaixo).

Observação 10.4. Quando a hipersuperfície M não é dada como imagem inversa $\varphi^{-1}(c)$ de um valor regular, os pontos críticos de $f|_M$ são simplesmente os pontos $p \in M$ nos quais $\text{grad } f(p)$ é normal a M , ou seja, $\text{grad } f(p) \perp v$ para todo $v \in T_p M$.

Exemplo 10.2. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função de classe C^∞ dada por $f(x, y) = ax + by$, com $a^2 + b^2 \neq 0$, e seja $S^1 = \varphi^{-1}(1)$, onde $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ é dada por $\varphi(x, y) = x^2 + y^2$. Como 1 é valor regular de φ , os pontos críticos de $f|_{S^1}$ são os pontos $(x, y) \in S^1$ onde $\text{grad } f(x, y) = (a, b)$ e $\text{grad } \varphi(x, y) = (2x, 2y)$ são múltiplos. Então

$(a, b) = \lambda(x, y)$ e $x^2 + y^2 = 1$. Isto nos dá

$$x = \frac{a}{\sqrt{a^2 + b^2}} \text{ e } y = \frac{b}{\sqrt{a^2 + b^2}},$$

ou

$$x = -\frac{a}{\sqrt{a^2 + b^2}} \text{ e } y = -\frac{b}{\sqrt{a^2 + b^2}}.$$

Nestes pontos, $f|_{S^1}$ assume, respectivamente, seu valor máximo igual a $\sqrt{a^2 + b^2}$, e seu valor mínimo igual a $-\sqrt{a^2 + b^2}$, pois

$$|f(x, y)| \leq \sqrt{a^2 + b^2} \text{ para todo } (x, y) \in S^1. \quad \square$$

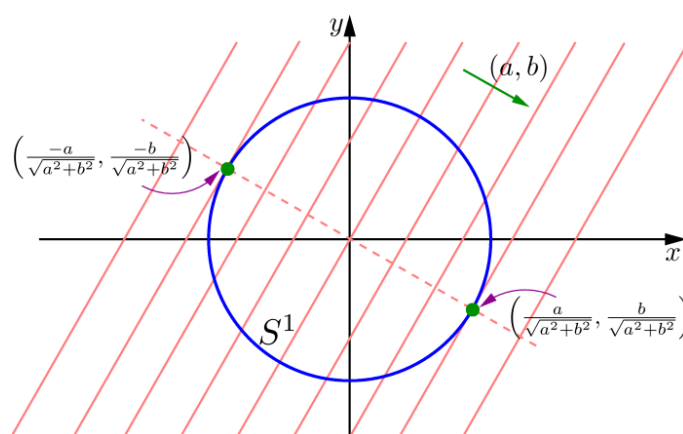


Fig. 15: Pontos críticos de $f|_{S^1}$.

Exemplo 10.3. Dados uma hipersuperfície $M \subset \mathbb{R}^{n+1}$ e um ponto $b \in \mathbb{R}^{n+1}$ tal que $b \notin M$, determinar o ponto $p \in M$ mais próximo a b . No caso em que M é fechada, um tal ponto sempre existe.

Consideremos a função $f : \mathbb{R}^{n+1} - \{b\} \rightarrow \mathbb{R}$ de classe C^∞ dada por $f(x) = \|x - b\|$. Os pontos onde $f|_M$ assume seu valor mínimo, caso existam, estão entre os pontos críticos de $f|_M$, isto é, entre os pontos $x \in M$ onde $\text{grad } f(x)$ é normal a M . Como $\text{grad } f(x) = \frac{x - b}{\|x - b\|}$, pois $\frac{\partial f}{\partial x_i}(x) = \frac{x_i - b_i}{\|x - b\|}$, para todo $i = 1, \dots, n$, os pontos críticos de $f|_M$, entre os quais se encontram os pontos de M situados a uma distância mínima do ponto b , são os pontos $x \in M$ tais que $x - b$ é normal a M . $\square$

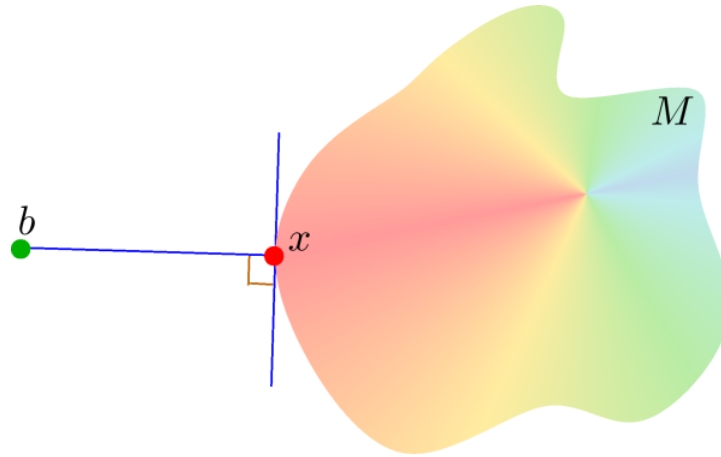


Fig. 16: $x - b$ é normal a M .

Exemplo 10.4. Seja $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma transformação linear *autoadjunta*, isto é, $\langle Ax, y \rangle = \langle x, Ay \rangle$ para quaisquer $x, y \in \mathbb{R}^n$. Isto equivale a dizer que a matriz (a_{ij}) de A com respeito à base canônica é simétrica, pois $a_{ij} = \langle Ae_j, e_i \rangle = \langle Ae_i, e_j \rangle = a_{ji}$.

Um número real λ é um *autovalor* de A quando existe um vetor $y \in \mathbb{R}^n - \{0\}$ tal que $Ay = \lambda y$. E os *autovetores associados ao autovalor λ* são os vetores $x \in \mathbb{R}^n$ tais que $Ax = \lambda x$.

Em geral, uma transformação linear $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ não precisa ter autovalores reais, como a rotação de ângulo $\theta \in (0, \pi)$ no plano.

Afirmção: Se $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma transformação linear autoadjunta, então existe uma base ortonormal de $\mathbb{R}^n$ formada por autovetores de A .

De fato, seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a forma quadrática dada por $f(x) = \langle Ax, x \rangle$ ou, em termos de coordenadas, $f(x) = \sum_{i,j=1}^n a_{ij}x_i x_j$.

Para determinarmos uma base ortonormal de autovetores de A estudaremos os pontos críticos

de f na esfera unitária $S^{n-1} \subset \mathbb{R}^n$. Como $S^{n-1} = \varphi^{-1}(1)$, onde 1 é valor regular da função $\varphi(x) = \langle x, x \rangle$, temos que $x \in S^{n-1}$ é um ponto crítico de $f|_{S^{n-1}}$ se, e só se, os vetores $\text{grad } f(x)$ e $\text{grad } \varphi(x) = 2x$ são múltiplos. Sendo $\frac{\partial f}{\partial x_i}(x) = 2 \sum_{j=1}^n a_{ij}x_j$, temos que $\text{grad } f(x) = 2Ax$. Logo os pontos críticos de $f|_{S^{n-1}}$ são os pontos $u \in S^{n-1}$ tais que $Au = \lambda u$ e, num tal ponto, temos $f(u) = \langle \lambda u, u \rangle = \lambda$, pois $\langle u, u \rangle = 1$.

Provamos, assim, que dada a forma quadrática $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \langle Ax, x \rangle$, onde $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é autoadjunta, um ponto $u \in S^{n-1}$ é um ponto crítico de $f|_{S^{n-1}}$ se, e só se, $Au = \lambda u$, onde $\lambda = f(u)$. Ou seja, $\lambda = f(u)$ é um autovalor de A e u é um autovetor de norma 1 associado ao autovalor λ .

Em particular, se λ_1 é o valor máximo de f no compacto S^{n-1} atingido no ponto $u_1 \in S^{n-1}$, então λ_1 é o maior autovalor de A e $Au_1 = \lambda_1 u_1$.

Seja $E = \{x \in \mathbb{R}^n \mid \langle x, u_1 \rangle = 0\}$ o complemento ortogonal do vetor u_1 . Se $x \in E$, então $\langle Ax, u_1 \rangle = \langle x, Au_1 \rangle = \lambda_1 \langle x, u_1 \rangle = 0$. Logo $A(E) \subset E$ e, portanto, por restrição, obtemos uma transformação linear autoadjunta $A : E \rightarrow E$.

Seja λ_2 o valor máximo da forma quadrática f entre os vetores unitários pertencentes a E , e seja $u_2 \in E$ tal que $|u_2| = 1$ e $f(u_2) = \lambda_2$. Então λ_2 é um autovalor de A e $Au_2 = \lambda_2 u_2$.

Prosseguindo desta maneira, obtemos uma base ortonormal de $\mathbb{R}^n$, $\{u_1, u_2, \dots, u_n\}$, formada por autovetores de A . $\square$

Exemplo 10.5. A *média geométrica* de n números reais positivos $x_1, \dots, x_n$ é menor do que ou igual à *média aritmética* destes números, isto é,

$$\sqrt[n]{x_1 \cdot \dots \cdot x_n} \leq \frac{x_1 + \dots + x_n}{n},$$

e a igualdade vale se, e só se, $x_1 = \dots = x_n$.

De fato, sejam $x_1, \dots, x_n$ n números reais positivos, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a função de classe C^∞ dada por $f(y_1, \dots, y_n) = y_1 \cdot \dots \cdot y_n$ e $c = x_1 + \dots + x_n$.

Vamos determinar o valor máximo de f na hipersuperfície

$$M_c = \{(y_1, \dots, y_n) \in \mathbb{R}^n \mid y_1 + \dots + y_n = c, y_1 > 0, \dots, y_n > 0\}.$$

Consideremos o aberto $U = \{(y_1, \dots, y_n) \in \mathbb{R}^n \mid y_1 > 0, \dots, y_n > 0\}$ e a função $\varphi : U \rightarrow \mathbb{R}$ de classe C^∞ dada por $\varphi(y_1, \dots, y_n) = y_1 + \dots + y_n$.

Então $\varphi^{-1}(c) = M_c$ é uma hipersuperfície de classe C^∞ de $\mathbb{R}^n$, pois $\text{grad } \varphi(y) = (1, 1, \dots, 1) \neq (0, 0, \dots, 0)$ para todo $y \in U$.

Como $\overline{M}_c$ é compacto, pois $\overline{M}_c \subset [0, c] \times \dots \times [0, c]$, existe $z \in \overline{M}_c$ tal que $f(z)$ é o valor máximo de $f|_{\overline{M}_c}$. Então $z \in M_c$, pois $f(y) = 0$ para todo $y \in \overline{M}_c - M_c$ e $f(y) > 0$ para todo $y \in M_c$.

Sendo $\frac{\partial f}{\partial y_i}(y) = \prod_{\substack{j=1 \\ j \neq i}}^n y_j$, para todo $i = 1, \dots, n$, temos, pelo método do multiplicador de Lagrange, que $\text{grad } f(z) = \lambda \text{ grad } \varphi(z) = (\lambda, \dots, \lambda)$. Então $z_1 + \dots + z_n = c$, $z_i > 0$ e $\prod_{j \neq i} z_j = \lambda$, para todo $i = 1, \dots, n$.

Afirmção: Se $z_1, \dots, z_n \in \mathbb{R} - \{0\}$ e $\prod_{\substack{j=1 \\ j \neq i}}^n z_j = \lambda$ para todo $i = 1, \dots, n$, então $z_1 = \dots = z_n$.

Vamos provar esta afirmação por indução sobre n .

Se $n = 2$, é claro que $z_1 = z_2$.

Suponhamos o resultado válido para $n - 1$, $n - 1 \geq 2$. Sejam $z_1, \dots, z_n$ n números reais não-nulos tais que $\prod_{\substack{j=1 \\ j \neq i}}^n z_j = \lambda$ para todo $i = 1, \dots, n$. Como, para todos $i, i' \in \{1, \dots, n - 1\}$,

$i \neq i'$, $\prod_{\substack{j=1 \\ j \neq i}}^n z_j = \prod_{\substack{j=1 \\ j \neq i'}}^n z_j$, e $z_n \neq 0$, temos $\prod_{\substack{j=1 \\ j \neq i}}^{n-1} z_j = \prod_{\substack{j=1 \\ j \neq i'}}^{n-1} z_j$. Logo, pela hipótese de indução, $z_1 = \dots = z_{n-1}$. Além disso, $z_1 = z_n$, pois $z_1 z_2 \dots z_{n-1} = z_2 z_3 \dots z_{n-1} z_n$.

Então $z_1 = z_2 = \dots = z_{n-1} = z_n$, provando a afirmação.

Como $z_1 + \dots + z_n = c$, temos $z_1 = \dots = z_n = \frac{c}{n}$.

Logo $f(x_1, \dots, x_n) \leq f(z_1, \dots, z_n) = \left(\frac{c}{n}\right)^n$, pois $(x_1, \dots, x_n) \in M_c$. Assim,

$$x_1 \dots x_n \leq \left(\frac{x_1 + \dots + x_n}{n}\right)^n,$$

ou seja,

$$\sqrt[n]{x_1 \dots x_n} \leq \frac{x_1 + \dots + x_n}{n},$$

para quaisquer números reais positivos $x_1, \dots, x_n$, e a igualdade vale se, e só se, $x_1 = \dots = x_n$.

□

Exemplo 10.6. (Desigualdade de Hadamard)

Se X é uma matriz $n \times n$ cujas linhas são os vetores $X_i = (x_{i1}, \dots, x_{in})$, então

$$|\det X| \leq \|X_1\| \dots \|X_n\|,$$

onde $\|\cdot\|$ é a norma Euclidiana.

Se $\det X = 0$, a desigualdade é evidente. Se $\det X \neq 0$, então todos os vetores-linhas são não-nulos. Neste caso, podemos considerar os vetores unitários $W_i = \frac{X_i}{\|X_i\|}$, $i = 1, \dots, n$. Então, como $X_i = \|X_i\| W_i$, temos que $\det X = \|X_1\| \dots \|X_n\| \det W$, onde W é a matriz cujas linhas são os vetores unitários $W_1, \dots, W_n$. A desigualdade ficará provada se mostrarmos que $|\det W| \leq 1$. Mais geralmente:

Afirmção: Se $W = (w_{ij})$ é uma matriz $n \times n$ tal que $\sum_{i,j=1}^n w_{ij}^2 = n$ então $|\det W| \leq 1$.

De fato, sejam $f, \varphi : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$ as funções de classe C^∞ dadas por $f(X) = \det X$ e $\varphi(X) = \sum_{i,j=1}^n (x_{ij})^2$. Então, para todos $i, j = 1, \dots, n$, $\frac{\partial \varphi}{\partial x_{ij}}(X) = 2x_{ij}$ e $\frac{\partial f}{\partial x_{ij}}(X) = (-1)^{i+j} X_{[i,j]}$, onde $X_{[i,j]}$ é o determinante da matriz $(n-1) \times (n-1)$, obtida de X pela omissão da i -ésima linha e da j -ésima coluna.

Assim, para todo $n \in \mathbb{N}$, $\varphi^{-1}(n) = M$ é uma hipersuperfície compacta de classe C^∞ em $\mathbb{R}^{n^2}$. Mais precisamente, M é a esfera em $\mathbb{R}^{n^2}$ de centro na origem e raio $\sqrt{n}$.

Então, pelo método do Multiplicador de Lagrange, uma matriz $W = (w_{ij})$ é um ponto crítico de $f|_M$ se, e só se, $\sum_{i,j=1}^n w_{ij}^2 = n$ e $\text{grad } f(W) = \lambda \text{ grad } \varphi(W)$ para algum λ real, ou seja,

$$(-1)^{i+j} W_{[i,j]} = 2\lambda w_{ij}, \quad (*)$$

para quaisquer $i, j = 1, \dots, n$.

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

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

Logo $\det W = 2\lambda$.

Multiplicando agora (*) por w_{ij} , fixando i e somando em relação a j , obtemos:

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

Se W é uma matriz onde $f|_M$ atinge seu valor máximo ou mínimo, então $\det W \neq 0$ e, pela igualdade acima, $\|X_i\|^2 = \sum_{j=1}^n w_{ij}^2 = 1$ para todo $i = 1, \dots, n$, ou seja, os vetores-linha têm norma igual a 1.

Multiplicando (*) por w_{kj} , $k \neq i$, e somando em relação a j , temos:

$$\sum_{j=1}^n (-1)^{i+j} w_{kj} W_{[i,j]} = 2\lambda \sum_{j=1}^n w_{kj} w_{ij} = 2\lambda \langle W_k, W_i \rangle.$$

Logo $\langle W_k, W_i \rangle = 0$ para $k \neq i$, pois $\sum_{j=1}^n (-1)^{i+j} w_{kj} W_{[i,j]} = 0$, por ser o desenvolvimento, em relação à i -ésima linha, do determinante de uma matriz com duas linhas (a i -ésima e a k -ésima) iguais a W_k .

Assim, todo ponto $W \in M$ onde $f|_M$ atinge seu valor máximo ou mínimo é uma matriz cujas linhas são vetores unitários dois a dois ortogonais, ou seja W é uma matriz ortogonal. Logo $\det W = +1$, se W é um ponto de máximo, e $\det W = -1$, se W é um ponto de mínimo. Então $-1 \leq \det W \leq 1$ para todo $W \in M$, ou seja, $-\|X_1\| \dots \|X_n\| \leq \det X \leq \|X_1\| \dots \|X_n\|$ para toda matriz X .

E a igualdade $|\det X| = \|X_1\| \dots \|X_n\|$ ocorre se, e só se, $X_1, \dots, X_n$ são vetores dois a dois ortogonais, no caso em que $\det X \neq 0$. $\square$

Observação 10.5. O valor absoluto de $\det X$ é o volume do paralelepípedo n -dimensional determinado pelos vetores-linha $X_1, \dots, X_n$ da matriz X . Assim, a desigualdade de Hadamard significa, geometricamente, que se mantivermos constantes (não-nulos) os comprimentos desses vetores, $|\det X|$ torna-se máximo quando eles forem 2 a 2 ortogonais e, neste caso, o volume do paralelepípedo é o produto $\|X_1\| \dots \|X_n\|$ dos comprimentos de suas arestas.