MAT 216 - CÁLCULO III - IFUSP

1 SEMESTRE de 2014

Professor Oswaldo Rio Branco de Oliveira

http://www.ime.usp.br/~oliveira

DIFERENCIABILIDADE, REGRA DA CADEIA, MATRIZ JACOBIANA E O TEOREMA DO VALOR MÉDIO EM VÁRIAS VARIÁVEIS

Seja n um número natural e $\{e_1, \ldots, e_n\}$ a base canônica de $\mathbb{R}^n$. Então, dados $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, ambos em $\mathbb{R}^n$, seu produto interno é $\langle x, y \rangle = x_1y_1 + \cdots + x_ny_n$, também denotado $x \cdot y$. A norma de x é $|x| = \sqrt{\langle x, x \rangle}$.

Lema 1. *Seja $T : \mathbb{R}^n \to R^m$ uma aplicação linear. Então,*

(a) Existe uma constante C tal que $|T(\overrightarrow{v})| \leq C|\overrightarrow{v}|$, para todo $\overrightarrow{v}$ em $\mathbb{R}^n$.

(b) T é contínua.

Prova.

(a) Dado um vetor $v = v_1e_1 + \cdots + v_ne_n$ em $\mathbb{R}^n$, com $v_1, \ldots, v_n$ em $\mathbb{R}$, temos

$$T(v) = v_1T(e_1) + \cdots + v_nT(e_n).$$

Seja $M = \max\left\{|T(e_1)|, \ldots, |T(e_n)|\right\}$. Logo, pela desigualdade triangular,

$$\begin{aligned} |T(v)| &\leq |v_1T(e_1)| + \cdots + |v_nT(e_n)| \\ &= |v_1|\,|T(e_1)| + \cdots + |v_n|\,|T(e_n)| \\ &\leq |v|M + \cdots + |v|M = nM\|v\|. \end{aligned}$$

(b) Fixado um ponto x em $\mathbb{R}^n$ e considerando um vetor h em $\mathbb{R}^n$ temos

$$0 \leq |T(x+h) - T(x)| = |T(x) + T(h) - T(x)| = |h)| \leq nM|h|.$$

Logo $T(x+h) \xrightarrow{h\to 0} T(x)$ e T é contínua em x, para todo x em $\mathbb{R}^n$■

A norma (operador) de T é definida por

$$\|T\| = \sup\{|T(v)| : |v| = 1\}.$$

Pelo lema, temos $|T(v)| \leq C$, para todo v com $|v| = 1$. Logo, tal sup é finito. Para uma aplicação linear T e um vetor $v \neq 0$ temos, por definição,

$$\left| T\left(\frac{v}{|v|}\right) \right| \leq \|T\|.$$

Donde segue

$$|T(v)| \leq \|T\| |v|, \quad \text{para todo } v \in \mathbb{R}^n.$$

Definição. *Uma função $F : \Omega \longrightarrow \mathbb{R}^m$, com Ω um aberto em $\mathbb{R}^n$, é* diferenciável *no ponto p em Ω se existe uma aplicação linear $T : \mathbb{R}^n \longrightarrow R^m$ tal que*

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

Equivalentemente, F é diferenciável em p se existe uma função E [dita "erro"] definida em uma bola aberta $B(0;r)$, com $r > 0$, tal que para h em $B(0;r)$ temos

$$F(p+h) = F(p) + T(h) + E(h)|h|, \quad \text{com } E(0) = 0 \text{ e } \lim_{h \to 0} E(h) = 0.$$

A aplicação linear T é a diferencial de F em p e é indicada

$$T = DF(p).$$

Proposição 2 (Diferenciabilidade Implica Continuidade). *Seja Ω aberto em $\mathbb{R}^n$ e $F : \Omega \longrightarrow \mathbb{R}^m$ diferenciável em $p \in \Omega$. Então, F é contínua em p.*

Prova.

Seja T a diferencial de F em p. Como T é contínua na origem, temos

$$F(p+h) - F(p) = \frac{F(p+h) - F(p) - T(h)}{|h|}|h| + T(h) \quad \xrightarrow{h \to 0} \quad 0.0 + T(0) = 0 \blacksquare$$

Diferenciabilidade é um conceito local e a seguir simplificamos os domínios.

Teorema 3 (Regra da Cadeia). *Consideremos a função $G : \mathbb{R}^m \to \mathbb{R}^p$, diferenciável no ponto x, e a função $F : \mathbb{R}^p \to \mathbb{R}^n$, diferenciável no ponto $y = G(x)$. Então, a função composta $F \circ G : \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável no ponto x e*

$$D(F \circ G)(x) = T \circ S, \ \text{onde } T = DF(G(x)) \text{ e } S = DG(x).$$

Prova.

Seja h em $\mathbb{R}^m$, com $h \neq 0$. Notemos que $T(0) = 0$.

Avaliemos o limite, para $h \to 0$, das duas parcelas no lado direito de

$$\frac{F(G(x+h))-F(G(x))-T(S(h))}{|h|} = \frac{F(G(x+h))-F(G(x))-T[G(x+h)-G(x)]}{|h|} + T\left(\frac{G(x+h)-G(x)-S(h)}{|h|}\right).$$

Como G é diferenciável em x e T é contínua, a última (segunda) parcela tende a zero se $h \to 0$.

Além disso, para $|h| \neq 0$ e pequeno o suficiente segue

$$\left|\frac{G(x+h)-G(x)}{|h|} - S\left(\frac{h}{|h|}\right)\right| \leq 1$$

e

$$\frac{|G(x+h)-G(x)|}{|h|} \leq 1 + \left|S\left(\frac{h}{|h|}\right)\right| \leq 1 + \|S\|.$$

Quanto à primeira parcela, escrevendo $v = v(h) = G(x+h) - G(x)$ temos

$$\frac{F(G(x)+v)-F(G(x))-T(v)}{|h|} = \begin{cases} \frac{F(G(x)+v)-F(G(x))-T(v)}{|v|}\frac{|v|}{|h|}, & \text{se } v \neq 0, \\ 0, & \text{caso contrário.} \end{cases}$$

Como G é contínua em x, temos $v(h) \to 0$ se $h \to 0$. Também temos

$$\frac{|v|}{|h|} \leq 1 + \|S\| \text{ para } |h| \neq 0 \text{ e pequeno o suficiente.}$$

Então, devido à diferenciabilidade de F no ponto $G(x)$, e o Teorema do Confronto, a primeira parcela tende a 0 se $h \to 0$ ■

A MATRIZ JACOBIANA

Consideremos uma função $F : \Omega \longrightarrow \mathbb{R}^m$, com Ω um aberto não vazio em $\mathbb{R}^n$, diferenciável em p pertencente a Ω. Seja $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ a aplicação linear tal que

$$F(p+h) = F(p) + T(h) + E(h)|h|\,,\ \forall h \in B(0;r),\ \text{com } \lim_{h\to 0} E(h) = E(0) = 0.$$

Proposição 4. *Com as hipótese acima, existem as derivadas parciais*

$$\frac{\partial F}{\partial x_j}(p) = \frac{\partial F}{\partial \overrightarrow{e_j}}(p),\ \textit{para todo } j \textit{ em } \{1,\ldots,m\}\,,\ \ e\ \ T(\overrightarrow{e_j}) = \frac{\partial F}{\partial x_j}(p).$$

Prova.

Fixemos j em $\{1,\ldots,m\}$. Seja $h = te_j$, com t em $(-r,r) \smallsetminus \{0\}$. Temos,

$$F(p+te_j) = F(p) + T(te_j) + E(te_j)|te_j|.$$

Logo,

$$\frac{F(p+te_j) - F(p)}{t} \ = \ T(e_j) \ + \ E(te_j)\frac{|t|}{t}.$$

Como

$$E(te_j) \xrightarrow{t\to 0} 0 \ \ \text{e} \ \ \frac{|t|}{t} = \pm 1,$$

segue que

$$\frac{\partial F}{\partial x_j}(p) \ = \ \lim_{t\to 0} \frac{F(p+te_j) - F(p)}{t} \ = \ T(e_j) \ \blacksquare$$

Notação. Fixemos $\{e_1,\ldots,e_n\}$ a base canônica ordenada de $\mathbb{R}^n$ e $\{f_1,\ldots,f_m\}$ a base canônica ordenada de $\mathbb{R}^m$. Dada uma aplicação linear

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m\,,$$

escrevemos

$$\begin{cases} T(\overrightarrow{e_1}) & = & a_{11}\overrightarrow{f_1} + \cdots + a_{m1}\overrightarrow{f_m} \\ & \vdots & \\ T(\overrightarrow{e_n}) & = & a_{1n}\overrightarrow{f_1} + \cdots + a_{mn}\overrightarrow{f_m}\,, \end{cases}$$

onde $a_{ij} \in \mathbb{R}$, se $1 \le i \le m$ e $1 \le j \le n$. Ainda mais, associamos à aplicação T a sua matriz $[T]$ de representação em relação às bases canônicas supra citadas:

$$[T] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \ = \ [a_{ij}].$$

A matriz $[T]$ pertence ao espaço vetorial das matrizes retangulares com m linhas e n colunas de números reais: $M_{m\times n}(\mathbb{R})$. A matriz formada pela primeira coluna de $[T]$ é a matriz dos coeficientes de $T(\overrightarrow{e_1})$. Se $1 \leq j \leq n$ temos,

$$[T(\overrightarrow{e_j})] = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in M_{m\times 1}(\mathbb{R}).$$

Esquematicamente escrevemos

$$[T] = \left[\begin{bmatrix} \vdots \\ T(\overrightarrow{e_1}) \\ \vdots \end{bmatrix} \cdots \begin{bmatrix} \vdots \\ T(\overrightarrow{e_n}) \\ \vdots \end{bmatrix} \right].$$

Definição. *A* norma (de Hilbert-Schmidt) *de T, indicada $|T|$, é dada por*

$$|T|^2 \;=\; \sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} |a_{ij}|^2.$$

Comentário. Vale a seguinte relação entre a norma de Hilbert-Schmidt e a norma do sup.

$$\|T\| \leq |T| \leq \sqrt{n}\,\|T\|.$$

Verificação. Mostremos duas desigualdades.

◇ Seja $v = v_1e_1 + \cdots + v_ne_n$ em $\mathbb{R}^n$. Pela desigualdade de Cauchy-Schwartz segue

$$|T(v)|^2 = \sum_{i=1}^{m} \left|\langle (a_{i1}, \ldots, a_{in}), (v_1, \ldots, v_n)\rangle\right|^2 \leq \sum_{i=1}^{m} |(a_{i1}, \ldots, a_{in})|^2 |v|^2.$$

Logo, $|T(v)|^2 \leq |T|^2|v|^2$. Donde segue

$$\|T\| \leq |T|.$$

◇ Temos $a_{ij} = \langle T(e_j), f_i\rangle$ e também $|T(e_j)| \leq \|T\|$. Logo,

$$|T|^2 = \sum_{j=1}^{n}\sum_{i=1}^{m} \left|\langle T(e_j), f_i\rangle\right|^2 = \sum_{j=1}^{n} |T(e_j)|^2 \leq n\|T\|^2 \blacksquare$$

Consideremos, a seguir, uma função $F : \Omega \longrightarrow \mathbb{R}^m$, com Ω um aberto não vazio de $\mathbb{R}^n$, diferenciável no ponto p pertencente a Ω.

Definição. *A* matriz jacobiana *de F no ponto p, denotada $JF(p)$, é a matriz do diferencial $DF(p) \in L(\mathbb{R}^n, \mathbb{R}^m)$, em relação às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$.*

Indiquemos um ponto x de $\mathbb{R}^n$ por $x = (x_1, \ldots, x_n)$ e, analogamente, um ponto y de $\mathbb{R}^m$ por $y = (y_1, \ldots, y_m)$. Então, pelo que já vimos na Proposição 2 temos:

$$DF(p)(\overrightarrow{e_j}) = \frac{\partial F}{\partial x_j}(p).$$

Apresentando as componentes de F como $F(x) = \big(F_1(x), \ldots, F_m(x)\big)$, obtemos

$$\frac{\partial F}{\partial x_j} = \left(\frac{\partial F_1}{\partial x_j}, \ldots, \frac{\partial F_m}{\partial x_j}\right).$$

Logo,

$$\frac{\partial F}{\partial x_j}(p) = \frac{\partial F_1(p)}{\partial x_j}\overrightarrow{e_1} + \cdots + \frac{\partial F_m(p)}{\partial x_j}\overrightarrow{e_m}.$$

Pela notação introduzida concluímos que a matriz jacobiana satisfaz

$$JF(p) = [DF(p)] = [a_{ij}] \in M_{m\times n}(\mathbb{R})\,, \text{ com } a_{ij} = \frac{\partial F_i}{\partial x_j}(p).$$

Isto é,

$$JF(p) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(p) & \cdots & \frac{\partial F_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(p) & \cdots & \frac{\partial F_m}{\partial x_n}(p) \end{bmatrix}_{m\times n} = \left[\frac{\partial F_i}{\partial x_j}(p)\right]_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}.$$

Notando que a i-ésima linha de $JF(p)$ é formada pelas coordenadas do vetor gradiente de F_i, se $1 \leq i \leq n$, em relação à base canônica, identificamos

$$JF = \begin{bmatrix} \left[\cdots\overrightarrow{\nabla}F_1\cdots\right] \\ \vdots \\ \left[\cdots\overrightarrow{\nabla}F_m\cdots\right] \end{bmatrix}_{m\times n} = \begin{bmatrix} \begin{bmatrix} \vdots \\ \frac{\partial F}{\partial x_1} \\ \vdots \end{bmatrix} & \cdots & \begin{bmatrix} \vdots \\ \frac{\partial F}{\partial x_n} \\ \vdots \end{bmatrix} \end{bmatrix}_{m\times n}.$$

Dado um vetor $h = \langle h_1, \ldots, h_n\rangle$ em $\mathbb{R}^n$ temos,

$$JF(p)(\overrightarrow{h}) = \begin{bmatrix} [\cdots\nabla F_1\cdots] \\ \vdots \\ [\cdots\nabla F_m\cdots] \end{bmatrix}_{m\times n} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} \nabla F_1(p)\cdot h \\ \vdots \\ \nabla F_m(p)\cdot h \end{bmatrix}_{m\times 1}.$$

As desigualdades

$$|\nabla F_i(p)\cdot h| \le |\nabla F_i(p)|\,|h|, \text{ para cada } 1\le i\le m,$$

implicam

$$|JF(p)(h)| \;\le\; |JF(p)|\,|h|.$$

TEOREMA DO VALOR MÉDIO

Teorema 5 (Teorema do Valor Médio). *Seja $F:\Omega\to\mathbb{R}$ um campo escalar diferenciável, com Ω um aberto em $\mathbb{R}^n$. Sejam a e b dois pontos em Ω tais que o segmento $\overline{ab}$ está contido em Ω. Então, existe um ponto p em $\overline{ab}$ tal que*

$$F(b)-F(a)=\vec{\nabla}F(p)\cdot(b-a).$$

Prova.

Consideremos, em Ω, a curva

$$\gamma(t)=a+t(b-a), \text{ com } t \text{ em } [0,1].$$

É claro que γ é diferenciável e que $\gamma'(t)=b-a$. Pela regra da cadeia a função

$$F\circ\gamma:[0,1]\longrightarrow\mathbb{R},$$

é derivável e

$$(F\circ\gamma)'(t)=\nabla F\big(\gamma(t)\big)\cdot\gamma'(t).$$

Pelo TVM em uma variável temos

$$\begin{aligned} F(b)-F(a) &= (F\circ\gamma)(1)-(F\circ\gamma)(0)\\ &= (F\circ\gamma)'(t_0)\\ &= \nabla F(\gamma(t_0))\cdot\gamma'(t_0), \end{aligned}$$

para algum t_0 em $[0,1]$.

Seja $p=\gamma(t_0)$, em $\overline{ab}$. Concluímos então,

$$F(b)-F(a)=\vec{\nabla}F(p)\cdot(b-a)\blacksquare$$

$F \in C^1$ IMPLICA F DIFERENCIÁVEL

Proposição 6. *Seja $F : \Omega \longrightarrow \mathbb{R}^m$, com Ω um aberto em $\mathbb{R}^n$, e p em Ω. Suponha que as derivadas parciais de primeira ordem de F existam em todo ponto de uma bola aberta $B(p;r)$, centrada em p e contida em Ω e com $r > 0$, e que tais derivadas sejam contínuas em p. Então, F é diferenciável no ponto p.*

Prova.

Escrevamos $F = F(x) = \big(F_1(x), \ldots, F_m(x)\big)$, para x em Ω. Consideremos um vetor h em $\mathbb{R}^n$ tal que $0 < |h| < r$. Seja T a aplicação linear cuja matriz em relação às bases usuais de $\mathbb{R}^n$ e $\mathbb{R}^m$ é $JF(p)$. Logo, $T_p(h) = JF(p)(h)$ e então,

$$F(p+h) - F(p) - T(h) \equiv \begin{bmatrix} F_1(p+h) - F_1(p) \\ \vdots \\ F_m(p+h) - F_m(p) \end{bmatrix}_{m\times 1} - \begin{bmatrix} \nabla F_1(p) \cdot h \\ \vdots \\ \nabla F_m(p) \cdot h \end{bmatrix}_{m\times 1} .$$

Pelo Teorema do Valor Médio, para cada índice i em $\{1, \ldots, m\}$ existe um ponto $p_i = p_i(h)$, dependendo do vetor h e no segmento que une os pontos p e $p + h$ [o segmento está contido em $B(p;r)$], tal que

$$F_i(p+h) - F_i(p) = \nabla F_i(p_i) \cdot h.$$

Consequentemente,

$$\frac{F(p+h) - F(p) - T(h)}{|h|} \equiv \begin{bmatrix} \big[\nabla F_1(p_1) - \nabla F_1(p)\big] \cdot \frac{h}{|h|} \\ \vdots \\ \big[\nabla F_m(p_m) - \nabla F_m(p)\big] \cdot \frac{h}{|h|} \end{bmatrix} \xrightarrow{h \longrightarrow 0} \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \blacksquare$$

REFERÊNCIAS

1. Apostol, T. M., *Cálculo*, Vol. 2, Editorial Reverté, 1999.

2. Fitzpatrick, P. M., *Advanced Calculus*, 2 ed., American Math. Soc., 2009.

3. Knapp, A. W., *Basic Real Analysis*, Birkhäuser, 2005.

4. Spivak, M., *O Cálculo em Variedades*, Ed. Ciência Moderna, 2003.

Departamento de Matemática - Universidade de São Paulo

oliveira@ime.usp.br

`http://www.ime.usp.br/~oliveira`