Curso de Especialização em Matemática

Universidade Tecnológica Federal do Paraná

Câmpus Campo Mourão

Disciplina: Análise no $\mathbb{R}^n$

Professor Me. Márcio Hiran Simões

UTFPR - Campo Mourão

Junho de 2010

Capítulo 1

Noções Topológicas do Espaço Euclidiano

1.1 Métrica, Norma e Produto Interno

Definição 1.1 *Dado um conjunto não-vazio M, uma métrica em M é uma função $d : M \times M \longrightarrow \mathbb{R}$ satisfazendo:*

(d1) $d(x, y) \geq 0$ *e* $d(x, y) = 0 \Longleftrightarrow x = y, \ \forall x, y \in M.$

(d2) $d(x, y) = d(y, x), \ \forall x, y \in M.$

(d3) $d(x, y) \leq d(x, z) + d(z, y), \ \forall x, y, z \in M.$

Nessas condições dizemos que o par (M, d) é um espaço métrico.

Definição 1.2 *Dado E um espaço vetorial real, uma norma em E é uma aplicação $\| \cdot \| : E \longrightarrow \mathbb{R}$ satisfazendo as seguintes propriedades:*

(N1) $\|x\| > 0$ *se* $x \neq 0, \ \forall x \in E.$

(N2) $\|\lambda x\| = |\lambda| \|x\|, \ \forall \lambda \in \mathbb{R}, \ \forall x \in E.$

(N3) $\|x + y\| \leq \|x\| + \|y\|, \ \forall x, y \in E.$

Nessas condições, dizemos que o par $(E, \| \cdot \|)$ é um espaço vetorial normado.

Observação 1.1 *Se* $\lambda = 0$ *em (N2) então* $\|0\| = 0$.

Observação 1.2 *Se* $\lambda = -1$ *em (N2) então* $\|-x\| = \|x\|,\ \forall x \in E$.

Observação 1.3 *De (N3) podemos ver que*

$$0 = \|0\| = \|x + (-x)\| \leq \|x\| + \|-x\| = 2\|x\|$$

isto é, $\|x\| \geq 0,\ \forall x \in E$.

Definição 1.3 *Seja* E *um espaço vetorial. Um produto interno em* E *é uma aplicação* $\langle \cdot, \cdot \rangle : E \times E \longrightarrow \mathbb{R}$ *satisfazendo as seguintes propriedades:*

(P1) $\langle x, x \rangle > 0\ \forall x \in E, x \neq 0$.

(P2) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle,\ \forall x, y, z \in E$.

(P3) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle,\ \forall x, y \in E$.

(P4) $\langle x, y \rangle = \langle y, x \rangle,\ \forall x, y \in E$.

Exemplo 1.1 *Seja* E *um espaço vetorial real e* $\langle \cdot, \cdot \rangle$ *um produto interno em* E*. Então a aplicação*

$$\begin{array}{rccl} \| \cdot \| : & E & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \|x\| = \sqrt{\langle x, x \rangle} \end{array}$$

é uma norma (comumente chamada de norma induzida do produto interno $\langle \cdot, \cdot \rangle$*). De fato,*

(i) de (P1) temos que $\|x\| > 0$*, se* $x \neq 0$*, logo a propriedade (N1) está satisfeita.*

(ii) de (P3) temos que:

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

logo (N2) está satisfeita.

(iii) Para provar que (N3) é satisfeita por $\| \cdot \|$ *precisamos do seguinte teorema:*

Teorema 1.1 - *Desigualdade de Cauchy-Schwarz.*

$$|\langle x, y \rangle \leq \|x\| \|y\|,\ \forall x, y \in E.$$

Prova: Sejam $x, y \in E$ e $\lambda \in \mathbb{R}$. Então

$$\|x+\lambda y\|^2 \geq 0,\ \forall x, y \in E \Longleftrightarrow \langle x+\lambda y, x+\lambda y\rangle \geq 0,\ \forall x, y \in E \Longleftrightarrow$$

$$\langle x, x\rangle + \langle x, \lambda y\rangle + \langle \lambda y, x\rangle + \langle \lambda y, \lambda y\rangle \geq 0,\ \forall x, y \in E \Longleftrightarrow \|x\|^2 + 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2 \geq 0,\ \forall x, y \in E,$$

logo $\Delta = 4(\langle x, y\rangle)^2 - 4\|x\|^2\|y\|^2 \leq 0$, donde

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

ou ainda,

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

$\square$

Agora que provamos a desigualdade de Cauchy - Schwarz vamos utilizá-la para provar (iii). Observe que:

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \\
&\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Portanto $\|x+y\| \leq \|x\| + \|y\|$ e (N3) está satisfeita. Segue que a aplicação do Exemplo 1.1 é uma norma.

Exemplo 1.2 *Seja $(E, \|\cdot\|)$ um espaço vetorial normado. Então, definindo*

$$\begin{array}{rccl}
d: & E \times E & \longrightarrow & \mathbb{R} \\
& (x, y) & \longmapsto & d(x, y) = \|x-y\|
\end{array}$$

temos que d é uma métrica (denominada métrica associada à norma $\|\cdot\|$).

De fato, devemos mostrar que as propriedades (d1), (d2) e (d3) da definição de métrica são válidas. Por (N1) temos que $d(x, y) \geq 0$ e $d(x, y) = 0$ se, e somente se, $x = y,\ \forall x, y \in E$. Assim, (d1) está verificada.

Usando (N2) obtemos a seguinte igualdade:

$$d(x, y) = \|x-y\| = \|(-1)(y-x)\| = \|y-x\| = d(y, x).$$

Portanto, (d2) está verificada.

Para provar (d3) usamos (N3) na desigualdade abaixo:

$$d(x,y) = \|x-y\| = \|x-z+z-y\| \leq \|x-z\| + \|z-y\| = d(x,z) + d(z,y).$$

Isto mostra que (d3) é válida e conclui a prova de que a aplicação d do Exemplo 1.2 é uma métrica.

Exemplo 1.3 *(Produto interno, métrica e norma euclidiana)*

Seja $E = \mathbb{R}^n$. O produto interno euclidiano (ou usual) no $\mathbb{R}^n$ é definido por

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

onde $x = (x_1, x_2, \cdots, x_n)$ e $y = (y_1, y_2, \cdots, y_n)$.

A norma euclidiana é a norma proveniente do produto interno e é dada por

$$\|x\| = \left[x_1^2 + x_2^2 + \cdots + x_n^2\right]^{1/2}.$$

Do mesmo modo, a métrica euclidiana provém do produto interno euclidiano, ou seja,

$$d(x,y) = \left[(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2\right]^{1/2}.$$

Observação 1.4 *Existem diferentes formas de se obter um produto interno no $\mathbb{R}^n$. Por exemplo, considere a matriz $A = [a_{ij}]_{n\times n}$ com as seguintes propriedades:*
(i) A é simétrica.
(ii) A é positiva definida, isto é,

$$(x_1, \ldots, x_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix} = \sum_{i,j=1}^{n} a_{ij}x_ix_j > 0,$$

se $x > 0$..

Então A define um produto interno no $\mathbb{R}^n$ da seguinte forma

$$\langle x, y\rangle_A = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

Note que o produto interno usual é obtido com $A = I$.

Observação 1.5 *Da mesma forma que existem vários produtos internos no* $\mathbb{R}^n$*, existem também várias normas no* $\mathbb{R}^n$*. Veja duas de grande importância:*

(i) Norma do máximo.

$$\|x\|_M = \max\{|x_1|, |x_2|, \cdots, |x_n|\}.$$

(ii) Norma da Soma.

$$\|x\|_S = |x_1| + |x_2| + \cdots + |x_n| = \sum_{i=1}^{n} |x_i|.$$

Exercício 1.1 *Prove que* $\|\cdot\|_M$ *e* $\|\cdot\|_S$ *são normas no* $\mathbb{R}^n$ *e verifique a desigualdade:*

$$\|x\|_M \leq \|x\| \leq \|x\|_S \leq M\|x\|_M, \ \forall x \in \mathbb{R}^n.$$

Definição 1.4 *Duas normas* $\|\cdot\|_1$ *e* $\|\cdot\|_2$ *em* E *são ditas equivalentes se, e somente se, existem constantes* $c_1 > 0$ *e* $c_2 > 0$ *tais que*

$$\|x\|_1 \leq c_1\|x\|_2 \ e \ \|x\|_2 \leq c_2\|x\|_1, \ \forall x \in E.$$

Observação 1.6 *Seja* $\|\cdot\|$ *uma norma em um espaço vetorial* E *proveniente de um produto interno* $\langle\cdot,\cdot\rangle$*. Então, vale a identidade do paralelogramo*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \ \forall x, y \in E.$$

De fato, note que

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

e

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

Daí,

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \ \forall x, y \in E.$$

Exercício 1.2 *A norma do máximo em* $\mathbb{R}^n$ *provém de algum produto interno? E a norma da soma em* $\mathbb{R}^n$*?*

Capítulo 2

Sequências

Definição 2.1 *Sejam* $(E, \|\cdot\|)$ *um espaço vetorial normado. Definimos:*

1- $B_r(a) = \{x \in E; d(x,a) = \|x - a\| < r\}$ *(bola aberta).*

2- $B_r[a] = \{x \in E; d(x,a) = \|x - a\| \leq r\}$ *(bola fechada).*

3- $S_r(a) = \{x \in E; \|x - a\| = r\}$ *(esfera).*

4- Dados $x, y \in E$*, o segmento de reta de extremos* x *e* y *é o conjunto*

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

Definição 2.2 *Sejam* $(E, \|\cdot\|)$ *um espaço normado e* $X \subset E$*.*

(i) Dizemos que X *é um subconjunto convexo quando* X *contém qualquer segmento de reta cujos extremos pertencem a* X*.*

(ii) Dizemos que X *é limitado quando existe* $c > 0$ *tal que* $\|x\| \leq c,\ \forall x \in X$*.*

Definição 2.3 *Uma sequência* $(x_k)_{k \in \mathbb{N}} \subset E$ *(isto é,* $x_k \in E,\ \forall k \in \mathbb{N}$*) converge para um ponto* $a \in E$ *se para todo* $\epsilon > 0$*, existe* $k_0 \in \mathbb{N}$ *(que depende de* ϵ*) tal que*

$$\|x_k - a\| < \epsilon, \quad \textit{sempre que} \ \ k \geq k_0.$$

Notação:

$$x_k \longrightarrow a \text{ em } E \text{ ou } \lim_{k \to +\infty} x_k = a \text{ em } E.$$

Teorema 2.1 *Sejam* $\|\cdot\|_1$ *e* $\|\cdot\|_2$ *duas normas em* E. *Então,* $\|\cdot\|_1$ *é equivalente a* $\|\cdot\|_2$ *se, e somente se, para toda sequência* $(x_k)_{k\in\mathbb{N}} \subset E$ *tem-se*

$$\lim_{k\to+\infty} x_k = x \ em \ (E, \|\cdot\|_1) \Longleftrightarrow \lim_{k\to+\infty} x_k = x \ em \ (E, \|\cdot\|_2).$$

Demonstração: Suponha que $\|\cdot\|_1$ seja equivalente a $\|\cdot\|_2$. Então existem constantes $c_1 > 0$ e $c_2 > 0$ tais que

$$\|x\|_1 \leq c_1\|x\|_2 \text{ e } \|x\|_2 \leq c_2\|x\|_1.$$

Logo, se $(x_k)_{k\in\mathbb{N}} \subset E$ temos

$$\|x_k - x\|_1 \leq c_1\|x_k - x\|_2 \ \text{ e } \ \|x_k - x\|_2 \leq c_2\|x_k - x\|_1 \tag{2.1}$$

Daí, se $x_k \longrightarrow x$ em $(E, \|\cdot\|_1)$ tem-se que $\|x_k - x\|_1 \longrightarrow 0$ e, portanto, $\|x_k - x\|_2 \longrightarrow 0$ por (2.1), isto é, $x_k \to x$ em $(E, \|\cdot\|_2)$. Isto prova a necessidade.

Reciprocamente, suponhamos que $x_k \to x$ em $(E, \|\cdot\|_1)$ se, e somente se, $x_k \to x$ em $(E, \|\cdot\|_2)$. Vamos provar que existe $c_1 > 0$ tal que $\|x\|_1 \leq c_1\|x\|_2, \ \forall x \in E$. Suponhamos, por absurdo, que isto não ocorra. Então, existe uma sequência $(x_k)_{k\in\mathbb{N}} \subset E$ tal que

$$\|x_k\|_1 \geq k\|x_k\|_2, \ \forall k \in \mathbb{N}.$$

Assim,

$$\frac{\|x_k\|_2}{\|x_k\|_1} < \frac{1}{k}, \ \forall k \in \mathbb{N}.$$

Denotemos $v_k = \frac{x_k}{\|x_k\|_1}, \ \forall k \in \mathbb{N}$ e teremos:

$$\|v_k\|_2 < \frac{1}{k} \Longrightarrow v_k \longrightarrow 0 \ \text{ em } \ (E, \|\cdot\|_2)$$

e

$$\|v_k\|_1 = 1, \ \forall k \in \mathbb{N} \Longrightarrow v_k \nrightarrow 0 \ \text{ em } \ (E, \|\cdot\|_1).$$

Isto contradiz a hipótese. A demonstração é análoga para o caso da constante c_2.

$\square$

Dada uma sequência $(x_k)_{k\in\mathbb{N}} \subset \mathbb{R}^n$ então $x_k = (x_{k1}, x_{k2}, \cdots, x_{kn}), \ k \in \mathbb{N}$, o que equivale a dar n sequências $(x_{ki})_{k\in\mathbb{N}}, \ i = 1, 2, \cdots, n$, reais. Portanto, são válidas todas as propriedades de sequências de números reais para as sequências do $\mathbb{R}^n$.

Teorema 2.2 *Sejam* $(x_k)_{k\in\mathbb{N}} \subset \mathbb{R}^n$ *e* $a = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n$. *Então*

$$\lim_{k\to+\infty} x_k = a \in \mathbb{R}^n \Longleftrightarrow \lim_{k\to+\infty} x_{ki} = a_i,\ \forall i = 1, 2, \cdots, n.$$

Demonstração: Para demonstrarmos este fato é conveniente considerar a norma do máximo. Suponhamos que $\lim\limits_{k\to+\infty} x_k = a$ em $\mathbb{R}^n$. Então, para $i = i, 2, \cdots, n$, temos

$$|x_{ki} - a_i| \leq \max\{|x_{k1} - a_1|, |x_{k2} - a_2|, \cdots, |x_{kn} - a_n|\} = \|x_k - a\|_M.$$

Por hipótese, temos que $\|x_k - a\|_M \longrightarrow 0$, quando $k \to +\infty$. Portanto, $|x_{ki} - a| \longrightarrow 0$, quando $k \to +\infty$, donde $\lim\limits_{k\to+\infty} x_{ki} = a_i$ para todo $i = 1, 2, \cdots, n$.

Reciprocamente, suponhamos que $\lim\limits_{k\to+\infty} x_{ki} = a_i$, para todo $i = 1, 2, \cdots, n$. Assim, dado $\epsilon > 0$, existem $k1, k2, \cdots, kn \in \mathbb{N}$ tais que

$$|x_{ki} - a_i| < \epsilon, \quad \text{sempre que } k \geq ki$$

para cada $i = 1, 2, \cdots, n$.

Seja $k_0 = \max_{1\leq i\leq n}\{ki\}$. Então,

$$|x_{ki} - a_i| < \epsilon, \quad \text{se } k \geq k_0,\ \forall i = 1, 2, \cdots, n.$$

Logo,

$$\|x_k - a\|_M = \max\{|x_1 - a_1|, |x_2 - a_2|, \cdots, |x_n - a_n|\} < \epsilon$$

sempre que $k \geq k_0$. Segue que

$$\lim_{k\to+\infty} x_k = a \in \mathbb{R}^n.$$

□

Corolário 2.1 *Sejam* $(x_k)_{k\in\mathbb{N}}$ *e* $(y_k)_{k\in\mathbb{N}}$ *sequências do* $\mathbb{R}^n$ *e* $(\lambda_k)_{k\in\mathbb{N}}$ *uma sequência real tais que*

$$\lim_{k\to+\infty} x_k = a, \quad \lim_{k\to+\infty} y_k = b \quad e \quad \lim_{k\to+\infty} \lambda_k = \lambda.$$

Então:

(1) $\lim\limits_{k\to+\infty} (x_k + y_k) = a + b$;

(2) $\lim\limits_{k\to+\infty} \lambda_k x_k = \lambda a$;

(3) $\lim\limits_{k\to+\infty} \langle x_k, y_k\rangle = \langle a, b\rangle$ *em* $\mathbb{R}$;

(4) $\lim\limits_{k\to+\infty} \|x_k\| = \|a\|$ *em* $\mathbb{R}$.

Demonstração: Sejam $x_k = (x_{k1}, \cdots, x_{kn})$, $y_k = (y_{k1}, \cdots, y_{kn})$, $a = (a_1, \cdots, a_n)$, $b = (b_1, \cdots, b_n)$, com $\lim\limits_{k\to+\infty} x_{ki} = a_i$ e $\lim\limits_{k\to\infty} y_{ki} = b_i$, para todo $i = 1, \cdots, n$.

Para provar (1) temos que

$$x_k + y_k = (x_{k1} + y_{k1}, \cdots x_{kn} - y_{kn}).$$

Pelo Teorema 2.2 e sabendo que, para sequências reais, vale a igualdade

$$\lim_{k\to+\infty}(x_{ki} + y_{ki}) = \lim_{k\to+\infty} x_{ki} + \lim_{k\to+\infty} y_{ki}$$

temos que

$$\lim_{k\to+\infty}(x_k + y_k) = \lim_{k\to+\infty}(x_{ki} + y_{ki}) = \lim_{k\to+\infty} x_{ki} + \lim_{k\to+\infty} y_{ki} = \lim_{k\to+\infty} x_k + \lim_{k\to+\infty} y_k = a + b$$

Para provar (2) temos que $\lambda_k x_k = \lambda_k(x_{k1}, \cdots, x_{kn}) = (\lambda_k x_{k1}, \cdots, \lambda_k x_{kn})$. Pelo Teorema 2.3 e sabendo que

$$\lim_{k\to+\infty} \lambda_k x_{ki} = \lim_{k\to+\infty} \lambda_k \lim_{k\to+\infty} x_{ki}$$

quando esses limites existem, temos

$$\lim_{k\to+\infty} \lambda_k x_k = \lim_{k\to+\infty} \lambda_k x_{ki} = \lim_{k\to+\infty} \lambda_k \lim_{k\to+\infty} x_{ki} = \lambda \lim_{k\to+\infty} x_k = \lambda a.$$

Para a prova de (3) temos que

$$\langle x_k, y_k\rangle = x_{k1}y_{k1} + \cdots + x_{kn}y_{kn}$$

logo

$$\begin{aligned}
\lim_{k\to+\infty}\langle x_k, y_k\rangle &= \lim_{k\to+\infty}(x_{k1}y_{k2} + \cdots + x_{kn}y_{kn}) \\
&= \lim_{k\to+\infty} x_{k1}y_{k1} + \cdots + \lim_{k\to+\infty} x_{kn}y_{kn} \\
&= \lim_{k\to+\infty} x_{k1} \lim_{k\to+\infty} y_{k1} + \cdots + \lim_{k\to+\infty} x_{kn} \lim_{k\to+\infty} y_{kn} \\
&= a_1b_1 + \cdots + a_nb_n \\
&= \langle a, b\rangle.
\end{aligned}$$

Resta-nos provar a igualdade (4). Para tal, necessitamos demonstrar a seguinte desigualdade

$$\Big|\|x_k\| - \|a\|\Big| \leq \|x_k - a\|$$

De fato, note que

$$\|x_k\| = \|x_k - a + a\| \leq \|x_k - a\| + \|a\|.$$

Daí,

$$\|x_k\| - \|a\| \leq \|x_k - a\|.$$

Por outro lado,

$$\|a\| = \|a - x_k + x_k\| \leq \|a - x_k\| + \|x_k\| = \|x_k - a\| + \|x_k\|.$$

Ou seja,

$$\|a\| - \|x_k\| \leq \|x_k - a\|$$

Portanto, da definição de valor absoluto, concluímos que,

$$\Big| \|x_k\| - \|a\| \Big| \leq \|x_k - a\|$$

Assim, dado $\epsilon > 0$, existe $k \in \mathbb{N}$ tal que se $k \geq k_0$ tem-se que

$$\Big| \|x_k\| - \|a\| \Big| \leq \|x_k - a\| < \epsilon$$

visto que $\lim\limits_{k \to +\infty} x_k = a$.

$\square$

Teorema 2.3 *(Bolzano - Weierstrass)*

Toda sequência limitada do $\mathbb{R}^n$ possui uma subsequência convergente.

Demonstração: Sabemos que o Teorema de Bolzano - Weierstrass é válido na reta real, isto é, toda sequência limitada de números reais possui uma subsequência convergente. Seja $(x_k)_{k \in \mathbb{N}}$ uma sequência limitada do $\mathbb{R}^n$. Então, a primeira coordenada $(x_{k1})_{k \in \mathbb{N}}$ de $(x_k)_{k \in \mathbb{N}}$ é uma sequência limitada e, portanto, podemos obter um subconjunto infinito $\mathbb{N}_1 \subset \mathbb{N}$ e um número real a_1 tal que $\lim\limits_{k \to +\infty} x_{k1} = a_1$.

Do mesmo modo, a sequência $(x_{k2})_{k \in \mathbb{N}_1}$ da segunda coordenada de $(x_k)_{k \in \mathbb{N}}$ é limitada e, portanto, podemos obter um subconjunto infinito $\mathbb{N}_2 \subset \mathbb{N}_1$ e um número real a_2 tal que $\lim\limits_{k \to \infty} x_{k2} = a_2$.

Seguindo o mesmo raciocínio, podemos obter conjuntos infinitos

$$\mathbb{N} \supset \mathbb{N}_1 \supset \mathbb{N}_2 \supset \cdots \supset \mathbb{N}_i$$

e números reais a_i tais que

$$\lim_{k \in \mathbb{N}_i} x_{ki} = a_i$$

Pondo $a = (a_1, a_2, \cdots, a_n)$ o Teorema 2.2 nos garante que

$$\lim_{k \in \mathbb{N}} x_k = a.$$

$\square$

Teorema 2.4 *Duas normas quaisquer no espaço $\mathbb{R}^n$ são equivalentes.*

Demonstração: Pela transitividade da equivalência das normas, basta provar que, dada uma norma $\| \cdot \|$ arbitrária no $\mathbb{R}^n$, ela é equivalente à norma da soma $\| \cdot \|_S$.

Dado $x \in \mathbb{R}^n$ então $x = x_1 e_1 + x_2 + e_2 + \cdots + x_n e_n$ onde $\{e_1, e_2, \cdots, e_n\}$ é a base canônica do $\mathbb{R}^n$. Então,

$$\begin{aligned} \|x\| &= \|x_1 e_1 + x_2 e_2 + \cdots + c_n e_n\| \\ &\leq |x_1| \|e_1\| + |x_2| \|e_2\| + \cdots + |x_n| \|e_n\| \\ &\leq \max_{1 \leq i \leq n} \{\|e_i\|\} (|x_1| + |x_2| + \cdots + |x_n|) \\ &= c_1 \|x\|_E. \end{aligned}$$

Resta provar que existe $c_2 > 0$ tal que $\|x\|_E \leq c_2 \|x\|$. Suponha, por absurdo, que isto não ocorra. Então, existe uma sequência $(x_k)_{k \in \mathbb{N}}$ tal que

$$\|x_k\|_S > k \|x_k\|, \ \forall k \in \mathbb{N}$$

ou seja,

$$\frac{\|x_k\|_S}{\|x_k\|} < \frac{1}{k}, \ \forall k \in \mathbb{N}.$$

Denotemos $v_k = \dfrac{x_k}{\|x_k\|_S}$ e teremos

$$\|v_k\|_S = 1, \ \forall k \in \mathbb{N} \tag{2.2}$$

e

$$\|v_k\| < \frac{1}{k}, \ \forall k \in \mathbb{N}$$

Logo, a sequência $(v_k)_{k\in\mathbb{N}}$ é limitada em $(\mathbb{R}^n, \|\cdot\|_S)$. Pelo Teorema de Bolzano - Weierstrass, existem uma subsequência $(v_{k_j})_{j\in\mathbb{N}}$ e $v \in \mathbb{R}^n$ tal que

$$\lim_{j\to+\infty} v_{k_j} = v \text{ em } (\mathbb{R}^n, \|\cdot\|_S)$$

E, pelo item 4 do corolário 2.1 temos que $\lim\limits_{j\to+\infty} \|v_{k_j}\|_S = \|v\|_S = 1$, de acordo com a equação(2.2). Assim, $\|v\|_S \neq 0$.

Por outro lado, $\|v_{k_j}\| \leq \frac{1}{k_j}$ implica que $\|v_{k_j}\| \longrightarrow 0$ quando $j \to +\infty$. Daí,

$$\|v\| = \|v - v_{k_j} + v_{k_j}\| \leq \|v - v_{k_j}\| + \|v_{k_j}\|.$$

Pela primeira parte da demonstração, nós temos que

$$\|v - v_{k_j}\| \leq c_1\|v - v_{k_j}\| \longrightarrow 0$$

quando $j \to +\infty$, assim como $\|v_{k_j}\| \to 0$ quando $j \to +\infty$. Assim

$$\|v\| \leq 0 \Longrightarrow \|v\| = 0$$

o que é uma contradição visto que tinhamos provado que $\|v\| \neq 0$. Isto mostra que existe $c_2 > 0$ tal que

$$\|x\|_S \leq c_2\|x\|.$$

$\square$

Capítulo 3

Noções Topológicas no $\mathbb{R}^n$

Definição 3.1 *Seja $X \subset \mathbb{R}^n$. O ponto $a \in \mathbb{R}^n$ é chamado* **ponto de acumulação de X** *quando, para todo $\epsilon > 0$,*

$$B_\epsilon(a) \cap (\{X - a\}) \neq \emptyset.$$

Definição 3.2 *Dizemos que o ponto $a \in X \subset \mathbb{R}^n$ é um ponto interior do conjunto X se existir $\epsilon > 0$ tal que*

$$B_\epsilon(a) \subset X.$$

Definição 3.3 *Seja $X \subset \mathbb{R}^n$. Denomina-se fronteira de X o conjunto dos pontos $y \in \mathbb{R}^n$ tais que, para todo $\epsilon > 0$, a bola $B_\epsilon(y)$ contém pontos de X e do complementar de X.*

Definição 3.4 *O conjunto $X \subset \mathbb{R}^n$ é aberto quando todo ponto de X é ponto interior de X ($X = intX$).*

Definição 3.5 *Seja $X \subset \mathbb{R}^n$. O ponto $a \in \mathbb{R}^n$ é chamado de ponto aderente a X quando existe uma sequência $(x_k)_{k\in\mathbb{N}} \subset X$ tal que*

$$\lim_{k\to+\infty} x_k = a.$$

Definição 3.6 *O conjunto de todos os pontos aderentes a X é denominado o fecho de X e é denotado por $\overline{X}$.*

Definição 3.7 *Um conjunto $X \subset \mathbb{R}^n$ é fechado quando X contém todos os seu pontos aderentes.*

Definição 3.8 *Dizemos que $X \subset \mathbb{R}^n$ é compacto se X é fechado e limitado.*

Definição 3.9 *Seja $X \subset \mathbb{R}^n$. Uma cobertura de X é uma família arbitraria de subconjuntos $(A_\lambda)_{\lambda \in L}$ do $\mathbb{R}^n$ tais que*

$$X \subset \bigcup_{\lambda \in L} A_\lambda.$$

Observação 3.1 *Se $A_\lambda \subset \mathbb{R}^n$ é aberto para todo $\lambda \in L$ então $(A_\lambda)_{\lambda \in L}$ é uma cobertura aberta.*

Observação 3.2 *Se L é enumerável então $(A_\lambda)_{\lambda \in L}$ é uma cobertura enumerável.*

Observação 3.3 *Se L é finito então $(A_\lambda)_{\lambda \in L}$ é uma cobertura finita.*

Definição 3.10 *Seja $Y \subset X \subset \mathbb{R}^n$. Dizemos que Y é denso em X quando para todo $a \in X$ existe uma sequência $(y_k)_{k \in \mathbb{N}} \subset Y$ tal que*

$$\lim_{k \to +\infty} y_k = a.$$

Proposição 3.1 *Todo conjunto $X \subset \mathbb{R}^n$ contém um subconjunto E enumerável e denso em X.*

Demonstração: Sabemos da análise real que o conjunto $\mathcal{Q}$ dos números racionais é enumerável e denso em $\mathbb{R}$. Seja $\mathcal{Q}^+ = \{r_1, r_2, \cdots, r_i, \cdots\}$ (uma enumeração de $\mathcal{Q}$) então $\mathcal{Q}^n = \{q_1, q_2, \cdots, q_i, \cdots\}$ (onde q_i é uma n-upla de números racionais) é enumerável e denso no $\mathbb{R}^n$.

Seja $\beta = \{B_{r_i}(q_j); i, j \in \mathbb{N}\} = \{B_1, B_2, \cdots, B_k, \cdots\}$ e tomemos β' uma subfamília de β tal que

$$\beta' = \{(B_{k_i})_{k_i \in \mathbb{N}'}, \mathbb{N}' \subset \mathbb{N}; B_{k_i} \cap X \neq \emptyset\}.$$

Definimos $E = \{x_{k_i}\}$ onde $x_{k_i} \in (B_{k_i} \cap X)$. Então E é enumerável e $E \subset X$.

Resta provar que E é denso em X. Sejam $x \in X$ e $\epsilon > 0$. Existe $r \in \mathcal{Q}$ tal que $r < \frac{\epsilon}{2}$. Como $\mathcal{Q}^n$ é denso em $\mathbb{R}^n$, existe $q \in \mathcal{Q}^n$ tal que

$$q \in B_r(x) \Longrightarrow \|x - q\| < r \Longrightarrow x \in B_r(q) = B_{k_i} \quad \text{para algum } i \in \mathbb{N}'$$

logo x e $x_{k_i} \in B_{k_i} \Rightarrow \|x - x_{k_i}\| < 2r < \epsilon$ e, portanto, $x \in B_\epsilon(x) \subset E$. Isto mostra que E é denso em X.

□

Proposição 3.2 *Seja $(K_i)_{i \in \mathbb{N}}$ uma sequência de conjuntos compactos não vazios. Se $K_1 \supset K_2 \supset \cdots \supset K_i \supset \cdots$ então*

$$\bigcap_{i=1}^{\infty} K_i = K$$

é compacta e não vazia.

Demonstração: É fácil ver que K é compacto pois cada K_i é compacto e $K_1 \supset K_2 \supset \cdots$. Vamos provar que $K \neq \emptyset$. Seja $(x_i)_{i \in \mathbb{N}}$ uma sequência do $\mathbb{R}^n$ tal que $x_i \in K_i$, $\forall i \in \mathbb{N}$. Então $(x_i)_{i \in \mathbb{N}} \subset K_1$ o que implica em $(x_i)_{i \in \mathbb{N}}$ limitada. Pelo Teorema de Bolzano - Weierstrass, existe uma subsequência $(x_{i_j})_{j \in \mathbb{N}}$ de $(x_i)_{i \in \mathbb{N}}$ que converge para um ponto $x \in \mathbb{R}^n$. Dado $i \in \mathbb{N}$ então $x_{i_j} \in K_i$, $\forall i_j \geq i$ donde

$$x \in \bigcap_{i=1}^{\infty} K_i.$$

□

Proposição 3.3 *Se $X \subset \mathbb{R}^n$ então toda cobertura aberta $(A_\lambda)_{\lambda \in L}$ de X possui uma subcobertura enumerável.*

Demonstração: Seja $E = \{x_1, x_2, \cdots, x_i, \cdots\} \subset X$ um subconjunto enumerável, denso em X. Consideremos o conjunto β de todas as bolas abertas $B_r(x)$, com centro em um ponto de E, raio racional e tais que cada uma delas está contida em algum A_λ. Observe que β é um conjunto enumerável de bolas abertas. Afirmamos que as bolas $B \in \beta$ cobrem X. Com efeito, dado $x \in X$, existe $\lambda \in L$ tal que $x \in A_\lambda$. Como A_λ é aberto, existe $r > 0$ racional tal que $B_{2r}(x) \subset A_\lambda$. Sendo E denso em X, podemos encontrar $x_i \in E$ com $|x - x_i| < r$. Então, $x \in B_r(x_i)$. Para mostrar que $B_r(x_i) \in \beta$, restar ver que esta bola está contida em A_λ. Ora,

$$y \in B_r(x_i) \Rightarrow |y - x_i| < r \Rightarrow |y - x| \leq |y - x_i| + |x_i - x| < 2r \Rightarrow y \in B_{2r}(x) \subset A_\lambda.$$

Isto conclui a verificação de que as bolas $B \in \beta$ cobrem X. Tomando uma enumeração $B_1, \cdots, B_i, \cdots$ para essas bolas e escolhendo para cada $i \in \mathbb{N}$ um índice $\lambda_i \in L$ tal que $B_i \subset A_{\lambda_i}$, concluímos que

$$X \subset A_{\lambda_1} \cup A_{\lambda_2} \cup \cdots \cup A_{\lambda_i} \cup \cdots$$

$\square$

Teorema 3.1 *O conjunto $K \subset \mathbb{R}^n$ é compacto se, e somente se, toda cobertura aberta de K admite uma subcobertura finita.*

Demonstração: Seja $K \subset \mathbb{R}^n$ um conjunto compacto, isto é, K é fechado e limitado. Seja $(A_\lambda)_{\lambda \in L}$ uma cobertura aberta de K. Pela proposição 3.3 existe uma subcobertura enumerável de K, isto é,

$$K \subset (A_1 \cup A_2 \cup \cdots \cup A_i \cup \cdots).$$

Para cada $i \in \mathbb{N}$ definimos

$$K_i = K \cap \complement_{\mathbb{R}^n}^{A_1 \cup A_2 \cup \cdots \cup A_i}.$$

Então, valem as seguintes propriedades para K_i:

(i) K_i é compacto, $\forall i \in \mathbb{N}$ pois K é fechado e $\complement_{\mathbb{R}^n}^{A_1 \cup A_2 \cup \cdots \cup A_i}$ é fechado.

(ii) $K \supset K_1 \supset K_2 \supset$ pois

- se $x \in K_1 \Rightarrow x \in K$ e $x \in \complement_{\mathbb{R}^n}^{A_1} \Rightarrow x \in K$ e $x \notin A_1$ logo $K \supset K_1$;

- se $x \in K_2 \Rightarrow x \in K$ e $x \in \complement_{\mathbb{R}^n}^{A_1 \cup A_2} \Rightarrow x \in K$ e $\notin A_1 \cup A_2 \Rightarrow x \in K$ e $x \notin A_1, x \notin A_2$. Como $x \in K$ e $x \notin A_1$ então $K_2 \supset K$.

Analogamente para os demais casos.

(iii) $\displaystyle\bigcap_{i=1}^{\infty} K_i = \emptyset$.

De fato, se $x \in K$ então $x \in A_i$ para algum i. Isto implica que $x \notin \complement_{\mathbb{R}^n}^{A_1 \cup A_2 \cup \cdots \cup A_i}$ e, portanto, $x \notin K_i$. Segue que

$$\bigcap_{i=1}^{\infty} K_i = \emptyset.$$

Pela proposição 3.2, devemos ter $K_{i_0} = \emptyset$ para algum i_0, já que $\cap_{i=1}^{\infty} K_i = \emptyset$. Daí, pela propriedade (ii) acima, $K_i = \emptyset$, $\forall i \geq i_0$. Logo $K \cap \complement_{\mathbb{R}^n}^{A_1 \cup \cdots \cup A_{i_0}} = \emptyset$. Se $x \in K$ então $x \notin \complement_{\mathbb{R}^n}^{A_1 \cup \cdots \cup A_{i_0}}$, isto é, $x \in A_1 \cup \cdots \cup A_{i_0}$. Segue que $\{A_i\}_{1 \leq i \leq i_0}$ é uma subcobertura finita para K.

Reciprocamente, suponhamos que toda cobertura aberta de K admite uma subcobertura finita. Provemos inicialmente que K é limitado. De fato, para cada $x \in K$ considere a bola $B_1(x)$. Então $\{B_1(x)\}_{x \in K}$ é uma cobertura aberta de K. Pela hipótese, existe uma subcobertura finita para K, ou seja,

$$K \subset \{B_1(x_1) \cup B_1(x_2) \cup \cdots \cup B_1(x_j)\}$$

logo K é limitado.

Provemos agora que K é um conjunto fechado. Para tanto, suponhamos por absurdo que K não é fechado. Isto é, existe $a \in \overline{K}$ tal que $a \notin K$. Para cada $i \in \mathbb{N}$ definimos

$$A_i = \complement_{\mathbb{R}^n}^{B_{1/i}[a]}.$$

Então, temos as seguintes propriedades para A_i:

1. A_i é aberto para todo $i \in \mathbb{N}$.

2. $A_1 \subset A_2 \subset A_3 \subset \cdots$

3. $K \subset (A_1 \cup A_2 \cup \cdots \cup A_i \cup \cdots)$

Com efeito, para provar 3, temos que $a \notin K$ e se $x \in K$ então $x \neq a$. Daí, existe $i \in \mathbb{N}$ suficientemente grande tal que $x \notin B_{1/i}[a]$. Isto implica que $x \in A_i$ e assim, $x \in (A_1 \cup A_2 \cup \cdots \cup A_i \cup \cdots)$.

Da hipótese, a cobertura $(A_i)_{i\in\mathbb{N}}$ de K admite uma subcobertura finita, isto é, $K \subset (A_{1_i} \cup A_{2_i} \cup \cdots \cup A_{i_p})$. Pela propriedade 2 acima, temos que

$$A_{i_1} \cup \cdots \cup A_{i_p}$$

é igual ao conjunto de maior índice. Então

$$K \subset A_i, \quad \text{para algum} \quad i$$

logo $B_{1/i}[a] \cap K = \emptyset$ o que é um absurdo pois a é ponto aderente de K. Segue então que K é fechado.

$\square$

Capítulo 4

Funções Contínuas

Definição 4.1 *Sejam* $X \subset \mathbb{R}^m$ *e* $f : X \longrightarrow \mathbb{R}^n$ *uma função. Dizemos que* f *é contínua no ponto* $a \in X$ *quando para todo* $\epsilon > 0$ *existe* $\delta > 0$ *(*δ *dependendo de* a *e de* ϵ*) tal que se* $x \in X$ *e* $\|x - a\|_{\mathbb{R}^m} < \delta$ *então* $\|f(x) - f(a)\|_{\mathbb{R}^n} < \epsilon$.

Exemplo 4.1 *As projeções* $\Pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1, \cdots, n$ *que a cada* $y \in \mathbb{R}^n$ *associa* $\Pi_i(y) = y_i$ *são contínuas.*

De fato, dado $\epsilon > 0$ tomemos $\delta = \epsilon$, e daí, se $0 < \|x - y\|_{\mathbb{R}^n} < \delta$ então

$$|\Pi_i(x) - \Pi_i(y)|_{\mathbb{R}} = |x_i - y_i| \leq \|x - y\|_{\mathbb{R}^n} < \delta = \epsilon.$$

Observação 4.1 *A composta de uma função contínua é contínua, isto é, dados* $X \subset \mathbb{R}^m, Y \subset \mathbb{R}^n, f : X \Rightarrow \mathbb{R}^n$ *contínua no ponto* $a \in X$, *com* $f(X) \subset Y$, $g : Y \Rightarrow \mathbb{R}^p$ *contínua no ponto* $b = f(a)$, *então composta* $g \circ f : X \Rightarrow \mathbb{R}^p$ *é contínua no ponto* a.

Demonstração: Dado arbitrariamente $\epsilon > 0$ existe, em virtude da continuidade de g, um número $\eta > 0$ tal que $y \in Y, |y - f(a)| < \eta \Longrightarrow |g(y) - g(f(a))| < \epsilon$. Por sua vez, a partir de η, a continuidade de f nos fornece $\delta > 0$ tal que $x \in X, |x - a| < \delta \Longrightarrow |f(x) - f(a)| < \eta$.

Segue-se que $x \in X, |x - a| < \delta \Longrightarrow |g(f(x)) - g(f(a))| < \epsilon$, logo $g \circ f$ é contínua no ponto a.

$\square$

Seja $X \subset \mathbb{R}^m$. Dar uma função $f : X \Rightarrow \mathbb{R}^n$ é o mesmo que dar n funções reais $f_1, f_2, \cdots, f_n : X \Rightarrow \mathbb{R}$ definidas por $f_i = \Pi_i \circ f$, as quais são chamadas as coordenadas da aplicação f. Para todo $x \in X$, temos $f(x) = (f_1(x), f_2(x), \cdots, f_n(x))$. Para indicar que as f_i são as funções coordenadas de f, escreve-se $f = (f_1, f_2, \cdots, f_n)$.

Teorema 4.1 *Uma função $f = (f_1, f_2, \cdots, f_n) : X \subset \mathbb{R}^m \Rightarrow \mathbb{R}^n$ é contínua no ponto $a \in X$ se, e somente se, para todo $i = 1, 2, \cdots, n$ tem-se que $f_i : X \Rightarrow \mathbb{R}$ é contínua no ponto a.*

Demonstração: Suponhamos que $f = (f_1, f_2, \cdots, f_n) : X \subset \mathbb{R}^m \Rightarrow \mathbb{R}^n$ é contínua no ponto a. Sabemos que as projeções $\Pi_i : \mathbb{R}^n \Rightarrow \mathbb{R}$ são contínuas. Além disso,

$$(\Pi_i \circ f)(x) = \Pi_i(f(x)) = f_i(x)$$

logo, f_i é contínua para todo $i = 1, 2, \cdots, n$.

Reciprocamente, suponhamos que $f_i : X \Rightarrow \mathbb{R}$ é contínua em a, qualquer que seja $i = 1, 2, \cdots, n$. Daí, dado $\epsilon > 0$, existe $\delta_i > 0$ tal que se $x \in X$ e $\|x - a\|_{\mathbb{R}^m} < \delta_i$ então $|f_i(x) - f_i(a)| < \epsilon$.

Tomando $\delta_0 = \min\{\delta_1, \cdots, \delta_n\}$ temos que, se $x \in X$ e $\|x - a\|_{\mathbb{R}^n} < \delta_0$ então $|f_i(x) - f_i(a)| < \epsilon$, para todo $i = 1, \cdots, n$.

Isto implica que

$$\|f(x) - f(a)\|_{M,\mathbb{R}^n} = \max\{|f_1(x) - f_1(a)|, \cdots, |f_n(x) - f_n(a)|\} < \epsilon.$$

Portanto, f é contínua no ponto a.

$\square$

Teorema 4.2 *Uma função $f : X \subset \mathbb{R}^m \Rightarrow \mathbb{R}^n$ é contínua em um ponto $a \in X$ se, e somente se, para toda sequência $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^m$ tal que $\lim_{k \to +\infty} x_k = a$ em $\mathbb{R}^m$ tem-se que $\lim_{k \to +\infty} f(x_k) = f(a)$ em $\mathbb{R}^n$.*

Demonstração: Suponhamos que f e contínua no ponto a e que $\lim_{k \to +\infty} x_k = a$. Assim, dado $\epsilon > 0$, existe $\delta > 0$ tal que se $x \in X, \|x - a\|_{\mathbb{R}^m} < \delta$ então $\|f(x) - f(a)\|_{\mathbb{R}^n} < \epsilon$. Como $\lim_{k \to +\infty} x_k = a$, existe $k_0 \in \mathbb{N}$ tal que $k > k_0$ implica $\|f(x_k) - f(a)\|_{\mathbb{R}^n} < \epsilon$. Logo $\lim_{k \to +\infty} f(x_k) = f(a)$.

Reciprocamente, suponhamos que f não seja contínua no ponto a. Então existe $\epsilon > 0$ tal que, para cada $k \in \mathbb{N}$, podemos obter $x_k \in X$ com $\|x_k - a\|_{\mathbb{R}^m} < 1/k$ e

$$\|f(x_k) - f(a)\|_{\mathbb{R}^m} \geq \epsilon.$$

Então, $\lim\limits_{k \to +\infty} x_k = a$ sem que seja $\lim\limits_{k \to +\infty} f(x_k) = f(a)$ o que contradiz a hipótese.

$\square$

Teorema 4.3 *A função $f : X \subset \mathbb{R}^m \Rightarrow \mathbb{R}^n$ é contínua se, e somente se, para todo aberto $A \subset \mathbb{R}^m$ tem-se que $f^{-1}(A) \subset \mathbb{R}^m$ é aberto em X.*

Demonstração: Suponhamos que f é contínua e $A \subset \mathbb{R}^n$ é um aberto. Tomemos um ponto $a \in f^{-1}(A)$. Então $f(a) \in A$. Pela definição de conjunto aberto, existe $\epsilon > 0$ tal que $B_\epsilon(f(a)) \subset A$. Sendo f contínua, existe $\delta > 0$ tal que se $x \in X, \|x - a\|^{\mathbb{R}^m} < \delta$ então $\|f(x) - f(a)\|_{\mathbb{R}^n} < \epsilon$. Isto significa que

$$f(B_\delta(a) \cap X) \subset B_\epsilon(f(a)) \subset A,$$

donde

$$(B_\delta(a) \cap X) \subset f^{-1}(A)$$

logo $f^{-1}(A)$ é aberto em X.

Reciprocamente, se a imagem inversa por f de todo aberto de $\mathbb{R}^n$ é aberto em X, então, dado $a \in X$ e $\epsilon > 0$, como $B_\epsilon(f(a))$ é aberto, concluímos que $A = \{x \in X; \|f(x) - f(a)\|_{\mathbb{R}^n} < \epsilon\}$ é aberto em X. Evidentemente, $a \in A$. Logo existe $\delta > 0$ tal que $(B_\delta(a) \cap X) \subset A$. Isto significa que $x \in X, \|x - a\|_{\mathbb{R}^m} < \delta$ então $\|f(x) - f(a)\|_{\mathbb{R}^n} < \epsilon$, ou seja, que f é contínua no ponto a. Como $a \in X$ é arbitrário, segue-se que f é contínua em X.

$\square$

Capítulo 5

Aplicações Lineares

5.1 Primeiros resultados

Seja $T : \mathbb{R}^m \Rightarrow \mathbb{R}^n$ uma função linear. Então, dado $x \in \mathbb{R}^m$ temos

$$
\begin{aligned}
T(x) &= T(x_1 e_1 + \cdots + x_m e_m) \\
&= T(x_1 e_1) + \cdots + T(x_m e_m) \\
&= x_1 T(e_1) + \cdots + x_m T(e_m) \\
&= \sum_{i=1}^{n} x_i T(e_i)
\end{aligned}
$$

onde $\beta = \{e_1, \cdots, e_m\}$ é uma base canônica do $\mathbb{R}^m$.

Seja $c = \max\{\|T(e_1)\|_{\mathbb{R}^n}, \|T(e_2)\|_{\mathbb{R}^n}, \cdots, \|T(e_m)\|_{\mathbb{R}^n}\}$ então

$$
\begin{aligned}
\|T(x)\|_{\mathbb{R}^n} &= \left\| \sum_{i=1}^{m} x_i T(e_i) \right\|_{\mathbb{R}^n} \\
&\leq \sum_{i=1}^{m} \|x_i\|_{\mathbb{R}^n} \|T(e_i)\|_{\mathbb{R}^n} \\
&\leq \sum_{i=1}^{m} c\|x_i\|_{\mathbb{R}^n} = c \sum_{i=1}^{m} |x_i|
\end{aligned}
$$

isto é,

$$\|T(x)\|_{\mathbb{R}^n} \leq c\|x\|_{\mathbb{R}^m}, \ \forall x \in \mathbb{R}^m.$$

Assim, podemos ver que

$$\|T(x) - T(y)\|_{\mathbb{R}^n} = \|T(x-y)\|_{\mathbb{R}^n} \leq c\|x-y\|_{\mathbb{R}^m}, \ \forall x, y \in \mathbb{R}^m.$$

Estas considerações mostram que:

Teorema 5.1 *Toda aplicação linear* $T : \mathbb{R}^m \Rightarrow \mathbb{R}^n$ *é lipschtziana e, consequentemente, contínua.*

Denotamos

$$\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) = \{T : \mathbb{R}^m \longrightarrow \mathbb{R}^n; T \text{ é linear } \}$$

Munindo $\mathcal{L}$ das operações usuais de adição e multiplicação por escalar, isto é,

- $\forall\, T, L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n); (T + L)(x) = T(x) + L(x)$.

- $\forall \lambda \in \mathbb{R},\ \forall T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n); (\lambda T)(x) = \lambda T(x)$,

segue que, $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ com essas operações, é um espaço vetorial.

Consideremos $\beta = \{e_1, \cdots, e_m\}$ e $\gamma = \{v_1, \cdots, v_n\}$ bases do $\mathbb{R}^m$ e do $\mathbb{R}^n$, respectivamente. Então

$$\begin{aligned} T(e_1) &= a_{11}v_1 + \cdots + a_{1n}v_n \\ T(e_2) &= a_{21}v_1 + \cdots + a_{2n}v_n \\ \vdots \quad & \qquad \vdots \\ T(e_m) &= a_{m1}v_1 + \cdots + a_{mn}v_n \end{aligned} \tag{5.1}$$

Daí,

$$\begin{aligned} T(x) &= \sum_{i=1}^{m} T(x_i e_i) \\ &= \sum_{i=1}^{m} x_i T(e_i) \\ &= x_1 T(e_1) + \cdots + x_m T(e_m) \\ &= x_1(a_{11}v_1 + \cdots + a_{1n}v_n) + \cdots + x_m(a_{m1}v_1 + \cdots + a_{mn}v_n) \\ &= (a_{11}x_1 + \cdots + a_{m1}x_m)v_1 + \cdots + (a_{1n}x_1 + \cdots + a_{mn}x_m)v_n \end{aligned}$$

logo

$$[T(x)]_\gamma = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}_{n\times m} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}_{m\times 1}$$

ou seja,

$$[T(x)]_\gamma = A^t \cdot [x]_\beta$$

onde A^t é a transposta da matriz dos coeficientes da equação (5.1).

Com isso, a função $F : \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \longrightarrow M_{n\times m}(\mathbb{R})$ que a cada $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ associa a matriz A^t, isto é, $F(T) = A^t$, é um isomorfismo.

Observação 5.1 *O espaço $\mathcal{L}(\mathbb{R}^m, \mathbb{R}) = (\mathbb{R}^m)^*$ é chamado de espaço dual de $\mathbb{R}^m$. Tal espaço é isomorfo a $M_{1\times m}(\mathbb{R}) = \mathbb{R}^m$.*

O espaço $\mathcal{L}(\mathbb{R}, \mathbb{R}^n)$ é isomorfo a $M_{n\times 1}(\mathbb{R})$.

O espaço $\mathbb{R}^* = \mathcal{L}(\mathbb{R}, \mathbb{R})$ é isomorfo a $M_{1\times 1}(\mathbb{R})$.

5.2 Norma de uma aplicação linear

Dada $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ temos que T é contínua. Então, como $B_1[0] \subset \mathbb{R}^m$ é um conjunto compacto vemos que $\|T(x)\|_{\mathbb{R}^n}$ assume máximo (e mínimo) na bola fechada $B_1[0]$, Portanto, podemos definir a seguinte aplicação:

$$\begin{array}{rcl} \|\cdot\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} : & \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \to & \mathbb{R} \\ & T \longmapsto & \|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} = \max\limits_{x\in B_1[0]} \left\{\|T(x)\|_{\mathbb{R}^n}\right\}. \end{array}$$

É fácil ver que $\|\cdot\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}$ é uma norma em $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$.

Além disso, dado $x \in \mathbb{R}^m, x \neq 0$ e $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, temos que

$$\begin{aligned} \|T(x)\|_{\mathbb{R}^n} &= \left\| T\left(\|x\|_{\mathbb{R}^m} \cdot \frac{x}{\|x\|_{\mathbb{R}^m}} \right) \right\|_{\mathbb{R}^n} \\ &= \|x\|_{\mathbb{R}^m} \cdot \left\| T\left(\frac{x}{\|x\|_{\mathbb{R}^n}} \right) \right\|_{\mathbb{R}^n} \\ &\leq \|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \cdot \|x\|_{\mathbb{R}^m} \end{aligned}$$

logo $\|T(x)\|_{\mathbb{R}^n} \leq \|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \cdot \|x\|_{\mathbb{R}^m},\ \forall x \in \mathbb{R}^m$.

Além disso, se $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k)$ e $L \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^n)$ então $(L \circ T) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ e vale a desigualdade

$$\|L \circ T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \leq \|L\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \cdot \|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}$$

Capítulo 6

Funções reais de m variáveis reais

6.1 Definições

Sabemos que se $f : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ e $a \in I$, f é diferenciável no ponto a quando existe o limite

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

e o número $f'(a)$ é a derivada de f no ponto a.

No que segue, adotaremos a seguinte notação:

$U \subset \mathbb{R}^m$ é um conjunto aberto;

$a \in U$;

$f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ uma função.

Para esta função f não faz sentido a expressão

$$\frac{f(a+h) - f(a)}{h}$$

pois não é possível "dividir por vetores".

Note entretanto que dado um vetor $v \in \mathbb{R}^m$ existe uma reta que passa por a e tem a direção de v, dada por

$$\{a + tv; t \in \mathbb{R}\}$$

Como $a \in U$ e U é aberto, existe $\delta > 0$ tal que $B_\delta(a) \subset U$. Portanto, se $|t| < \epsilon = \dfrac{\delta}{\|v\|}$ então

$$\|(a + tv) - a\| = \|tv\| = |t|\|v\| < \delta$$

Isto implica que $(a + tv) \in B_\delta(a) \subset U$. Logo, tomando

$$\begin{array}{rlll} \lambda: & (-\epsilon, \epsilon) & \longrightarrow & \mathbb{R}^m \\ & t & \longmapsto & \lambda(t) = a + tv \end{array}$$

teremos

isto é, $\lambda(t) \in U, \ \forall t \in (-\epsilon, \epsilon)$ com $\lambda(0) = a$.

Portanto, faz sentido definir a seguinte composição:

$$\begin{array}{rlll} (f \circ \lambda): & (-\epsilon, \epsilon) \subset \mathbb{R} & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & (f \circ \lambda)(t) = f(a + tv) \end{array}$$

Definição 6.1 *Dado $v \in \mathbb{R}^m$, definimos a derivada direcional de f no ponto a, na direção de v como sendo*

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

Equivalentemente,

$$\begin{aligned} \tfrac{\partial f}{\partial v}(a) &= \lim_{t \to 0} \frac{(f \circ \lambda)(t) - (f \circ \lambda)(0)}{t} \\ &= \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} \end{aligned}$$

Definição 6.2 *Seja $\beta = \{e_1, \cdots, e_m\}$ a base canônica do $\mathbb{R}^m$, então a derivada direcional de f num ponto a na direção de e_i é a i-ésima derivada parcial de f no ponto a.*

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

6.2 Três exemplos importantes

Exemplo 6.1 *(Existência das derivadas parciais não implica em continuidade)*

Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por

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

e seja $v = (\alpha, \beta) \in \mathbb{R}^2$. Então, temos

(i) Se $\alpha \neq 0$ e $\beta = 0$

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t\to 0} \frac{f(\alpha t, 0)}{t} = 0.$$

(ii) Se $\alpha = 0$ e $\beta \neq 0$

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t\to 0} \frac{f(0, \beta t)}{t} = 0.$$

(iii) Se $\alpha \neq 0$ e $\beta \neq 0$

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t\to 0} \frac{f(\alpha t, \beta t)}{t} = \lim_{t\to 0} \left[\frac{t^2\alpha\beta}{t^2\alpha^2 + t^2\beta^2} \cdot \frac{1}{t} \right] = \frac{\alpha}{\alpha^2+\beta^2} \lim_{t\to 0} \frac{1}{t} = +\infty$$

(iv) $\nexists \lim\limits_{(x,y)\to(0,0)} f(x,y)$.

De fato, se $y = x$, então

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{x\to 0} \frac{x^2}{2x^2} = \frac{1}{2}$$

e, se $x = 0$ então $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$.

De (i) e (ii) existem as derivadas $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{\partial f}{\partial y}(0,0) = 0$ e de (iv), f não é contínua em $(0,0)$.

Aparentemente isto se deve ao fato de que não existem as derivadas direcionais exceto nas direções dos eixos. Entretanto, considere o próximo exemplo.

Exemplo 6.2 *(Existência de derivadas direcionais não implica continuidade)*

Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} \dfrac{x^2 y}{x^4+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}.$$

Dado $v = (\alpha, \beta) \in \mathbb{R}^2$ então, temos

(i) Se $\alpha \neq 0$ e $\beta = 0$

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t \to 0} \frac{f(\alpha t, 0)}{t} = 0.$$

(ii) Se $\alpha = 0$ e $\beta \neq 0$

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t \to 0} \frac{f(0, \beta t)}{t} = 0.$$

(iii) Se $\alpha \neq 0$ e $\beta \neq 0$

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t \to 0} \frac{f(\alpha t, \beta t)}{t} = \lim_{t \to 0} \left[\frac{\alpha^2 t^2 \beta t}{\alpha^4 t^4 + \beta^2 t^2} \cdot \frac{1}{t} \right] = \lim_{t \to 0} \frac{\alpha^2 \beta}{\alpha^4 t^2 + \beta^2} = \frac{\alpha^2 \beta}{\beta^2} = \frac{\alpha^2}{\beta}$$

Em resumo,

$$\frac{\partial f}{\partial v}(0,0) = \begin{cases} 0, & \text{se } \beta = 0 \\ \dfrac{\alpha^2}{\beta}, & \text{se } \beta \neq 0. \end{cases}$$

Portanto, existem as derivadas direcionais segundo qualquer direção. Entretanto,

$$\nexists \lim_{(x,y) \to (0,0)} f(x,y)$$

De fato, se $y = x^2$ então

$$\lim_{(x,y) \to (0,0)} f(x,y) = \lim_{x \to 0} \frac{x^4}{2x^4} = \frac{1}{2}$$

e, se $y = 0$ então $\displaystyle\lim_{(x,y) \to (0,0)} f(x,y) = 0$.

Exemplo 6.3 *(Existência de derivadas direcionais não implica dependência linear da direção)*

Para cada $v \in \mathbb{R}^m$ podemos associar $\dfrac{\partial f}{\partial v}(a)$ Isto define uma aplicação

$$\begin{array}{rccc} \dfrac{\partial f}{\partial \bullet}(a) : & \mathbb{R}^m & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & \dfrac{\partial f}{\partial v}(a) \end{array}$$

Será que:

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

e

$$\frac{\partial f}{\partial \lambda v}(a) = \lambda \frac{\partial f}{\partial v}(a)$$

Na verdade, sempre temos

$$\begin{aligned}
\frac{\partial f}{\partial(\lambda v)}(a) &= \lim_{t\to 0}\frac{f(a+t\lambda v)-f(a)}{t}\\
\\
&\lim_{t\to 0}\lambda\frac{f(a+t\lambda v)-f(a)}{t\lambda}\\
\\
&= \lambda\lim_{t\to 0}\frac{f(a+(t\lambda)v)-f(a)}{(\lambda t)} = \lambda\frac{\partial f}{\partial v}(a)
\end{aligned}$$

Por outro lado, considere

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

É fácil ver que existem as derivadas direcionais e

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

Com efeito, se $v=(a_1,b_1)$ e $w=(a_2,b_2)$ então $v+w=(a_1+a_2,b_1+b_2)$ e

$$\begin{aligned}
\text{(i)}\quad \frac{\partial f}{\partial(v+w)}(0,0) &= \lim_{t\to 0}\frac{f(t(a_1+a_2),t(b_1+b_2))}{t}\\
\\
&= \lim_{t\to 0}\frac{t^2(a_1+a_2)^2\cdot t(b_1+b_2)}{t^2(a_1+a_2)^2+t^2(b_1+b_2)^2}\cdot\frac{1}{t}\\
\\
&= \lim_{t\to 0}\frac{(a_1+a_2)^2(b_1+b_2)^2}{(a_1+a_2)^2+(b_1+b_2)^2} = \frac{(a_1+a_2)^2(b_1+b_2)^2}{(a_1+a_2)^2+(b_1+b_2)^2}.
\end{aligned}$$

$$\begin{aligned}
\text{(ii)}\quad \frac{\partial f}{\partial v}(0,0) &= \lim_{t\to 0}\frac{f(ta_1,tb_1)}{t}\\
\\
&\lim_{t\to 0}\frac{(t^2a_1^2)(tb_1)}{t^2a_1^2+t^2b_1^2}\cdot\frac{1}{t}\\
\\
&= \frac{a_1^2b_1}{a_1^2+b_1^2}.
\end{aligned}$$

(iii) $$\frac{\partial f}{\partial w}(0,0) = \lim_{t\to 0}\frac{f(ta_2,tb_2)}{t}$$

$$= \lim_{t\to 0}\frac{(t^2a_2^2)(tb_2)}{t^2a_2^2+t^2b_2^2}\cdot\frac{1}{t}$$

$$= \frac{a_2^2b_2}{a_2^2+b_2^2}.$$

Portanto,

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

Teorema 6.1 *(Teorema do Valor Médio)*

Sejam $U \subset \mathbb{R}^m$ um aberto, $a \in U, v \in \mathbb{R}^m$ e $f : U \longrightarrow \mathbb{R}$. Se o segmento de reta $[a,a+v] \subset U, f\big|_{[a,a+v]}$ é contínua e existe $\frac{\partial f}{\partial v}(x)$ para todo $x \in [a,a+v]$, então existe $\theta \in (0,1)$ tal que

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

Demonstração: Seja

$$\begin{array}{rccl} \lambda: & [0,1] & \longrightarrow & U \\ & t & \longmapsto & \lambda(t) = a+tv \end{array}$$

uma parametrização do segmento $[a,a+v]$ e definamos $\xi = (f\circ\lambda) : [0,1] \longrightarrow \mathbb{R}$. Então

(1) ξ é contínua em $[0,1]$;

(2) ξ é derivável em $(0,1)$;

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

$$\xi(1)-\xi(0) = \xi'(\theta).$$

Mas, observe que

$$
\begin{aligned}
\xi'(\theta) &= \lim_{t\to 0}\frac{\xi(\theta+t)-\xi(\theta)}{t}\\
&= \lim_{t\to 0}\frac{f(a+\theta v)-f(a+\theta v)}{t}\\
&= \frac{\partial f}{\partial v}(a+\theta v),\\
\xi(1) &= f(a+v) \quad \text{e} \quad \xi(0)=f(a).
\end{aligned}
$$

Isto conclui a prova do teorema. □

Corolário 6.1 *Sejam $U \subset \mathbb{R}^m$ um conjunto aberto e conexo, $f : U \longrightarrow \mathbb{R}$. Se $\frac{\partial f}{\partial v}(x) = 0$ para todo $x \in U$ e para todo $v \in \mathbb{R}^m$ então f é contante.*

Demonstração: De fato, fixemos $a \in U$. A existência de $\frac{\partial f}{\partial v}$ garante a continuidade da restrição $f\big|_{[a,b]}$ para todo segmento de reta $[a,b] \subset U$. Pelo Teorema do Valor Médio, tem-se que $[a,b] \subset U$ implica $f(b) - f(a) = 0$, isto é, $f(a) = f(b)$. Ora, qualquer ponto $x \in U$ pode ser ligado ao ponto a por uma poligonal contida em U, com vértices $a_0 = a, a_1, \cdots, a_{k-1}, a_k = x$, em virtude da conexidade do aberto U. Assim, temos sucessivamente, $f(a) = f(a_1) = \cdots = f(x)$. Logo $f(x) = f(a)$ para todo $x \in U$, donde f é constante. □

6.3 Funções diferenciáveis

Sejam $I \subset \mathbb{R}, a \in I$ e $f : I \longrightarrow \mathbb{R}$. Por definição f é diferenciável em a quando existe o limite

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h} = f'(a)$$

e, neste caso, $f'(a)$ é a derivada de f no ponto a.

Equivalentemente,

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=f'(a)\Longleftrightarrow \lim_{h\to 0}\left|\frac{f(a+h)-f(a)}{h}-f'(a)\right|=0\Longleftrightarrow$$

$$\lim_{h\to 0}\frac{|f(a+h)-f(a)-f'(a)h|}{|h|}=0.$$

logo, se f é diferenciável em $a\in I$, existe uma transformação linear

$$\begin{array}{rccl} T_{f'(a)}: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & h & \longmapsto & T_{f'(a)}(h)=f'(a)h \end{array}$$

tal que

$$\lim_{h\to 0}\frac{|f(a+h)-f(a)-T_{f'(a)}(h)|}{|h|}=0.$$

Reciprocamente, suponha que exista uma transformação linear $T\in\mathcal{L}(\mathbb{R},\mathbb{R})=\mathbb{R}^*$ tal que

$$\lim_{h\to 0}\frac{|f(a+h)-f(a)-T(h)|}{|h|}=0.$$

Como $T\in\mathbb{R}^*$ vemos que $T(x)=\alpha x$ para algum $\alpha\in\mathbb{R}$. Logo

$$\lim_{h\to 0}\frac{|f(a+h)-f(a)-\alpha h|}{|h|}=0$$

isto é, existe

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=\alpha$$

e, portanto, f é diferenciável em a e $f'(a)=\alpha$.

O que acabamos de fazer demonstra a seguinte proposição:

Proposição 6.1 *Sejam $I\subset\mathbb{R}$ um intervalo aberto, $a\in I$ e $f:I\longrightarrow\mathbb{R}$. Então, f é diferenciável no ponto a se, e somente se, existe $T\in\mathcal{L}(\mathbb{R},\mathbb{R})=\mathbb{R}^*$ tal que*

$$\lim_{h\to 0}\left|\frac{f(a+h)-f(a)-T(h)}{h}\right|=0.$$

No caso afirmativo, a transformação $T=f'(a)$ é a derivada de f no ponto a.

Equivalentemente, f é diferenciável em a se, e somente se, existe $\alpha\in\mathbb{R}$ tal que

$$\lim_{h\to 0}\left|\frac{f(a+h)-f(a)-\alpha h}{h}\right|=0.$$

No caso afirmativo, $\alpha = f'(a)$ é a derivada de f no ponto a.

Se $a \in I$ e I é uma aberto na reta $\mathbb{R}$, existe $\epsilon > 0$ tal que $B_\epsilon(a) \subset I$. Denotamos por $\mathcal{F}(\epsilon, \mathbb{R}) = \{u : (-\epsilon, \epsilon) \longrightarrow \mathbb{R}; u$ é uma função $\}$. Desta forma, dada $T \in \mathcal{L}(\mathbb{R}, \mathbb{R}) = \mathbb{R}^*$ podemos associar a função $r_T \in \mathcal{F}(\epsilon, \mathbb{R})$, denominada função resto (ou erro) relativa a T, dada por

$$r_T(h) = f(a+h) - f(a) - T(h) = f(a+h) - (T(h) + f(a)),\ h \in (-\epsilon, \epsilon).$$

A reta paralela a T passando pelo ponto $(a, f(a))$ é:

$$L(x) = \alpha x + f(a) - \alpha a,$$

onde $T(x) = \alpha x$.

Então

$$L(a+h) = \alpha h + f(a) = T(h) + f(a).$$

Suponha que f é diferenciável em a. Então

$$\lim_{h \to 0} \left| \frac{r_{f'(a)}(h)}{h} \right| = \lim_{h \to 0} \left| \frac{f(a+h) - f(a) - f'(a)h}{h} \right| = 0$$

Reciprocamente, dada $T \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ suponha que

$$\lim_{h \to 0} \left| \frac{r_T(h)}{h} \right| = 0.$$

Pela definição de resto

$$\frac{r_T(h)}{h} = \frac{f(a+h) - f(a) - T(h)}{h}$$

Então,

$$\lim_{h \to 0} \left| \frac{f(a+h) - f(a) - T(h)}{h} \right| = 0$$

Segue que f é diferenciável em a e $T = f'(a)$.

Proposição 6.2 *Sejam $I \subset \mathbb{R}$ um intervalo aberto, $a \in I$ e $f : I \longrightarrow \mathbb{R}$. Então, f é diferenciável no ponto a se, e somente se, existe $T \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ tal que o resto relativo a T satisfaz*

$$\lim_{h \to 0} \frac{r_T(h)}{h} = 0.$$

Definição 6.3 *Sejam $I \subset \mathbb{R}^m$ um aberto, $a \in U$ e $f : U \longrightarrow \mathbb{R}$. Dizemos que f é diferenciável no ponto a quando existir $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}) = (\mathbb{R}^m)^*$ tal que*

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - T(h)|}{\|h\|} = 0$$

Observamos que, se f é diferenciável em a, a transformação T satisfazendo a definição acima é única.

De fato, sejam $T_1, T_2 \in \mathcal{L}(\mathbb{R}^m, \mathbb{R})$ tais que

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - T_1(h)|}{\|h\|} = 0$$

e

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - T_2(h)|}{\|h\|} = 0.$$

Então, note que

$$\begin{aligned} \frac{|T_1(h) - T_2(h)|}{\|h\|} &= \frac{|T_1(h) - (f(a+h) - f(a)) + f(a+h) - f(a) - T_2(h)|}{\|h\|} \\ &\leq \frac{|f(a+h) - f(a) - T_1(h)|}{\|h\|} + \frac{f(a+h) - f(a) - T_2(h)|}{\|h\|} \end{aligned}$$

Tomando o limite com $h \to 0$ temos

$$\lim_{h \to 0} \frac{|T_1(h) - T_2(h)|}{\|h\|} = 0,$$

isto é, para todo $\epsilon > 0$, existe $\delta > 0$ tal que $0 < \|h\| < \delta$ então

$$|T_1(h) - T_2(h)| < \epsilon \|h\|$$

Dado $x \in \mathbb{R}^m$ podemos escrever $x = \frac{2\|x\|}{\delta} \frac{\delta}{2} \frac{x}{\|x\|}$. Então

$$\begin{aligned} |T_1(x) - T_2(x)| &= \left| T_1\left(\frac{2\|x\|}{\delta} \frac{\delta}{2} \frac{x}{\|x\|}\right) - T_2\left(\frac{2\|x\|}{\delta} \frac{\delta}{2} \frac{x}{\|x\|}\right) \right| \\ &= 2\frac{\|x\|}{\delta} \left| T_1\left(\frac{\delta}{2} \frac{x}{\|x\|}\right) - T_2\left(\frac{\delta}{2} \frac{x}{\|x\|}\right) \right| \quad \text{mas } \left\| \frac{\delta}{2} \frac{x}{\|x\|} \right\| = \frac{\delta}{2} \\ &< \frac{2\|x\|}{\delta} \epsilon \frac{\delta}{2} = \|x\| \epsilon \end{aligned}$$

Assim, para todo $\epsilon > 0$

$$|T_1(x) - T_2(x)| < \epsilon\|x\|, \;\; \forall x \in \mathbb{R}^m$$

e, pela norma da transformação linear, temos

$$0 \leq \|T_1 - T_2\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R})} = \max_{x \in B_1[0]} |T_1(x) - T_2(x)| < \epsilon$$

Como ϵ é arbitrário, temos que $T_1 - T_2 = 0$, ou seja, $T_1 = T_2$.

Definição 6.4 *A transformação linear T da definição anterior é a derivada de f no ponto a denotada por $df(a)$.*

Analogamente, dada $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R})$ podemos definir

$$r_T(h) = f(a+h) - f(a) - T(h).$$

Teorema 6.2 *Sejam $U \subset \mathbb{R}^m$ aberto, $a \in U$ e $f : U \to \mathbb{R}$. Então f é diferenciável no ponto a se, e somente se, existe $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R})$ tal que o resto associado a T satisfaz*

$$\lim_{h \to 0} \frac{r_T(h)}{\|h\|} = 0.$$

Demonstração: Suponhamos f diferenciável no ponto $a \in U$. Isto, por definição, quer dizer que existe $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R})$ tal que

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - T(h)|}{\|h\|} = 0$$

isto é,

$$\lim_{h \to 0} \frac{|r_T(h)|}{\|h\|} = 0$$

□

Teorema 6.3 *Sejam $U \subset \mathbb{R}^m$ aberto e $a \in U$. Se $f : U \to \mathbb{R}$ é diferenciável em a então:*

(i) *f é contínua em a.*

(ii) *f possui derivada direcional $\dfrac{\partial f}{\partial v}(a)$ segundo qualquer direção $v \in \mathbb{R}^m$ e $\dfrac{\partial f}{\partial v}(a) = df(a)(v)$.*

(iii) *A derivada de f no ponto a é*

$$df(a) = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_m}\right)$$

isto é,

$$\begin{aligned} df(a): \quad \mathbb{R}^m &\longrightarrow \mathbb{R} \\ v &\longmapsto df(a)(v) = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_m}\right) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \\ &= \sum_{j=1}^{m} \frac{\partial f(a)}{\partial x_j} v_j. \end{aligned}$$

(iv) *As derivadas direcionais $\dfrac{\partial f(a)}{\partial v}$ dependem linearmente de v.*

Demonstração:

(i) Como f é diferenciável em a, dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$|f(a+h) - f(a) - (df(a))(h)| \le \epsilon \|h\|_{\mathbb{R}^m}, \quad \text{se} \quad \|h\|_{\mathbb{R}^m} < \delta$$

Em particular, se $\epsilon = 1$ existe $\delta > 0$ tal que

$$|f(a+h) - f(a) - (df(a))(h)| \le \|h\|_{\mathbb{R}^m}, \quad \text{se} \quad \|h\|_{\mathbb{R}^m} < \delta$$

Logo,

$$\begin{aligned} |f(a+h) - f(a)| &= |f(a+h) - f(a) - df(a)(h) + df(a)(h)| \\ &\le |f(a+h) - f(a) - df(a)(h)| + |df(a)(h)| \\ &\le \|h\|_{\mathbb{R}^m} + c\|h\|_{\mathbb{R}^m} = (c+1)\|h\|_{\mathbb{R}^m} \\ &= k\|h\|_{\mathbb{R}^m} \end{aligned}$$

isto é, se $\|h\|_{\mathbb{R}^m} < \delta$ então $|f(a+h) - f(a)| \le k\|h\|_{\mathbb{R}^m}$. Segue que f é contínua no ponto a.

(ii) Como f é diferenciável em a então, para todo $\epsilon > 0$, existe $\delta > 0$ tal que, se $0 < \|h\|_{\mathbb{R}^m} < \delta$ então

$$|f(a+h) - f(a) - (df(a))(h)| \leq \epsilon \|h\|_{\mathbb{R}^m}.$$

Logo, dado $v \in \mathbb{R}^m, v \neq 0$ tomamos $h = tv$ para $0 < |t| < \dfrac{\delta}{\|v\|_{\mathbb{R}^m}}$, ou seja, para $0 < \|h\|_{\mathbb{R}^m} = |t|\|v\|_{\mathbb{R}^m} < \delta$, então

$$|f(a+h) - f(a) - (df(a))(tv)| \leq \epsilon |t| \|v\|_{\mathbb{R}^m}$$

donde

$$\left| \frac{f(a+h) - f(a)}{t} - (df(a))(v) \right| \leq \epsilon \|v\|_{\mathbb{R}^m}$$

Em resumo, para todo $\epsilon > 0$, existe $\eta > 0$ $\left(\eta = \dfrac{\delta}{\|v\|}\right)$ tal que se $0 < |t| < \eta$ então

$$\left| \frac{f(a+tv) - f(a)}{t} - df(a)(v) \right| < \|v\|_{\mathbb{R}^m} \epsilon$$

e isto implica que

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

por sua vez, isto implica que existe

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

(iii) Sabemos que $\mathcal{L}(\mathbb{R}^m, \mathbb{R})$ é isomorfo a $M_{1 \times m}(\mathbb{R})$ (o conjunto das matrizes $1 \times m$ com entradas reais) e que $df(a) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R})$. Daí,

$$\begin{array}{rcl} (df(a))(e_1) & = & A_1 \\ (df(a))(e_2) & = & A_2 \\ \vdots & \vdots & \vdots \\ (df(a))(e_m) & = & A_m \end{array}$$

Pelo item anterior

$$\begin{array}{rcc} A_1 = & \dfrac{\partial f}{\partial e_1}(a) = & \dfrac{\partial f}{\partial x_1}(a) \\ \vdots & \vdots & \vdots \\ A_m = & \dfrac{\partial f}{\partial e_m}(a) = & \dfrac{\partial f}{\partial x_m}(a) \end{array}$$

Logo, $df(a) = \left(\dfrac{\partial f}{\partial x_1}(a), \cdots, \dfrac{\partial f}{\partial x_m}(a)\right)_{1\times m}$ e

$$(df(a))(v) = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_m}(a)\right)_{1\times m} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \sum_{j=1}^{m} \frac{\partial f(a)}{\partial x_j} v_j.$$

(iv) Imediato de (ii) pois $df(a)$ é linear. $\square$

Teorema 6.4 *(Condição de diferenciabilidade)*

Sejam $U \subset \mathbb{R}^m$ aberto e $a \in U$. Se $f : U \longrightarrow \mathbb{R}$ é tal que existem as derivadas parciais $\dfrac{\partial f}{\partial x_j}$ numa vizinhança de a e estas são contínuas no ponto a, então f é diferenciável em a.

Demonstração: Seja $\eta > 0$ tal que $\dfrac{\partial f}{\partial x_j}(x)$ existe, para todo $x \in B_\eta(a)$, $\forall j = 1, \cdots, m$. Então, dado $\epsilon > 0$, existe $\delta > 0 (0 < \delta < \eta)$ tal que se $\|\xi\|_{\mathbb{R}^m} < \delta$ então

$$\left| \frac{\partial f}{\partial x_j}(a+\xi) - \frac{\partial f}{\partial x_j}(a) \right| < \epsilon$$

Agora, seja $h \in \mathbb{R}^m$ tal que $\|h\| < \delta$. Considere

$$\begin{aligned} p^{(0)} &= (a_1, a_2, a_3, \cdots, a_m) \\ p^{(1)} &= (a_1 + h_1, a_2, a_3, \cdots, a_m) \\ p^{(2)} &= (a_1 + h_1, a_2 + h_2, a_3, \cdots, a_m) \\ &\vdots \\ p^{(m)} &= (a_1 + h_1, a_2 + h_2, a_3 + h_3, \cdots, a_m + h_m) = a + h \end{aligned}$$

Então,

$$\begin{aligned} f(a+h) - f(a) &= f(p^{(m)}) - f(p^{(0)}) \\ &= f(p^{(m)}) - f(p^{(m-1)}) + f(p^{(m-1)} - \cdots - f(p^{(1)}) + f(p^{(1)}) - f(p^{(0)}) \\ &= \sum_{j=1}^{m} \left[f(p^{(j)}) - f(p^{(j-1)}) \right] \end{aligned}$$

Note que $\left[p^{(j-1)}, p^{(j)}\right]$ é um segmento paralelo ao j-ésimo eixo, ou seja,

$$\left[p^{(j-1)}, p^{(j)}\right] = \left\{p^{(j-1)} + t(h_j e_j),\ t \in [0,1]\right\}$$

onde $\{e_1, \cdots, e_m\}$ é a base canônica do $\mathbb{R}^m$. Pelo Teorema do Valor Médio, existe $\theta_j \in (0,1)$ tal que

$$\begin{aligned}
f(p^{(j)}) - f(p^{(j-1)}) &= \frac{\partial f}{\partial (h_j e_j)}(p^{(j-1)} + \theta_j(h_j e_j)) \\
&= h_j \frac{\partial f}{\partial e_j}(p^{(j-1)} + \theta_j(h_j e_j)) \\
&= h_j \frac{\partial f}{\partial x_j}(p^{(j-1)} + \theta_j(h_j e_j))
\end{aligned}$$

Portanto

$$f(a+h) - f(a) = \sum_{j=1}^{m} h_j \frac{\partial f}{\partial x_j}(p^{(j-1)} + \theta_j(h_j e_j))$$

donde

$$\begin{aligned}
\left| f(a+h) - f(a) - \sum_{j=1}^{m} h_j \frac{\partial f(a)}{\partial x_j} \right| &= \left| \sum_{j=1}^{m} \left(\frac{\partial f}{\partial x_j}(p^{(j-1)} + \theta_j(h_j e_j)) - \frac{\partial f(a)}{\partial x_j} \right) h_j \right| \\
&\leq \sum_{j=1}^{m} \left| \frac{\partial f}{\partial x_j}(p^{(j-1)} + \theta_j(h_j e_j)) - \frac{\partial f(a)}{\partial x_j} \right| |h_j| \\
&< \epsilon \sum_{j=1}^{m} |h_j| = \epsilon \|h\|_{S,\mathbb{R}^m} < m\epsilon \|h\|_{\mathbb{R}^m}
\end{aligned}$$

Resumindo, temos que para todo $\epsilon > 0$ existe $\delta > 0$ tal que se $0 < \|h\|_{\mathbb{R}^m} < \delta$ então

$$\frac{\left| f(a+h) - f(a) - \sum_{j=1}^{m} \frac{\partial f(a)}{\partial x_j} h_j \right|}{\|h\|_{\mathbb{R}^m}} < m\epsilon$$

isto é,

$$\lim_{h \to 0} \frac{\left| f(a+h) - f(a) - \sum_{j=1}^{m} \frac{\partial f(a)}{\partial x_j} h_j \right|}{\|h\|_{\mathbb{R}^m}} = 0,$$

o que prova que f é diferenciável em a e $df(a)(h) = \displaystyle\sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(a) h_j$. □

Observação 6.1 *Toda função de classe C^1 é diferenciável.*

6.3.1 Gradiente

Para cada $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}) = (\mathbb{R}^m)^*$ existe um único $v_T \in \mathbb{R}^m$ tal que

$$T(u) = \langle v_T, u \rangle.$$

Definição 6.5 *Seja $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ uma função diferenciável no ponto a do aberto U. Então definimos o gradiente de f no ponto a, denotado por $\nabla f(a)$, como sendo o vetor do $\mathbb{R}^m$ tal que*

$$df(a)(u) = \langle \nabla f(a), u \rangle, \ \forall u \in \mathbb{R}^m.$$

Teorema 6.5 *Seja $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ diferenciável no ponto a do aberto U. Então*

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

6.4 Diferenciabilidade de funções $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$

Definição 6.6 *Sejam $U \subset \mathbb{R}^m$ um conjunto aberto, $a \in U$ e $f = (f_1, \cdots, f_n) : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$. Dizemos que f é diferenciável no ponto a quando existe $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ tal que*

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - T(h)\|_{\mathbb{R}^n}}{\|h\|_{\mathbb{R}^m}} = 0$$

No caso afirmativo, a aplicação T é única e é denominada a derivada de f no ponto a.

Denota-se $T = df(a)$.

Para cada $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ podemos definir a função resto associada a T

$$\begin{array}{rccl} r_T : & B_\epsilon(0) \subset \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ & h & \longmapsto & r_T(h) = f(a+h) - f(a) - T(h) \end{array}$$

Aqui $\epsilon > 0$ é tal que $B_\epsilon(a) \subset U$.

Teorema 6.6 *Sejam $U \subset \mathbb{R}^m$ um aberto e $a \in U$. Uma função $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é diferenciável em a se, e somente se, existe $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ tal que o resto r_T satisfaz*

$$\lim_{h \to 0} \frac{r_T(h)}{\|h\|} = 0.$$

Demonstração: Imediata da definição.

Definição 6.7 *Sejam $U \subset \mathbb{R}^m$ um aberto, $a \in U$, $v \in \mathbb{R}^m$ e $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$. A derivada direcional de f no ponto a, na direção do vetor v, é o vetor $\dfrac{\partial f}{\partial v}(a) \in \mathbb{R}^n$ dado pelo limite*

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

quando o limite existir.

Equivalentemente, um vetor $w \in \mathbb{R}^n$ é a derivada direcional de f no ponto a, na direção do vetor v, quando para todo $\epsilon > 0$, existir $\delta > 0$ tal que se $0 < |t| < \delta$ então

$$\left\| \frac{f(a + tv) - f(a)}{t} - w \right\|_{\mathbb{R}^n} < \epsilon.$$

Teorema 6.7 *Sejam $U \subset \mathbb{R}^m$ um aberto, $a \in U$ e $f = (f_1, \cdots, f_n) : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$.*

(a) *Se f é diferenciável em a, então f é contínua.*

(b) *Se f é diferenciável em a então existe $\dfrac{\partial f}{\partial v}(a)$ segundo qualquer direção $v \in \mathbb{R}^m$ e, além disso, $\dfrac{\partial f}{\partial v}(a) = df(a)(v)$.*

(c) *f é diferenciável em a se, e somente se, para cada $i = 1, 2, \cdots, n$ tem-se que*

$f_i : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ *é diferenciável em* a. *No caso afirmativo temos*

$$
\begin{array}{rccl}
df(a): & \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\
& v & \longmapsto & df(a)(v) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \cdots & \dfrac{\partial f_n}{\partial x_m}(a) \end{bmatrix}_{n\times m} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}_{m\times 1} \\
& & & = \begin{pmatrix} \displaystyle\sum_{j=1}^{m} \frac{\partial f_1}{\partial x_j}(a)v_j \\ \vdots \\ \displaystyle\sum_{j=1}^{m} \frac{\partial f_n}{\partial x_j}(a))v_j \end{pmatrix}_{n\times 1} .
\end{array}
$$

Demonstração:

(a) Desde que f é diferenciável em a, dado $\epsilon > 0$ existe $\delta > 0$ tal que se $0 < \|h\| < \delta$ então

$$\|f(a+h) - f(a) - df(a)(h)\| < \epsilon\|h\|.$$

Para $\epsilon = 1$, existe $\delta > 0$ tal que

$$\|f(a+h) - f(a) - df(a)(h)\|_{\mathbb{R}^n} < \epsilon\|h\|_{\mathbb{R}^m}.$$

Então

$$
\begin{aligned}
\|f(a+h) - f(a)\|_{\mathbb{R}^n} &= \|f(a+h) - f(a) - df(a)(h) + d(f(a)(h)\|_{\mathbb{R}^n} \\
&\leq \|f(a+h) - f(a) - df(a)(h)\|_{\mathbb{R}^n} + \|d(f(a)(h)\|_{\mathbb{R}^n} \\
&\leq (1+c)\|h\|_{\mathbb{R}^m}
\end{aligned}
$$

se $0 < \|h\|_{\mathbb{R}^m} < \delta$. Isto mostra que f é contínua no ponto a.

(b) Suponha f diferenciável em a. Então dado $\epsilon > 0$ existe $\delta > 0$ tal que se $0 < \|h\| < \delta$ segue que

$$\|f(a+h) - f(a) - df(a)(h)\| < \epsilon\|h\|$$

Dado $v \in \mathbb{R}^m, v \neq 0$, seja $h = tv$ com $0 < |t| < \dfrac{\delta}{\|v\|}$. Assim, $0 < \|h\| = |t|\|v\| < \delta$ e, então

$$\|f(a+tv) - f(a) - tdf(a)(v)\| < \epsilon|t|\|v\|$$

ou seja,

$$\left\|\frac{f(a+tv) - f(a)}{t} - df(a)(v)\right\| < \epsilon\|v\|.$$

Logo, existe

$$\lim_{t\to 0}\left(\frac{f(a+tv) - f(a)}{t}\right) = df(a)(v) = \frac{\partial f}{\partial v}(a).$$

(c) Suponha que $f = (f_1, \cdots, f_n) : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é diferenciável no ponto a. Então dado $\epsilon > 0$ existe $\delta > 0$ tal que se $0 < \|h\| < \delta$ então

$$\|f(a+h) - f(a) - df(a)(h)\|_{\mathbb{R}^n} < \epsilon\|h\|_{\mathbb{R}^m}.$$

Agora, note que, fixado $i \in \{1, \cdots, n\}$ temos

$$\begin{aligned} |f_i(a+h) - f_i(a) - df(a)_i(h)| &\leq \max_{1\leq k\leq n}\{|f_k(a+h) - f_k(a) - df(a)_k|\} \\ &= \|f(a+h) - f(a) - df(a)(h)\|_{M,\mathbb{R}^n} \\ &\leq \|f(a+h) - f(a) - df(a)(h)\|_{\mathbb{R}^n} \\ &< \epsilon\|h\|_{\mathbb{R}^m} \end{aligned}$$

Em resumo, para todo $\epsilon > 0$ existe $\delta > 0$ tal que se $0 < \|h\| < \delta$ então

$$|f_i(a+h) - f_i(a) - df(a)_i(h)| < \epsilon\|h\|_{\mathbb{R}^m}$$

com $df(a)_i \in \mathcal{L}(\mathbb{R}^m, \mathbb{R})$. Logo $f_i : U \longrightarrow \mathbb{R}$ é diferenciável em a e, além disso, $df_i(a) = df(a)_i$ ou seja,

$$\begin{aligned} df_i(a) : \quad \mathbb{R}^m &\longrightarrow \mathbb{R} \\ v &\longmapsto df_i(a)(v) = \left(\frac{\partial f_i}{\partial x_1}(a), \cdots, \frac{\partial f_i}{\partial x_m}(a)\right)\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \\ &\qquad\qquad = \sum_{j=1}^{m}\frac{\partial f_i(a)}{\partial v_j}v_j \end{aligned}$$

Reciprocamente, suponhamos que para cada $i = 1, \cdots, n$, $f_i : U \longrightarrow \mathbb{R}$ é diferenciável. Dado $\epsilon > 0$ existe $\delta > 0$ tal que se $0 < \|h\| < \delta$ então

$$|f_i(a+h) - f_i(a) - df_i(a)(h)| < \epsilon\|h\|, \ \forall i = 1, \cdots, n$$

Definimos $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ onde $T_i = df_i(a)$ para $i = 1, \cdots, n$. É claro que $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ e

$$T(v) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \cdots & \dfrac{\partial f_n}{\partial x_m}(a) \end{pmatrix}_{n\times m} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}_{m\times 1}$$

Logo

$$\begin{aligned} \|f(a+h) - f(a) - T(h)\|_{\mathbb{R}^n} &\le n \max_{1\le i\le n}\{|f_i(a+h) - f_i(a) - T_i(h)|\} \\ &\le n\epsilon\|h\|_{\mathbb{R}^m} \end{aligned}$$

se $0 < \|h\| < \delta$, isto é, existe $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ tal que

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - T(h)\|_{\mathbb{R}^n}}{\|h\|_{\mathbb{R}^m}} = 0$$

. Logo f é diferenciável em a e $df(a) = T$. $\square$

Observação 6.2 *A matriz* $\left(\dfrac{\partial f_i}{\partial x_j}\right)_{n\times m}$, $1 \le i \le n, 1 \le j \le m$, *é denominada matriz jacobiana de* f *no ponto* a. *Quando* $n = m$ *o determinante da matriz jacobiana é denominado o Jacobiano de* f *no ponto* a.

6.4.1 Regra da Cadeia

Teorema 6.8 *Sejam* $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^\nu$ *e* $g : V \subset \mathbb{R}^\nu \longrightarrow \mathbb{R}^n$ *tais que* $f(U) \subset V$ *(* U *e* V *são abertos). Se* f *é diferenciável no ponto* $a \in U$ *e* g *é diferenciável no ponto* $b = f(a)$ *então* $(g \circ f) : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ *é diferenciável no ponto* a *e*

$$d(g \circ f)(a) = (dg(b) \circ df(a))$$

Demonstração: Como $df(a) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\nu)$ e $dg(a) \in \mathcal{L}(\mathbb{R}^\nu, \mathbb{R}^n)$ temos que $(dg(b) \circ df(a)) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. Além disso,

$$\|(g\circ f)(a+h)-(g\circ f)(a)-(dg(b)\circ df(a))(h)\|_{\mathbb{R}^n}=$$

$$=\|(g\circ f)(a+h)-(g\circ f)(a)+dg(b)(f(a+h)-f(a))-dg(b)(f(a+h)-f(a))-(dg(b)(df(a)(h))\|$$

$$=\|g(f(a+h))-g(f(a))-dg(b)(f(a+h)-f(a))+dg(b)(f(a+h)-f(a)-df(a)(h))\|$$

$$\leq\|g(f(a+h))-g(b)-dg(b)(f(a+h)-f(a))\|+\|dg(b)(f(a+h)-f(a)-df(a)(h))\|$$

$$\leq c_1\|f(a+h)-f(a)-df(a)(h)\| \tag{6.1}$$

Agora, como f é diferenciável no ponto a e g é diferenciável no ponto $b=f(a)$, então dado $\epsilon>0$ existe $\delta_1>0$ e $\delta_2>0$ tais que

$$\|f(a+h)-f(a)-df(a)(h)\|_{\mathbb{R}^\nu}\leq\epsilon\|h\|_{\mathbb{R}^m},\quad \text{se } \|h\|<\delta_1 \tag{6.2}$$

e

$$\|g(b+\xi)-g(b)-dg(b)(\xi)\|_{\mathbb{R}^n}\leq\epsilon\|\xi\|_{\mathbb{R}^\mu},\quad \text{se } \|\xi\|<\delta_2 \tag{6.3}$$

Além disso, sabemos que existem $\delta_3>0$ e $c_2>0$ tais que

$$\|f(a+h)-f(a)\|<c_2\|h\|\quad \text{se } \|h\|<\delta_3 \tag{6.4}$$

Escolhendo $\delta=\min\{\delta_1,\dfrac{\delta_2}{c_2},\delta_3\}$ teremos

$$\text{se } \|h\|<\delta\Rightarrow\|f(a+h)-f(a)\|_{\mathbb{R}^\nu}\leq c_2\|h\|_{\mathbb{R}^m}<c_2\frac{\delta_2}{c_2}=\delta_2\Rightarrow$$

$$\|g\left(b+f(a+h)-f(a)\right)-g(b)-dg(b)\left(f(a+h)-f(a)\right)\|_{\mathbb{R}^n}\leq\epsilon\|f(a+h)-f(a)\|_{\mathbb{R}^\nu}$$

ou seja

$$\|g\left(b+f(a+h)-f(a)\right)-g(b)-dg(b)\left(f(a+h)-f(a)\right)\|_{\mathbb{R}^n}\leq\epsilon c_2\|h\|_{\mathbb{R}^\nu}. \tag{6.5}$$

Logo,

$$\begin{aligned}\|(g\circ f)(a+h)-(g\circ f)(a)-\left(dg(b)\circ df(a)\right)(h)\|_{\mathbb{R}^n} &\leq\epsilon c_2\|h\|_{\mathbb{R}^m}+\epsilon\|h\|_{\mathbb{R}^m}\\ &=\epsilon(c_2+1)\|h\|_{\mathbb{R}^m},\quad \text{se } \|h\|<\delta\end{aligned}$$

Resumindo, para todo $\epsilon > 0$ existe $\delta > 0$ tal que se $\|h\| < \delta$ então

$$\|(g \circ f)(a+h) - (g \circ f)(a) - (dg(b) \circ df(a))\,(h)\|_{\mathbb{R}^n} \leq \epsilon(c_2+1)\|h\|_{\mathbb{R}^m}, \quad \text{se } \|h\| < \delta.$$

$\square$

Exemplo 6.4 *Seja* $f : [0, 2\pi] \longrightarrow \mathbb{R}^2$ *dada por* $f(t) = (\cos t, \operatorname{sen} t)$*. É claro que* f *é diferenciável em* $(0, 2\pi)$ *e*

$$\begin{array}{rcccl} df(t_0): & \mathbb{R} & \longrightarrow & \mathbb{R}^2 & \\ & t & \longmapsto & df(t_0)(t) & = \begin{pmatrix} -\operatorname{sen} t_0 \\ \cos t_0 \end{pmatrix}_{2\times 1} (t)_{1\times 1} \\ & & & & = \begin{pmatrix} -t \operatorname{sen} t_0 \\ t \cos t_0 \end{pmatrix} \end{array}$$

isto é,

$$df(t_0)(t) = (-t \operatorname{sen} t_0, t \cos t_0)$$

ou ainda,

$$df(t_0) = (-\operatorname{sen} t_0, \cos t_0).$$

Note ainda que

$$\|df(t_0)\|_{\mathcal{L}(\mathbb{R},\mathbb{R}^2)} = \max_{|t|\leq 1} \|df(t_0)(t)\|_{\mathbb{R}^2} = \max_{|t|\leq 1} |t| = 1$$

e

$$\|df(t_0)\|_{\mathbb{R}^2} = \|(-\operatorname{sen} t_0, \cos t_0)\|_{\mathbb{R}^2} = 1$$

logo $df(t_0) \neq 0,\ \forall t_0 \in (0, 2\pi)$*. Entretanto,*

$$f(2\pi) - f(0) = 0.$$

Segue que $f(2\pi) - f(0) \neq f'(t_0)(2\pi)$*.*

Portanto, não pode haver uma fórmula análoga ao teorema do valor médio quando $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n,\ n > 1$*.*

Teorema 6.9 *(Desigualdade do valor médio para funções vetoriais)*

Seja $f : [a,b] \subset \mathbb{R} \longrightarrow \mathbb{R}^n$ *uma função contínua em* $[a,b]$ *e diferenciável em* (a,b). *Então, existe* $x \in (a,b)$ *tal que*

$$\|f(b) - f(a)\|_{\mathbb{R}^n} \leq \|df(x)\|_{\mathcal{L}(\mathbb{R},\mathbb{R}^n)}(b-a),$$

ou, equivalentemente

$$\|f(a+u) - f(a)\|_{\mathbb{R}^n} \leq \|df(a+\theta u)\|_{\mathcal{L}(\mathbb{R},\mathbb{R}^n)}|u|.$$

Demonstração: Considere $v = (f(b) - f(a))$. Temos que $0 \neq v \in \mathbb{R}^n$. Definamos

$$\begin{array}{rccl} \varphi: & [a,b] & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & \varphi(t) = \langle v, f(t)\rangle_{\mathbb{R}^n} \end{array}$$

Observe que φ é contínua em $[a,b]$ e diferenciável em (a,b). Pelo Teorema do Valor Médio, existe $x \in (a,b)$ tal que

$$|\varphi(b) - \varphi(a)| = \varphi'(x)(b-a),$$

isto é,

$$|\varphi(b) - \varphi(a)| = \langle v, f'(x)\rangle(b-a) \leq \|v\|\|f'(x)\|(b-a).$$

Por outro lado,

$$\begin{aligned} |\varphi(b) - \varphi(a)| = \langle v, f(b), \rangle - \langle v, f(a)\rangle \\ &= \langle v, f(b) - f(a)\rangle \\ &= \langle v, v\rangle = \|v\|^2. \end{aligned}$$

Logo

$$\|v\|^2 \leq \|v\|\|f'(x)\|(b-a).$$

Como $v \neq 0$, então, da desigualdade acima resulta que

$$\|v\| \leq \|f'(x)\|(b-a)$$

que, por sua vez, implica que

$$\|f(b) - f(a)\| \leq \|f'(x)\|_{\mathbb{R}^n}(b-a)$$

ou, ainda,

$$\|f(b)-f(a)\| \leq \|df(x)\|_{\mathcal{L}(\mathbb{R},\mathbb{R}^n)}(b-a).$$

$\square$

Teorema 6.10 *Seja $f: U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ uma função diferenciável no conjunto aberto e convexo U. Se $a, b \in U$ então existe um ponto c no segmento (a,b) tal que*

$$\|f(b)-f(a)\|_{\mathbb{R}^n} \leq \|df(c)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|b-a\|_{\mathbb{R}^m}.$$

Demonstração: A função $\lambda : [0,1] \longrightarrow \mathbb{R}^m$ que a cada $t \in [0,1]$ associa $\lambda(t) = a+t(b-a)$ é uma parametrização do segmento retilíneo $[a,b]$ com $\lambda(0) = a$ e $\lambda(1) = b$. Além disso, $\lambda([0,1]) \subset U$ pois U é convexo.

Tomando $(f \circ \lambda) : [0,1] \subset \mathbb{R} \longrightarrow \mathbb{R}^n$ segue que $f \circ \lambda$ é contínua em $[0,1]$ e diferenciável em $(0,1)$. Pelo teorema anterior, existe $x \in (0,1)$ tal que

$$\|(f \circ \lambda(1) - (f \circ \lambda)(0)\|_{\mathbb{R}^n} \leq \|d(f \circ \lambda)(x)\|_{\mathcal{L}(\mathbb{R},\mathbb{R}^n)}.$$

Daí,

$$\|f(b)-f(a)\|_{\mathbb{R}^n} \leq \|d(f \circ \lambda)(x)\|_{\mathcal{L}(\mathbb{R},\mathbb{R}^n)}$$

Agora, observamos que

$$\begin{aligned}
(d(f \circ \lambda)(x))(v) &= (df(\lambda(x)) \circ d\lambda(x))v \\
&= [df(\lambda(x))(d\lambda(x)v) \\
&= [df(\lambda(x))](v(b-a)), \ \forall v \in \mathbb{R}.
\end{aligned}$$

Então,

$$\begin{aligned}
\|d(f \circ \lambda)(x)\| &= \max_{|v|\leq 1} \|d(f \circ \lambda)(x)(v)\|_{\mathbb{R}^m} \\
&= \max_{|v|\leq 1} \|df(\lambda(x))(v(b-a))\|_{\mathbb{R}}^n \\
&\leq \max_{|v|\leq 1} \|df(\lambda(x))\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}|v|\|b-a\|_{\mathbb{R}^m} \\
&\leq \|f(\lambda(x))\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|b-a\|.
\end{aligned}$$

Como $x \in (0,1)$ e $\lambda(x) = c \in (a,b)$, segue que

$$\|f(b) - f(a)\|_{\mathbb{R}^m} \leq \|df(c)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \|b - a\|_{\mathbb{R}^m}$$

$\square$

6.5 Funções Continuamente Diferenciáveis e Derivadas de Ordem Dois

Definição 6.8 *Uma função $f = (f_1, \cdots, f_n) : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é dita continuamente diferenciável no aberto U (ou de classe C^1 em U) quando para todo $i = 1, \cdots, n$ e $j = 1, \cdots, m$ existe $\dfrac{\partial f_i}{\partial x_j}$ em todo ponto de U e $\dfrac{\partial f_i}{\partial x_j} : U \longrightarrow \mathbb{R}$ são contínuas.*

Observação 6.3 $f = (f_1, \cdots, f_n) \in C^1(U, \mathbb{R}^n) \Leftrightarrow f_i \in C^1(U, \mathbb{R}),\ \forall i = 1, \cdots, n.$

Se $f = (f_1, \cdots, f_n) : U \longrightarrow \mathbb{R}^n$ é de classe C^1, então f é diferenciável em todo ponto U. Desta forma, fica definida a aplicação

$$\begin{array}{rccl} df : & U & \longrightarrow & \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ & x & \longmapsto & df(x). \end{array}$$

Teorema 6.11 *A função $f = (f_1, \cdots, f_n) : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é de classe C^1 se, e somente se, f é diferenciável em U e $df : U \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é contínua.*

Demonstração: Suponha que $f : U \longrightarrow \mathbb{R}^n$ é de classe C^1. Então, existe $\dfrac{\partial f_i}{\partial x_j}$ em todo ponto de U e $\dfrac{\partial f_i}{\partial x_j} : U \longrightarrow \mathbb{R}$ é contínua para todo i e todo j. Isto e os teoremas 6.4 e 6.7 implicam que f é diferenciável em U. Além disso, dado arbitrariamente $a \in U$ e $\epsilon > 0$, existe $\delta > 0$ tal que, se $\|x - a\|_{\mathbb{R}^m} < \delta$ então

$$\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) \right| < \frac{\epsilon}{nm},\ \ \forall i, j$$

Observe que,

$$df(a)(v) = \left(\sum_{j=1}^{m} \frac{\partial f_1}{\partial x_j}(a) v_j, \cdots, \sum_{j=1}^{m} \frac{\partial f_n}{\partial x_j}(a) v_j \right);$$

$$df(x)(v) = \left(\sum_{j=1}^{m} \frac{\partial f_1}{\partial x_j}(x)v_j, \cdots, \sum_{j=1}^{m} \frac{\partial f_n}{\partial x_j}(x)v_j\right);$$

logo

$$df(x)(v) - df(a)(v) = \left(\sum_{j=1}^{m} \left(\frac{\partial f_1}{\partial x_j}(x) - \frac{\partial f_1}{\partial x_j}(a)\right) v_j, \cdots, \sum_{j=1}^{m} \left(\frac{\partial f_n}{\partial x_j}(x) - \frac{\partial f_n}{\partial x_j}(a)\right) v_j\right);$$

e, assim,

$$\begin{aligned}
\|df(x)(v) - df(a)(v)\|_{\mathbb{R}^n} &\leq n\|df(x)(v) - df(a)(v)\|_{M,\mathbb{R}^n} \\
&= n \max\left\{\left|\sum_{j=1}^{m} \left(\frac{\partial f_1}{\partial x_j}(x) - \frac{\partial f_1}{\partial x_j}(a)\right) v_j\right|, \cdots, \left|\sum_{j=1}^{m} \left(\frac{\partial f_n}{\partial x_j}(x) - \frac{\partial f_n}{\partial x_j}(a)\right) v_j\right|\right\} \\
&\leq n \max\left\{\sum_{j=1}^{m} \left|\frac{\partial f_1}{\partial x_j}(x) - \frac{\partial f_1}{\partial x_j}(a)\right| |v_j|, \cdots, \sum_{j=1}^{m} \left|\frac{\partial f_n}{\partial x_j}(x) - \frac{\partial f_n}{\partial x_j}(a)\right| |v_j|\right\}
\end{aligned}$$

Portanto,

$$\begin{aligned}
\|df(x) - df(a)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} &= \max_{\|v\|_{\mathbb{R}^m}<1} \|(df(x) - df(a))(v)\|_{\mathbb{R}^n} \\
&\leq n \max\left\{\sum_{j=1}^{m} \left|\frac{\partial f_1}{\partial x_j}(x) - \frac{\partial f_1}{\partial x_j}(a)\right|, \cdots, \sum_{j=1}^{m} \left|\frac{\partial f_n}{\partial x_j}(x) - \frac{\partial f_n}{\partial x_j}(a)\right|\right\} \\
&= n \sum_{j=1}^{m} \left|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right| \\
&< n \cdot m \cdot \left(\tfrac{\epsilon}{m \cdot n}\right) = \epsilon
\end{aligned}$$

ou seja, se $\|x-a\| < \delta$ então $\|df(x)-df(a)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} < \epsilon$. Portanto, $df : U \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é contínua no ponto a. Como a é arbitrário, conclui-se que $df : U \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é contínua em U.

Reciprocamente, suponhamos que f é diferenciável em U e $df : U \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é contínua. Logo pelo Teorema 6.7 existe $\dfrac{\partial f}{\partial v}$ em U, segundo qualquer direção $v \in \mathbb{R}^m$ e $\dfrac{\partial f}{\partial v}(x) = df(x)(v),\ \forall x \in U$. Em particular, se $\{e_1, \cdots, e_m\}$ é a base canônica do $\mathbb{R}^m$,

existe
$$\frac{\partial f}{\partial e_j}(x) = df(x)(e_j) = \left(\frac{\partial f_1}{\partial x_j}(x), \cdots, \frac{\partial f_n}{\partial x_j}(x)\right), \ \forall j = 1, \cdots, n$$

Além disso, se $\{u_1, \cdots, u_n\}$ é a base canônica do $\mathbb{R}^n$, temos
$$\langle df(x)(e_j), u_i\rangle_{\mathbb{R}^n} = \frac{\partial f_i}{\partial x_j}(x).$$

Isto mostra que existe $\dfrac{\partial f_i}{\partial x_j}(x)$ em todo ponto de U, para todo $i = 1, \cdots, m$ e todo $j = 1, \cdots, n$ e, além disso,
$$\frac{\partial f_i}{\partial x_j}(x) = \langle df(x)(e_j), u_i\rangle_{\mathbb{R}^n}.$$

Para provar a continuidade de
$$\frac{\partial f_i}{\partial x_j} : U \longrightarrow \mathbb{R}$$
façamos
$$\begin{aligned}
\left|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right| &= |\, \langle df(x)(e_j), u_i\rangle - \langle df(a)(e_j), u_i\rangle \,| \\
&= |\langle df(x)(e_j) - df(a)(e_j), u_i\rangle| \\
&= |\langle \left(df(x) - df(a)\right)(e_j), u_i\rangle| \\
&\leq \|\left(df(x) - df(a)\right) e_j\|_{\mathbb{R}^n} \|u_i\|_{\mathbb{R}^n} \\
&\leq \|df(x) - df(a)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} \|e_j\|_{\mathbb{R}^n} \|u_i\|_{\mathbb{R}^n}
\end{aligned}$$
isto é,
$$\left|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right| \leq \|df(x) - df(a)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)}$$
o que mostra a continuidade de $\dfrac{\partial f_i}{\partial x_j} : U \longrightarrow \mathbb{R}$. □

De acordo com o teorema anterior, se $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é de classe C^1 no aberto U então está definida a aplicação
$$\begin{array}{rccc} df: & U & \longrightarrow & \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ & x & \longmapsto & df(x) \end{array}$$

a qual é contínua.

Sabemos que $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é isomorfo a $M_{n\times m}(\mathbb{R})$ que, por sua vez, é isomorfo a $\mathbb{R}^{n\cdot m}$.

Via este isomorfismo, podemos escrever

$$\begin{array}{rccl} df: & U & \longrightarrow & \mathbb{R}^{nm} \\ & x & \longmapsto & df(x) \end{array}$$

onde as nm funções $\dfrac{\partial f_i}{\partial x_j} : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ são as coordenadas de df. Logo $df : U \longrightarrow \mathbb{R}^{nm}$ é diferenciável se, e somente se, as (nm) funções $\dfrac{\partial f_i}{\partial x_j} : U \longrightarrow \mathbb{R}$ são diferenciáveis.

Definição 6.9 *Uma função $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ definida em um aberto U, diz-se duas vezes diferenciável no ponto $a \in U$ quando existe um aberto V, com $a \in V \subset U$ tal que f é diferenciável em V e a aplicação $df : V \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é diferenciável no ponto a.*

Observação 6.4 *Uma função $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é duas vezes diferenciável em a se, e somente se, as (nm) funções $\dfrac{\partial f_i}{\partial x_j} : V \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ são diferenciáveis em a, ou seja, se, e somente se,*

$$\frac{\partial f}{\partial v} : V \longrightarrow \mathbb{R}^n$$

é diferenciável no ponto a, para todo $v \in \mathbb{R}^m$.

Dada $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ duas vezes diferenciável no ponto $a \in U$, a derivada $df : V \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ no ponto a é denominada derivada segunda de f no ponto a e é denotada por $d^2f(a)$. Então

$$d^2f(a) = d(df(a)) \in \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)) \approx \mathcal{L}(\mathbb{R}^m, \mathbb{R}^{nm})$$

Exemplo 6.5 *Seja $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ duas vezes diferenciável. Isto é equivalente a dizer que as m funções $\dfrac{\partial f}{\partial x_j} : V \longrightarrow \mathbb{R}$ são diferenciáveis no ponto a, com*

$$\begin{array}{rcccl} d\left(\frac{\partial f}{\partial x_j}\right)(a): & \mathbb{R}^m & \longrightarrow & \mathbb{R} & \\ & v & \longmapsto & d\left(\dfrac{\partial f}{\partial x_j}\right)(a)(v) & = \left(\dfrac{\partial^2 f}{\partial x_1 \partial x_j}(a), \cdots, \dfrac{\partial^2 f}{\partial x_n \partial x_j}(a)\right) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \\ & & & & = \displaystyle\sum_{k=1}^{m} \frac{\partial^2 f}{\partial x_k \partial x_j}(a) v_k \end{array}$$

Então,

$$d^2f(a) = d(df(a)) \in \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R})) \approx \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$$

onde

$$\begin{array}{rccl} d^2f(a): & \mathbb{R}^m & \longrightarrow & (\mathbb{R}^m)^* \\ & v & \longmapsto & d^2f(a)(v) = \left(\sum_{k=1}^{m} \frac{\partial^2 f}{\partial x_k \partial x_1}(a)(v_k), \cdots, \sum_{k=1}^{m} \frac{\partial^2 f}{\partial x_k \partial x_m}(a)(v_k)\right) \end{array}$$

$$= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_1}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_m}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_m}(a) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m}(a) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

$$= \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right)_{m\times m} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}_{m\times 1}$$

e

$$\begin{array}{rccl} d^2f(a)(v): & \mathbb{R}^m & \longrightarrow & \mathbb{R} \\ & w & \longmapsto & \left(d^2f(a)(v)\right)(w) = \left(\sum_{k=1}^{m} \frac{\partial^2 f(a)(v_k)}{\partial x_k \partial x_1}, \cdots, \sum_{k=1}^{m} \frac{\partial^2 f(a)(v_k)}{\partial x_k \partial x_m}\right)_{1\times m} \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}_{m\times 1} \end{array}$$

$$= w_1 \sum_{k=1}^{m} \frac{\partial^2 f(a)(v_k)}{\partial x_k \partial x_1} + \cdots + w_m \sum_{k=1}^{m} \frac{\partial^2 f(a)(v_k)}{\partial x_k \partial x_m}$$

Assim,

$$\mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R})) \approx \mathcal{L}_2(\mathbb{R}^m, \mathbb{R}) = \{B : \mathbb{R}^m \times \mathbb{R}^m \longrightarrow \mathbb{R}; B \text{ é bilinear}\}$$

$$\begin{array}{ccccc} d^2f(a) & & & & d^2f(a) \\ & \searrow & & \nearrow & \\ & & M_{m\times m}(\mathbb{R}) & & \\ & & \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right) & & \end{array}$$

Assim,

$$\left(d^2f(a)(v)\right)(w) = (w_1, \cdots, w_m) \underbrace{\left(\frac{\partial^2 f(a)}{\partial x_i \partial x_j}\right)}_{\text{matriz Hessiana}} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

isto é, fica definido o funcional linear

$$\begin{array}{rccl} d^2f(a): & \mathbb{R}^m \times \mathbb{R}^m & \longrightarrow & \mathbb{R} \\ & (v, w) & \longmapsto & d^2f(a)(v,w) = (w_1, \cdots, w_m) \left(\dfrac{\partial^2 f(a)}{\partial x_i \partial x_j}\right) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \end{array}$$

Para o caso geral, se $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é duas vezes diferenciável no ponto $a \in U$ então

$$d^2f(a) \in \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)).$$

Mas, temos o isomorfismo

$$\mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)) \approx \mathcal{L}_2(\mathbb{R}^m, \mathbb{R}^n) = \{B : \mathbb{R}^m \times \mathbb{R}^m \longrightarrow \mathbb{R}; B \text{ é bilinear } \}$$

que a cada $T \in \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ associa B_T onde $B_T(v, w) = (T(v))(w)$. Logo, $d^2f(a)$ é uma aplicação bilinear

$$\begin{array}{rccl} d^2f(a): & \mathbb{R}^m \times \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ & (v, w) & \longmapsto & d^2f(a)(v,w) = (d^2f(a)(v))(w) \end{array}$$

Definição 6.10 *Uma função $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ é de classe C^2 no aberto U quando f é duas vezes diferenciável em todo o ponto de U e a aplicação*

$$d^2f : U \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$$

é contínua.

Teorema 6.12 *(Schwarz) Sejam $U \subset \mathbb{R}^m$ um aberto e $a \in U$. Se $f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ é duas vezes diferenciável no ponto a então*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a), \ \forall\ i, j (1 \leq i, j \leq k)$$

Demonstração: Sem perda de generalidade, podemos considerar $U \subset \mathbb{R}^2$ e $a = (x_0, y_0)$. Como U é aberto, existe $c > 0$ tal que o quadrado $(x_0 - \epsilon, x_0 + \epsilon) \times (y_0 - \epsilon, y_0 + \epsilon) \subset U$. Definimos $\varphi : (-\epsilon, \epsilon) \longrightarrow \mathbb{R}$ dada por

$$\varphi(t) = f(x_0 + t, y_0 + t) - f(x_0 + t, y_0) - f(x_0, y_0 + t) + f(x_0, y_0)$$

Desta forma, para cada $t \in (-\epsilon, \epsilon)$ fixo, a função $\xi : [x_0 - t, x_0 + t] \longrightarrow \mathbb{R}$ que a cada $x \in [x_0 - t, x_0 + t]$ associa $\xi(x) = f(x, y_0 + t) - f(x, y_0)$ satisfaz

(i) $\xi(x_0 + t) - \xi(x_0) = \varphi(t)$

(ii) ξ é diferenciável no intervalo $(x_0, x_0 + t)$ e é contínua em $[x_0, x_0 + t]$

Assim, pelo teorema do valor médio, temos que existe $\theta \in (0, 1)$ tal que

$$\xi(x + 0 + t) - \xi(x_0) = \xi'(x + \theta t)t$$

logo

$$\varphi(t) = \left[\frac{\partial f}{\partial x}(x_0 + \theta t, y_0 + t) - \frac{\partial f}{\partial x}(x_0 + \theta t, y_0)\right] t$$

Como $\dfrac{\partial f}{\partial x} : U \longrightarrow \mathbb{R}$ é diferenciável em (x_0, y_0) então temos que:

$$\lim_{t \to 0} \frac{\left|\dfrac{\partial f}{\partial x}(x_0 + \theta t, y_0 + t) - \dfrac{\partial f}{\partial x}(x_0, y_0) - d\left(\dfrac{\partial f}{\partial x}\right)(x_0, y_0)(\theta t, t)\right|}{\sqrt{\theta^2 + 1}\ |t|} = 0$$

isto é,

$$\lim_{t \to 0} \frac{\dfrac{\partial f}{\partial x}(x_0 + \theta t, y_0 + t) - \dfrac{\partial f}{\partial x}(x_0, y_0) - d\left(\dfrac{\partial f}{\partial x}\right)(x_0, y_0)(\theta t, t)}{t} = 0$$

e também temos

$$\lim_{t \to 0} \frac{\dfrac{\partial f}{\partial x}(x_0 + \theta t, y_0) - \dfrac{\partial f}{\partial x}(x_0, y_0) - d\left(\dfrac{\partial f}{\partial x}\right)(x_0, y_0)(\theta t, 0)}{t} = 0$$

Daí,

$$\lim_{t \to 0} \frac{\dfrac{\partial f}{\partial x}(x_0 + \theta t, y_0 + t) - \dfrac{\partial f}{\partial x}(x_0 + \theta t, y_0) - d\left(\dfrac{\partial f}{\partial x}\right)(x_0, y_0)(0, t)}{t} = 0$$

ou,

$$\lim_{t\to 0}\left(\frac{\left[\frac{\partial f}{\partial x}(x_0+\theta t, y_0+t)-\frac{\partial f}{\partial x}(x_0+\theta t, y_0)\right]t}{t^2}-\frac{d\left(\frac{\partial f}{\partial x}\right)(x_0,y_0)(0,t)}{t}\right)=0$$

ou, ainda,

$$\lim_{t\to 0}\left(\frac{\varphi(t)}{t^2}-\frac{d\left(\frac{\partial f}{\partial x}\right)(x_0,y_0)(0,t)}{t}\right)=0$$

Mas,

$$d\left(\frac{\partial f}{\partial x}\right)(x_0,y_0)(0,t)=\frac{\partial^2 f}{\partial x^2}(x_0,y_0)0+\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0)t$$

Assim,

$$\lim_{t\to 0}\left[\frac{\varphi(t)}{t^2}-\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0)\right]=0$$

o que implica

$$\lim_{t\to 0}\frac{\varphi(t)}{t^2}=\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0)$$

Analogamente, considerando, no lugar de ξ, a função

$$\begin{array}{rcl} \psi: \ [y_0-t, y_0+t] & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \psi(y)=f(x_0+t,y)-f(x_0,y_0) \end{array}$$

obtemos

$$\lim_{t\to 0}\frac{\varphi(t)}{t^2}=\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0).$$

Segue que

$$\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)=\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0)$$

$\square$