Diferenciação

©2011 Vinicius Cifú Lopes

UFABC, 2º quad. 2011

Este capítulo completa a resposta às perguntas que fizemos em "Derivação Espacial".

Já vimos derivação de curvas (com uma variável escalar e valor vetorial), derivação parcial de funções de valor escalar, mas várias variáveis, e diversas formas de derivação de campos. Agora, derivaremos uma função $f\colon \mathbb{R}^n \to \mathbb{R}^m$ *em geral* e veremos como aquelas derivações todas eram casos particulares dessa mesma operação.

Também daremos significado, enfim, a condições de diferenciabilidade ("suficiente") sobre curvas e campos que você pode encontrar em outros textos. O capítulo ainda contém várias demonstrações, algumas em slide e outras no texto adicional: como em todo o Cálculo, cada uma não é apenas importante por provar alguma tese, mas muito mais por conter ao menos uma técnica ou raciocínio chave.

Funções de 1ª ordem

Matriz $A \in \mathrm{M}_{mn}(\mathbb{R})$ induz função

$$A\colon \mathbb{R}^n \to \mathbb{R}^m,\ x \mapsto Ax$$

onde x é visto como vetor coluna.

Função de 1ª ordem:

$$f\colon \mathbb{R}^n \to \mathbb{R}^m,\ f(x) = u + Ax$$

onde $u \in \mathbb{R}^m$ e $A \in \mathrm{M}_{mn}(\mathbb{R})$.

Veja que toda função de 1ª ordem é contínua!

A respeito das funções induzidas por matrizes, note que a soma de matrizes $m\times n$ corresponde à soma dessas funções $\mathbb{R}^n \to \mathbb{R}^m$; não há um produto específico dessas funções; ao produto de matrizes $k \times m$ e $m \times n$ corresponde a composição $\mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}^k$ das funções induzidas.

Exercício: Sendo f como antes e $g(y) = v + By$, em que condições podemos formar $f + g$ ou $g \circ f$? Mostre que, então, cada função também é de 1ª ordem.

Melhor aproximação de 1ª ordem

Dados $D \subseteq \mathbb{R}^n$, pto. a interior de D e $f\colon D \to \mathbb{R}^m$.

Queremos substituir $f(x)$, ao redor de a, por uma expressão de 1ª ordem $u + Ax$.

O raciocínio, aqui, será análogo ao que fizemos em "Derivação" com funções de uma variável; ao substituir f por uma função, que pretendemos que aproxime a primeira, convém estudar o erro cometido:

Erro absoluto: $E(x) = f(x) - (u + Ax)$.
Impomos $E(a) = 0$ (exatidão em a), donde

$$u + Aa = f(a).$$

(Diagrama na lousa.) Não distinguimos ainda a melhor aproximação!

Erro relativo:

$$\frac{E(x)}{\|x-a\|} = \frac{f(x) - (u + Ax)}{\|x-a\|} = \underset{\text{com } u+Aa=f(a)}{=\!=\!=} \frac{f(x) - f(a) - A(x-a)}{\|x-a\|}.$$

Não posso calcular em $x = a$: tomo limite.

(Observe que precisamos tomar a norma no denominador; o limite será vetorial.)

Proposição: Se existir uma matriz A tal que

$$\lim_{x\to a} \frac{1}{\|x-a\|}[f(x) - f(a) - A(x-a)] = 0,$$

então A é única.
Assim, distinguimos uma aproximação.

Para demonstrá-la, suponha que tenhamos

$$\lim_{x\to a} \frac{f(x) - f(a) - A(x-a)}{\|x-a\|} = 0 \text{ e } \lim_{x\to a} \frac{f(x) - f(a) - B(x-a)}{\|x-a\|} = 0$$

para duas matrizes A, B. Subtraindo as equações, obtemos

$$\lim_{x\to a} \tfrac{1}{\|x-a\|}\{[f(x) - f(a) - A(x-a)] - [f(x) - f(a) - B(x-a)]\} = 0$$

ou, simplesmente, $\lim_{x\to a} \frac{1}{\|x-a\|}(B - A)(x - a) = 0$. Em particular, se $x = a + he_i$ com h real, temos $\lim_{h\to 0} \frac{1}{\|he_i\|}(B - A)(he_i) = 0$, ou seja,

$$\lim_{h\to 0} \frac{h}{|h|}(B - A)e_i = 0.$$

Como sempre $h/|h| = \pm 1$, para que esse limite seja 0 é preciso que $(B - A)e_i = 0$. Contudo, veja que

$$(B - A)e_i = (B - A)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \text{a } i\text{-ésima coluna da matriz } B - A.$$

Como o índice i é arbitrário, concluímos que todas as colunas de $B - A$ são nulas e que $A = B$.

Essa proposição, ao especificar uma única aproximação de 1ª ordem como a melhor, permite fazermos a seguinte definição:

Definição: Se A existe tal que

$$\lim_{x\to a} \frac{1}{\|x-a\|}[f(x) - f(a) - A(x-a)] = 0,$$

então:

- diz-se que f é *diferenciável* em a;
- escreve-se $f'(a) = A$, sua *(matriz) diferencial;*
- diz-se que $[f(a) - Aa] + Ax$ é a *melhor aproximação de 1ª ordem* de f ao redor de a.

(Há várias notações para a matrix $f'(a)$.) Aqui, é importante o termo "diferenciável": ao contrário de nosso estudo para funções de uma variável, aqui "diferenciável" e "derivável" *não* são a mesma propriedade!

Note que $f'(a)$ é matrix $m \times n$ e induz função $\mathbb{R}^n \to \mathbb{R}^m$.

Matrizes $m \times n$ são (mn)-uplas de números, com somas e limites entrada a entrada, donde $\mathrm{M}_{mn}(\mathbb{R}) \cong \mathbb{R}^{mn}$.

Se f é diferenciável em todo o D, obtemos

$$f'\colon D \to \mathrm{M}_{mn}(\mathbb{R}) \cong \mathbb{R}^{mn},\ a \mapsto f'(a),$$

e f' é função também!

A propriedade de diferenciabilidade

Exercício: Se f é diferenciável em a, então é contínua em a.

Sugestão: assuma $\lim_{x\to a} \frac{E(x)}{\|x-a\|} = 0$, mostre $\lim_{x\to a}\big(\|x-a\| \frac{E(x)}{\|x-a\|}\big) = 0$ e use $E(x) = f(x) - f(a) - A(x-a)$.

Mas afinal, qual é a tal matrix $f'(a)$? Eis como determiná-la:

Proposição: Se f é diferenciável em a, então existem todas as derivadas parciais em a e

$$f'(a) = \left[\frac{\partial f_i}{\partial x_j}(a)\right]_{i,j}.$$

Exemplo: Dada $f(x,y,z) = (x^2y, x + 3\operatorname{sen} z)$, temos

$$f'(1,2,0) = \begin{bmatrix} 2xy & x^2 & 0 \\ 1 & 0 & 3\cos z \end{bmatrix}_{\substack{x=1\\y=2\\z=0}} = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}.$$

Exercício: Quem é $f'(a)$ e quem é f' nos casos $m = 1$, $n = 1$, $m = n = 1$?

Para demonstrá-la, assuma f dif. em a...
Escreva $A = f'(a)$, matriz com linhas A_i.
Limite vetorial: $\lim\limits_{x \to a} \frac{1}{\|x-a\|}[f(x) - f(a) - A(x-a)] = 0$.
Na i-ésima componente: $\lim\limits_{x \to a} \frac{1}{\|x-a\|}[f_i(x) - f_i(a) - A_i(x-a)] = 0$.
Se $x = a + he_j$ com h real, temos $x \xrightarrow{h \to 0} a$ e

$$\lim_{h \to 0} \tfrac{1}{\|he_j\|}[f_i(a+he_j) - f_i(a) - A_i h e_j] = 0.$$

Ou seja, $\lim\limits_{h \to 0} \frac{1}{|h|}[f_i(a+he_j) - f_i(a) - hA_{ij}] = 0$.
Então

$$\lim_{h \to 0} \overset{\pm 1}{\cancel{\tfrac{h}{|h|}}} \underbrace{\left(\frac{f_i(a+he_j) - f_i(a)}{h} - A_{ij} \right)}_{\text{deverá } \to 0} = 0.$$

Portanto,

$$\lim_{h \to 0} \frac{f_i(a+he_j) - f_i(a)}{h} = A_{ij}.$$

Assim, o limite que define $\frac{\partial f_i}{\partial x_j}(a)$ existe e é igual ao elemento ij de A!

Exercício: Assuma $u \in \mathbb{R}^n$ unitário, $m = 1$ e f diferenciável. Use a técnica acima para mostrar que

$$\frac{\partial f}{\partial u}(a) \overset{(1^{\underline{o}})}{=\!=} f'(a)u \overset{(2^{\underline{o}})}{=\!=} \langle \nabla f(a) | u \rangle.$$

Note que há duas coisas a mostrar nesse exercício: os itens ($1^{\underline{o}}$) e ($2^{\underline{o}}$). Ele apresenta que, no caso de funções escalares, a melhor aproximação pode ser escrita como $\Delta f \approx \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \cdot \Delta x_j$, onde Δ significa sempre "variação".

Exemplo na lousa (Guidorizzi, $n = 2$ e $m = 1$):

$$f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{se } (x,y) \neq (0,0), \\ 0 & \text{se } (x,y) = (0,0). \end{cases}$$

f não é diferenciável em $(0,0)$, mas é contínua.

Exemplo na lousa (Guidorizzi, $n = 2$ e $m = 1$):

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

f não é diferenciável nem contínua em $(0,0)$.

Teorema: Se $\frac{\partial f_i}{\partial x_j}$ existem ao redor de a e são contínuas em a, então f é diferenciável em a. Assim, basta checar continuidade das derivadas parciais.

Recíproca não vale (Guidorizzi):

$$f(x) = \begin{cases} \|x\| \operatorname{sen} \frac{1}{\|x\|} & \text{se } n \neq 0, \\ 0 & \text{se } x = 0. \end{cases}$$

Essa f é diferenciável (também em 0), mas suas derivadas parciais não são contínuas em 0.

Sumário:

(1) Se podemos formar $A = [\frac{\partial f_i}{\partial x_j}(a)]_{i,j}$ e $\lim_{x \to a} \frac{1}{\|x-a\|}[f(x) - f(a) - A(x-a)] = 0$, então f é diferenciável em a e $f'(a) = A$. (A é única candidata!)

(2) Se não podemos formar A, então f *não* é diferenciável em a.

(3) Se podemos formar A, mas $\lim \neq 0$ ou não existe, então f *não* é diferenciável em a.

(4) Se f não é contínua em a, *não* é diferenciável em a. (Não vale recíproca.)

(5) Se as $\frac{\partial f_i}{\partial x_j}$ são contínuas em a, então f é diferenciável em a. (Não vale recíproca.)

(6) $f' \colon D \to \mathbb{R}^{mn}$ é função, talvez diferenciável.

f é de *classe* C^k (diferenciável k vezes **e** $f^{(k)}$ contínua) $\Leftrightarrow$
$\Leftrightarrow$ todas der. parciais de f de ordem k existem **e** são contínuas.

(7) ($m = 1$) $z = f(a) + f'(a)(x-a)$ só é melhor aprox. 1ª ordem, só é eq. plano tangente, se f é diferenciável. (Diagrama na lousa.)

(Lembre que $f^{(k)} = f'^{\cdots\prime}$, onde o sinal $\prime$ aparece k vezes, é a k-ésima diferencial de f. O valor de k pode ser qualquer inteiro $0, 1, 2, \ldots$ ou ∞.)

Exemplo na lousa (Guidorizzi, $n = 2$ e $m = 1$): Retomamos

$$f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{se } (x,y) \neq (0,0), \\ 0 & \text{se } (x,y) = (0,0). \end{cases}$$

Aparece fórmula para plano "tangente" que não é tangente.

Regra da Cadeia

Teorema: Suponha $D_f \subseteq \mathbb{R}^n$, $D_g \subseteq \mathbb{R}^m$, $f \colon D_f \to D_g$ e $g \colon D_g \to \mathbb{R}^k$. Se f é diferenciável em a e g é diferenciável em $f(a)$, então $g \circ f$ é diferenciável em a e

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

O produto indicado é o de matrizes; posto isso, a regra é idêntica àquela para funções de uma variável e valores escalares. É preciso que a esteja no interior de $D_f = D_{g \circ f}$ e que $f(a)$ no interior de D_g.

Para demonstrar a regra, assuma as hipóteses dadas no enunciado e escreva $b = f(a)$, $A = f'(a)$, $B = g'(b)$. Queremos mostrar que

$$\lim_{x\to a} \tfrac{1}{\|x-a\|}[g(f(x)) - g(b) - BA(x-a)] = 0.$$

Sabemos que

$$\begin{aligned} f(x) - f(a) &= A(x-a) + E_f(x) \text{ com } \frac{E_f(x)}{\|x-a\|} \xrightarrow{x\to a} 0 \text{ e} \\ g(y) - g(b) &= B(y-b) + E_g(y) \text{ com } \frac{E_g(y)}{\|y-b\|} \xrightarrow{y\to b} 0. \end{aligned}$$

Em particular, porque f é contínua em a, se fizermos $y = f(x) \xrightarrow{x\to a} f(a) = b$ temos

$$g(f(x)) - g(b) = B(f(x) - b) + E_g(f(x)) \text{ com } \frac{E_g(f(x))}{\|f(x)-b\|} \xrightarrow{x\to a} 0.$$

Portanto,

$$\begin{aligned} g(f(x)) - g(b) &= B(A(x-a) + E_f(x)) + E_g(f(x)) = \\ &= BA(x-a) + BE_f(x) + E_g(f(x)). \end{aligned}$$

Agora, tomando $x \to a$, sabemos que $x \neq a$, mas é preciso ainda considerar os casos $f(x) \neq f(a)$ e $f(x) = f(a) = b$. Faremos os cálculos separadamente, mas de fato as duas situações podem ser simultâneas. Quando $f(x) = b$, já temos $E_g(f(x)) = 0$ e então

$$\tfrac{1}{\|x-a\|}[g(f(x)) - g(b) - BA(x-a)] = B\frac{E_f(x)}{\|x-a\|} \xrightarrow{x\to a} 0$$

porque B induz uma função contínua. Quando $f(x) \neq f(a)$, podemos escrever

$$\tfrac{1}{\|x-a\|}[g(f(x)) - g(b) - BA(x-a)] = B\frac{E_f(x)}{\|x-a\|} + \frac{E_g(f(x))}{\|f(x)-b\|} \cdot \frac{\|f(x)-f(a)\|}{\|x-a\|}.$$

O primeiro fator do produto tende a 0 já que $y = f(x) \to f(a)$; falta então mostrar que $\frac{\|f(x)-f(a)\|}{\|x-a\|}$ é limitado enquanto $x \to a$.

Para tanto, observe que

$$\begin{aligned} 0 \leqslant \frac{\|f(x)-f(a)\|}{\|x-a\|} &= \frac{\|A(x-a) + E_f(x)\|}{\|x-a\|} \leqslant \\ &\leqslant \frac{\|A(x-a)\|}{\|x-a\|} + \frac{\|E_f(x)\|}{\|x-a\|}; \end{aligned}$$

o segundo termo é a norma de $E_f(x)/\|x-a\| \to 0$ e o segundo é

$$\leqslant \frac{\|A\| \cdot \|x-a\|}{\|x-a\|} = \|A\|$$

onde utilizamos uma norma de matrizes similar à de vetores e uma desigualdade similar à de Cauchy. Por exemplo, assumindo $m = 1$, temos

$$|\langle \nabla f(a) | x-a \rangle| = |\, \|\nabla f(a)\|.\|x-a\|.\cos\theta \,| \leqslant \|\nabla f(a)\|.\|x-a\|.$$

Exercício: Dadas $V\colon \mathbb{R}^n \to \mathbb{R}$ e $\gamma\colon I \to \mathbb{R}^n$, com V diferenciável em $\gamma(t_0)$ e γ derivável em t_0, mostre que
$$(V \circ \gamma)'(t_0) = \langle \nabla V(\gamma(t_0)) | \gamma'(t_0) \rangle.$$
Basta reescrever a regra neste caso!

Fórmula usual: Dadas $V(x, y, z)$ e $\gamma(t) = (x(t), y(t), z(t))$, seja $\alpha(t) = V(\gamma(t)) = f(x(t), y(t), z(t))$. Então
$$\alpha' = \frac{\partial V}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial V}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial V}{\partial z} \cdot \frac{dz}{dt}.$$

Exemplo: $V(x, y) = 5y^2 - 3x^3 y$ e $\gamma(t) = (\cos t, 2t^3)$ em $t_0 = \pi$.
$$(V \circ \gamma)'(\pi) = \underbrace{\begin{bmatrix} -9x^2 y & 10y - 3x^3 \end{bmatrix}}_{(x,y)=\gamma(\pi)=(-1,2\pi^3)} \begin{bmatrix} -\operatorname{sen} t \\ 6t^2 \end{bmatrix}_{t=\pi} =$$
$$= \begin{bmatrix} -18\pi^3 & 20\pi^3 + 3 \end{bmatrix} \begin{bmatrix} 0 \\ 6\pi^2 \end{bmatrix} = 120\pi^5 + 18\pi^2.$$

Ou: $(V \circ \gamma)'(\pi) = \langle \nabla V(-1, 2\pi^3) | \gamma'(\pi) \rangle = \ldots = 120\pi^5 + 18\pi^2$.

Ou:
$$(V \circ \gamma)'(\pi) = \underbrace{-9x^2 y(-\operatorname{sen} t) + (10y - 3x^3)6t^2}_{t=\pi \Rightarrow x=-1,\, y=2\pi^3} = 120\pi^5 + 18\pi^2.$$

Ou diretamente:
$$(V \circ \gamma)(t) = 20t^6 - 6t^3 \cos^3 \Rightarrow$$
$$\Rightarrow (V \circ \gamma)'(t) = 120t^5 - 18t^2 \cos^3 t + 18t^3 \cos^2 t \operatorname{sen} t \Rightarrow$$
$$\Rightarrow (V \circ \gamma)'(\pi) = 120\pi^5 + 18\pi^2.$$

Teorema do Valor Médio

Notação usual: $f\colon \mathbb{R}^n \to \mathbb{R}^m$.

Caso $m = n = 1$ (FUV):

Se $f\colon [a, b] \to \mathbb{R}$ é contínua em $[a, b]$ e derivável em $]a, b[$, então existe $x_0 \in\,]a, b[$ tal que
$$\frac{f(b) - f(a)}{b - a} = f'(x_0).$$

Lembre que o TVM que estudamos em "Derivação" tem a seguinte interpretação: A velocidade média $\frac{f(b)-f(a)}{b-a}$ é de fato realizada como a velocidade instantânea $f'(x_0)$ em algum instante x_0 entre a e b.

Caso $n = 1$ (curvas): Considere a hélice $\gamma\colon [-1,1] \to \mathbb{R}^3$,

$$\gamma(t) = (\cos 2\pi t, \operatorname{sen} 2\pi t, t).$$

(Diagrama na lousa.)

Então

$$\frac{\gamma(1) - \gamma(-1)}{1-(-1)} =$$
$$= \tfrac{1}{2}[(\cos 2\pi, \operatorname{sen} 2\pi, 1) - (\cos(-2\pi), \operatorname{sen}(-2\pi), -1)] =$$
$$= (0,0,1) \text{ vertical.}$$

Mas $\forall t_0 \in \,]-1,1[$

$$\gamma'(t_0) = (-2\pi \operatorname{sen} 2\pi t_0, 2\pi \cos 2\pi t_0, 1)$$

de modo que $\gamma'(t_0)$ nunca é vertical.

Assim, o TVM precisa ser adaptado!

(O vetor nunca é vertical porque seno e cosseno nunca se anulam simultaneamente.)

Outros exemplos são possíveis: por exemplo, no espaço $\mathbb{R}^2$, o círculo $(\cos 2\pi t, \operatorname{sen} 2\pi t)$, $t \in [0,1]$, tem os pontos inicial e final iguais, de modo que o deslocamento total é nulo, mas o vetor tangente nunca é nulo.

O que vale é a seguinte propriedade:

Proposição: Se $\gamma : [a,b] \to \mathbb{R}^m$ é contínua em $[a,b]$ e derivável em $]a,b[$, então existe $t_0 \in \,]a,b[$ tal que

$$\frac{\|\gamma(b) - \gamma(a)\|}{b-a} \leqslant \|\gamma'(t_0)\|.$$

Demonstração usa TVM de FUV...

A proposição indica que, no caso de uma curva (com várias voltas), o deslocamento vetorial pode ser pequeno e, justamente por isso, em algum momento a velocidade vetorial deverá ser mais alta para compensar as voltas.

Tome $\varphi(t) = \langle \gamma(t) - \gamma(a) | \gamma(b) - \gamma(a) \rangle$.

Temos $\varphi\colon [a,b] \to \mathbb{R}$ contínua e derivável; $\varphi(a) = 0$; $\varphi(b) = \|\gamma(b) - \gamma(a)\|^2$; $\varphi'(t) = \langle \gamma'(t) | \gamma(b) - \gamma(a) \rangle$.

Então $\exists t_0 \in \,]a,b[$ tal que

$$\varphi(b) - \varphi(a) = \varphi'(t_0)(b-a).$$

Assim,

$$\frac{\|\gamma(b) - \gamma(a)\|^2}{b-a} = \langle \gamma'(t_0) | \gamma(b) - \gamma(a) \rangle \leqslant$$
$$\leqslant \|\gamma'(t_0)\| . \|\gamma(b) - \gamma(a)\|.$$

(Se $\|\gamma(b)-\gamma(a)\| = 0$ não podemos cancelar nos dois lados da inequação, mas a desigualdade é imediatamente satisfeita!)

Caso $m = 1$ (valores escalares):
Notação para segmentos de reta:

$$a, b \in \mathbb{R}^n \Rightarrow [a, b] = \{\, (1-t)a + tb \mid 0 \leqslant t \leqslant 1 \,\} =$$
$$= \{\, a + t(b-a) \mid 0 \leqslant t \leqslant 1 \,\}$$

(Diagrama na lousa.)
Não confundir com $[[a, b]]$!

Proposição: Se $D \subseteq \mathbb{R}^n$ é aberto, $[a, b] \subseteq D$ e $f\colon D \to \mathbb{R}$ é diferenciável, então $\exists t_0 \in \,]0, 1[$

$$f(b) - f(a) = \langle \nabla f((1-t_0)a + t_0 b) | b - a \rangle.$$

Note que $(1-t_0)a + t_0 b$ é um ponto alinhado entre a e b. Agora, não podemos dividir por $b - a$, que é um vetor, então devemos tê-lo "do outro lado", multiplicando através do produto interno. O gradiente de f desempenha o papel da derivada.

Esta demonstração também invoca o TVM para funções de uma variável: Tome $\varphi(t) = f((1-t)a + tb)$; note que o argumento de f é uma curva cuja derivada é sempre $b - a$. Temos $\varphi\colon [0, 1] \to \mathbb{R}$ contínua derivável, $\varphi(0) = f(a)$; $\varphi(1) = f(b)$ e $\varphi'(t) = \langle \nabla f((1-t)a + tb) | b - a \rangle$. Então existe $t_0 \in \,]0, 1[$ tal que $\varphi(1) - \varphi(0) = \varphi'(t_0)(1 - 0)$ e basta, novamente, substituir as expressões nessa equação.

Em geral (valores vetoriais): prova-se a
Proposição: Se $D \subseteq \mathbb{R}^n$ é aberto, $[a, b] \subseteq D$ e $f\colon D \to \mathbb{R}^m$ é diferenciável com

$$(\exists K > 0)(\forall x \in [a, b])\, \|f'(x)\| \leqslant K,$$

então

$$\|f(b) - f(a)\| \leqslant K\|b - a\|.$$

Como tópico opcional, apresentamos os Teoremas das Funções Implícita e Inversa: Lembre que, para

$$xyf(x, y) + (f(x, y))^3 = x,$$

aplicamos $\frac{\partial}{\partial x}$ ou $\frac{\partial}{\partial y}$ aos dois lados e isolamos $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$. Fazemos o mesmo com matrizes de diferenciais, e mais:

(1) Suponha $D \subseteq \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$, $f\colon D \to \mathbb{R}^m$ de classe C^k e $(a, b) \in D$ com $f(a, b) = 0$. Escreva $f'(a, b) = [A\ B]$, onde A é $m \times n$ e B é $m \times m$. Se B é invertível (como matriz) então existem bolas abertas $W \ni (a, b)$ e $U \ni a$ tais que $W \subseteq D$ e

$$(\forall x \in U)\exists! y_x \in \mathbb{R}^m)\, (x, y_x) \in W \text{ e } f(x, y_x) = 0$$

(note que forçosamente $y_a = b$) e a função $g\colon U \to \mathbb{R}^m$, $g(x) = y_x$, é de classe C^k.

Uma demonstração desse teorema pode ser encontrada no Guidorizzi ou no Rudin. Aqui, vejamos o que a Regra da Cadeia tem a dizer-nos: Temos $h(x) = f(x, g(x)) = 0$ em U, de modo que $h'(a) = 0$ e $h(a) = f(a, g(a)) = f(a, b)$; portanto,

$$0 = h'(a) = f'(a,b) \cdot (x, g(x))'|_{x=a} = \begin{bmatrix} A & B \end{bmatrix} \cdot \begin{bmatrix} (x)'|_{x=a} \\ g'(a) \end{bmatrix} = A \cdot 1_n + B \cdot g'(a),$$

donde podemos isolar $g'(a) = -B^{-1}A$.

(2) Suponha $D \subseteq \mathbb{R}^n$, $\Psi\colon D \to \mathbb{R}^n$ de classe C^k e $a \in D$. Escreva $A = \Psi'(a)$. Se A é invertível ($J_\Psi(a) \neq 0$), então existem bolas abertas $U \ni a$, $V \ni \Psi(a)$ tais que $\Psi|_U\colon U \to V$ é bijeção e $\Psi|_U^{-1}\colon V \to U$ tem classe C^k.

Agora, escreva $\Psi = \Psi|_U$ e $\Phi = \Psi^{-1}$: temos $\Phi(\Psi(x))$ em U, donde $(\Phi \circ \Psi)'(a) =!_n$ a matriz identidade. A Regra da Cadeia dá

$$1_n = (\Phi \circ \Psi)'(a) = \Phi'(\Psi(a)) \cdot \Psi'(a) = \Phi'(\Psi(a)) \cdot A \Rightarrow \Phi'(\Psi(a)) = A^{-1}.$$

Exercício: Mostre que (1)$\Leftrightarrow$(2), isto é, use (1) para provar (2) e vice-versa. Sugestão para (1)$\Rightarrow$(2): tome a função $x - \Psi(y) = 0$.

Polinômios de Taylor

Objetivo: substituir $f\colon \mathbb{R}^n \to \mathbb{R}^m$ por aproximações polinomiais.
Tratar cada componente separada: assumiremos $m = 1$.
Caso $n = 1$: visto em "Derivação" (FUV).
Se f derivável até ordem $d+1$, melhor aprox. polinomial a f de grau d ao redor de a é

$$\sum_{k=0}^{d} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Resto de Lagrange (erro cometido):

$$\frac{f^{(d+1)}(\xi_x)}{(d+1)!}(x-a)^{d+1} \text{ para algum } \xi_x \text{ entre } a \text{ e } x.$$

(Note: para $d = 1$, é melhor aprox. 1ª ordem: $f(a) + f'(a)(x-a)$.)

Caso n qualquer:
Sejam $[a, b]$ segm. $\subseteq D$ aberto $\subseteq \mathbb{R}^n$ e $f\colon D \to \mathbb{R}$ de classe C^{d+1}.
Melhor aprox. polinomial a f de grau d ao redor de a:

$$\sum_{k=0}^{d} \sum_{\substack{s_1+\ldots+s_n=k \\ s_1,\ldots,s_n \geqslant 0}} \frac{1}{s_1!\ldots s_n!} \cdot \frac{\partial^k f}{\partial x_1^{s_1}\ldots\partial x_n^{s_n}}(a) \prod_{i=1}^{n}(x_i - a_i)^{s_i}$$

Resto de Lagrange:

$$\sum_{\substack{s_1+\ldots+s_n=d+1 \\ s_1,\ldots,s_n \geqslant 0}} \frac{1}{s_1!\ldots s_n!} \cdot \frac{\partial^{d+1} f}{\partial x_1^{s_1}\ldots\partial x_n^{s_n}}(\xi_x) \prod_{i=1}^{n}(x_i - a_i)^{s_i}$$

para algum ξ_x no segmento $[a, x]$.

Para aplicações, deixamos a cargo do leitor procurá-las em sua área de interesse e em cursos de Cálculo Numérico; os exemplos que demos em “Derivação” para funções de uma variável já devem tê-lo convencido da importância do assunto. Aqui, é importante compreender toda a formulação utilizada: o próximo exercício trata disso.

Exercício: Expanda explicitamente o polinômio de Taylor de grau 3 para f arbitrária quando $n = 2$, usando centro (a, b) e variáveis (x, y). Aplique-o à função x^5y^7.

Solução: Escrevendo f_{xy} em vez de $\frac{\partial^2 f}{\partial x \partial y}$ etc., temos

$$\begin{aligned}
&f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + \\
&\qquad + \tfrac{1}{2}f_{xx}(a,b)(x-a)^2 + f_{xy}(a,b)(x-a)(y-b) + \tfrac{1}{2}f_{yy}(a,b)(y-b)^2 + \\
&+ \tfrac{1}{6}f_{xxx}(a,b)(x-a)^3 + \tfrac{1}{2}f_{xxy}(a,b)(x-a)^2(y-b) + \tfrac{1}{2}f_{xyy}(a,b)(x-a)(y-b)^2 + \tfrac{1}{6}f_{yyy}(a,b)(y-b)^3.
\end{aligned}$$

No caso de $f(x, y) = x^5y^7$, basta substituir:

$$\begin{aligned}
&a^5b^7 + 5a^4y^7(x-a) + 7a^5b^6(y-b) + \\
&\qquad + 10a^3b^7(x-a)^2 + 35a^4b^6(x-a)(y-b) + 21a^5b^5(y-b)^2 + \\
&\qquad + 10a^2b^7(x-a)^3 + 60a^3b^6(x-a)^2(y-b) + 105a^4b^5(x-a)(y-b)^2 + 35a^5b^4(y-b)^3.
\end{aligned}$$

Já este exercício tem solução mais elaborada; porém, é apenas um exercício combinatório e não envolve Análise:

Exercício (Demidovich — notação): Mostre que o k-ésimo termo da soma é igual a

$$\frac{1}{k!}\left[\left(\sum_{i=1}^{n}(x_i - a_i)\frac{\partial}{\partial x_i}\right)^k f\right]_{\text{calculado em } a}.$$