Capítulo 2

Cálculo Diferencial. Derivação de primeira ordem

Neste capítulo começamos estudo dos principais conceitos do cálculo diferencial para funções de várias variáveis (sejam escalares que vetoriais) tais que derivadas parciais, diferencibilidade, diferencial e gradiente.

2.1 Derivadas direccionais e parciais

Sejam $f : D \subset \mathbb{R}^n \to \mathbb{R}$ e $\mathbf{a} \in \operatorname{int} D$. Pretendemos avaliar a "taxa de variação" de f quando se atribuem "pequenos acréscimos" ao vetor $\mathbf{x}$, partindo da posição $\mathbf{a}$. Convém observar já que, enquanto no caso das funções reais de variável real os "acréscimos" possíveis tinham apenas duas direcções - a parte positiva e a parte negativa do eixo das abcissas - agora podemos considerar acréscimos $\mathbf{h}$ com qualquer das direcções de $\mathbb{R}^n$ (deverá naturalmente exigir-se que o ponto $\mathbf{a}+\mathbf{h}$ pertença ainda a D, mas isso decerto se verificará se a norma de $\mathbf{h}$ for suficientemente pequena, visto que $\mathbf{a} \in \operatorname{int} D$). É natural esperar que a "taxa de variação" de f em $\mathbf{a}$ dependa da direcção considerada (partindo de $\mathbf{a}$). Assim, por exemplo, se $f(\mathbf{a})$ for a temperatura num ponto $\mathbf{a}$ de uma sala aquecida e com uma janela aberta, é de esperar que a temperatura diminua na direcção da janela, e aumente no sentido contrário.

Definição 2.1 *Sejam $f : D \subset \mathbb{R}^n \to \mathbb{R}$ e $\mathbf{a} \in \operatorname{int} D$. Diz-se que f tem derivada em* **a** *segundo o vector $\mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq \mathbf{0}$, se existir e for finito o limite*

$$\lim_{h\to 0^+} \frac{f(\mathbf{a}+h\mathbf{v}) - f(\mathbf{a})}{h}. \tag{2.1}$$

O valor finito (2.1) chamado mesmo derivada *(ou* derivada direccional*) de* f em **a** segundo o vector **v** *(ou* na direcção de **v***) denota-se por $f'(\mathbf{a};\mathbf{v})$ ou $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$.*

Observações 2.2 *1) A igualdade (2.1) implica, em particular, que*

$$f'(\mathbf{a};\lambda\mathbf{v}) = \lambda f'(\mathbf{a};\mathbf{v})$$

para todos $\lambda > 0$. Por isso na definição da derivada direccional podemos supôr que $\|\mathbf{v}\| = 1$ (caso contrário sendo $\mathbf{v} \neq 0$ considera-se o vetor $\frac{\mathbf{v}}{\|\mathbf{v}\|}$).

2) Considerando a função real $g : \mathbb{R}^+ \to \mathbb{R}$,

$$g(t) = f(\mathbf{a} + t\mathbf{v})$$

definida numa vizinhança direita de $t = 0$, vemos que $f'(\mathbf{a}; \mathbf{v}) = g'_+(0)$ (a derivada à direita da função g no ponto 0). Fisicamente, esta derivada dá a taxa de variação de f em torno do ponto $\mathbf{a}$ na direcção do vetor $\mathbf{v}$.

Exemplo 2.3 *Seja $f(x, y) = x^2 + y^2$, $\mathbf{a} = (1, 1)$ e $\mathbf{v} = (1, 2)$. Então*

$$f(\mathbf{a} + t\mathbf{v}) = f(1 + t, 1 + 2t) = (1 + t)^2 + (1 + 2t)^2,$$

e portanto a derivada de f no ponto $\mathbf{a}$ segundo o vetor $\mathbf{v}$ é dada por

$$\left.\frac{d}{dt} f(\mathbf{a} + t\mathbf{v})\right|_{t=0} = [2(1 + t) + 4(1 + 2t)]_{t=0} = 6.$$

Para além das derivadas direccionais merece uma atenção especial "taixa de variação" de uma função ao longo dos eixos coordenados. Nomeadamente, sendo $f : D \subset \mathbb{R}^n \to \mathbb{R}$ e $\mathbf{a} \in \operatorname{int} D$ definimos a *derivada parcial* de f no ponto $\mathbf{a}$ em ordem de i-ésima coordenada (simbolicamente $\frac{\partial f}{\partial x_i}(\mathbf{a})$ ou $f_{x_i}(\mathbf{a})$) o limite

$$\lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{e}_i) - f(\mathbf{a})}{h} \tag{2.2}$$

onde está variada só a variável x_i (que admete o acrescimo h) enquanto todas outras variáveis (coordenadas) x_j, $j \neq i$, ficam constantes ($= a_j$). Assim para cálculo de uma derivada parcial em ordem de x_i basta aplicar as regras conhecidas de derivação (para funções de variável real) considerando todas coordenadas excepto x_i constantes.

Seguinte afirmação dá ligação entre as derivadas parciais e direccionais.

Proposição 2.4 *Sejam $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $\mathbf{a} \in \operatorname{int} D$ e $i = 1, 2, ..., n$. A derivada parcial $\frac{\partial f}{\partial x_i}(\mathbf{a})$ existe sse existem as derivadas direccionais $f'(\mathbf{a}; \mathbf{e}_i)$ e $f'(\mathbf{a}; -\mathbf{e}_i)$ e tem lugar a igualdade*

$$f'(\mathbf{a}; -\mathbf{e}_i) = -f'(\mathbf{a}; \mathbf{e}_i).$$

Demonstração. Pela definição a derivada parcial $\frac{\partial f}{\partial x_i}(\mathbf{a})$ é o limite (2.2) que existe sse existem ambos limites laterais

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

$$\begin{aligned} \lim_{h \to 0-} \frac{f(\mathbf{a} + h\mathbf{e}_i) - f(\mathbf{a})}{h} &= \lim_{h \to 0+} \frac{f(\mathbf{a} - h\mathbf{e}_i) - f(\mathbf{a})}{-h} \\ = \quad -\lim_{h \to 0+} \frac{f(\mathbf{a} + h(-\mathbf{e}_i)) - f(\mathbf{a})}{h} &= -f'(\mathbf{a}; -\mathbf{e}_i) \end{aligned}$$

e são iguais. ∎

Claro que as derivadas direccionais (e parciais como caso particular) obedecem as regras de derivação usuais conhecidas da Análise Matemática I.

Proposição 2.5 *Sejam* $f, g : D \subset \mathbb{R}^n \to \mathbb{R}$, $\mathbf{a} \in \operatorname{int} D$ *e* $\alpha \in \mathbb{R}$. *Se no ponto* $\mathbf{a}$ *existirem as derivadas direccionais* $\frac{\partial f}{\partial \mathbf{v}}$ *e* $\frac{\partial g}{\partial \mathbf{v}}$ *para algum vector* $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq \mathbf{0}$, *então nesse ponto existem também as derivadas direccionais* $\frac{\partial (f+g)}{\partial \mathbf{v}}$, $\frac{\partial (\alpha f)}{\partial \mathbf{v}}$, $\frac{\partial (fg)}{\partial \mathbf{v}}$ *e* $\frac{\partial \left(\frac{f}{g}\right)}{\partial \mathbf{v}}$ *(a última, naturalmente, só quando* $g(\mathbf{a}) \neq 0$*) e têm lugar as fórmulas:*

(i) $\frac{\partial (f+g)}{\partial \mathbf{v}}(\mathbf{a}) = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) + \frac{\partial g}{\partial \mathbf{v}}(\mathbf{a})$;

(ii) $\frac{\partial (\alpha f)}{\partial \mathbf{v}}(\mathbf{a}) = \alpha \frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$;

(iii) $\frac{\partial (fg)}{\partial \mathbf{v}}(\mathbf{a}) = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})\, g(\mathbf{a}) + f(\mathbf{a}) \frac{\partial g}{\partial \mathbf{v}}(\mathbf{a})$;

(iv) $\frac{\partial \left(\frac{f}{g}\right)}{\partial \mathbf{v}}(\mathbf{a}) = \frac{\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) g(\mathbf{a}) - f(\mathbf{a}) \frac{\partial g}{\partial \mathbf{v}}(\mathbf{a})}{(g(\mathbf{a}))^2}$.

Exemplos 2.6 *1) Seja* $f : \mathbb{R}^2 \to \mathbb{R}$ *uma função definida por* $f(x, y) = 2x + y$.

A derivada de f *no ponto* $\mathbf{a} = (1, 2)$ *segundo o vector* $\mathbf{v} = (1, 1)$ *é dada por*

$$\begin{aligned} f'((1,2);(1,1)) &= \lim_{h \to 0^+} \frac{f((1,2) + h(1,1)) - f(1,2)}{h} \\ &= \lim_{h \to 0^+} \frac{f(1+h, 2+h) - f(1,2)}{h} \\ &= \lim_{h \to 0^+} \frac{[2(1+h) + (2+h)] - (2 \cdot 1 + 2)}{h} = \lim_{h \to 0^+} \frac{3h}{h} = 3. \end{aligned}$$

2) Seja $f : \mathbb{R}^2 \to \mathbb{R}$ *uma função definida por* $f(x, y) = x^2 y + y^3$.

As derivadas parciais de f *num qualquer ponto* $(a, b) \in \mathbb{R}^2$ *são dadas por*

$$\frac{\partial f}{\partial x}(a, b) = 2ab \quad e \quad \frac{\partial f}{\partial y}(a, b) = a^2 + 3b^2.$$

3) Seja $f : \mathbb{R}^3 \to \mathbb{R}$ *uma função definida por* $f(x, y, z) = x^2 + y^3 + 2z$.

As derivadas parciais de f *num qualquer ponto* $(a, b, c) \in \mathbb{R}^3$ *são dadas por*

$$\frac{\partial f}{\partial x}(a, b, c) = 2a, \;\; \frac{\partial f}{\partial y}(a, b, c) = 3b^2 \quad e \quad \frac{\partial f}{\partial z}(a, b, c) = 2.$$

Exercício 2.7 *Determine todas as derivadas parciais da função definida em* $\mathbb{R}^2$ *por* $f(x, y) = x^2 + y^2 \sin(xy)$ *num ponto qualquer* $(a, b) \in \mathbb{R}^2$.

Interpretação geométrica

Restringimo-nos apenas ao caso $n = 2$. Consideremos a função $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ e $(a, b) \in \operatorname{int} D$. Recordemos que o gráfico de f é o conjunto

$$Gr\,(f) = \left\{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D \text{ e } z = f\,(x, y)\right\}.$$

Fixando y em b restringimos o domínio da função aos pontos que se encontram sobre a recta $y = b$ do plano-xy. O plano vertical de equação $y = b$ intersecta o gráfico da função numa curva que é o gráfico da função definida por $\varphi\,(x) := f\,(x, b)$. Se esta função for diferenciável no ponto $x = a$, a sua derivada será dada por

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

isto é, pela derivada parcial de f em ordem a x no ponto (a, b). Portanto, geometricamente, $f'_x\,(a, b)$ sendo a derivada da função real de variável real $\varphi\,(x) = f\,(x, b)$, representa o declive da recta tangente ao seu gráfico no ponto $(a, b, f\,(a, b))$. Analogamente, do ponto de vista geométrico $f'_y\,(a, b)$, sendo a derivada parcial da função real de variável real $\psi\,(y) := f\,(a, y)$, representa o declive da recta tangente ao seu gráfico no ponto $(a, b, f\,(a, b))$.

2.2 Diferenciabilidade e diferencial

Na teoria das funções reais de variável real, a existência de derivada de uma função num ponto implica a continuidade da função nesse ponto. No caso multidimensional a existência de derivadas direccionais, mesmo quando elas existem em todas as direcções, não garante a continuidade da função que está ilustrado no exemplo seguinte.

Exemplo 2.8 *Seja* $f : \mathbb{R}^2 \to \mathbb{R}$,

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

Vamos determinar a derivada direccional $f'\,(\mathbf{a}; \mathbf{v})$ *onde* $\mathbf{a} = (0, 0)$ *e* $\mathbf{v} = (u, v) \in \mathbb{R}^2$, $\mathbf{v} \neq \mathbf{0}$. *Há dois casos:*

(**i**) *se* $u \neq 0$, *logo para* $h > 0$ *temos*

$$\begin{aligned} \frac{f\,(\mathbf{a} + h\mathbf{v}) - f\,(\mathbf{a})}{h} &= \frac{f\,(h\mathbf{v}) - f\,(0, 0)}{h} = \frac{f\,(hu, hv)}{h} \\ &= \frac{\frac{h^3uv^2}{h^2u^2+h^4v^4}}{h} = \frac{uv^2}{u^2 + h^2v^4}, \end{aligned}$$

donde passando ao limite quando $h \to 0^+$ *concluímos que* $f'\,(\mathbf{a}; \mathbf{v}) = \frac{v^2}{u}$;

(**ii**) *se* $\mathbf{v} = (0, v)$, $v \neq 0$, *logo obtemos, de maneira análoga,* $f'\,(\mathbf{a}; \mathbf{v}) = 0$.

Portanto $f'\,(\mathbf{a}; \mathbf{v})$ *existe em todas as direcções* $\mathbf{v}$. *É facil ver também que as derivadas parciais* $f'_x\,(0, 0) = f'_y\,(0, 0) = 0$.

Observamos que $f(x,y) \to 0$ *quando* $(x,y) \to (0,0)$ *ao longo de qualquer reta que passa pela origem (ver Capítulo 1). No entanto, em cada ponto da parábola* $x = y^2$ *(excepto da origem)* f *toma o valor* $\frac{1}{2}$*. Como tais pontos existem tão próximos da origem quanto se queira e* $f(0,0) = 0$*, a função* f *não é contínua em* $(0,0)$*.*

Porque acontece isso? Porque existência da derivada no caso unidimensional implica a continuidade mas em dois dimensões já não? A razão disso pode ser vista de forma geral. Pois, se $n = 1$, pela escolha de um $h \neq 0$ e escrevendo

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

temos que quando $h \to 0$ o segundo membro tende para $f'(a) \times 0 = 0$ e por isso $f(a+h) \to f(a)$. Isto mostra que a existência de $f'(a)$ implica a continuidade de f em a. Apliquemos agora o mesmo raciocínio a uma função qualquer de n variáveis, $n > 1$. Suponhamos que a derivada $f'(\mathbf{a};\mathbf{v})$ existe para um certo $\mathbf{v}$. Então se $h \neq 0$ podemos escrever

$$f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a}) = \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h} h.$$

Quando $h \to 0^+$ o segundo membro tende para $f'(\mathbf{a};\mathbf{v}) \times 0 = 0$; logo a existência de $f'(\mathbf{a};\mathbf{v})$ para um dado $\mathbf{v}$ implica que

$$\lim_{h \to 0^+} f(\mathbf{a} + h\mathbf{v}) = f(\mathbf{a}).$$

Tal significa que $f(\mathbf{x}) \to f(\mathbf{a})$ quando $\mathbf{x} \to \mathbf{a}$ ao longo da direcção $\mathbf{v}$. Se $f'(\mathbf{a};\mathbf{v})$ existe para todo o vector $\mathbf{v}$, então $f(\mathbf{x}) \to f(\mathbf{a})$ quando $\mathbf{x} \to \mathbf{a}$ ao longo de todas as direcções. Mas isso ainda não implica a continuidade como nos vemos no Capítulo 1 e que mostra o exemplo anterior.

Conforme tudo dito acima as derivadas parciais e direccionais não constituem uma extensão satisfatória do conceito unidimensional de derivada. Existe uma generalização mais adequada que implica a continuidade. De facto, tal generalização está a ver com aproximação de uma função numa vizinhança de um ponto dado a custa de uma função linear. Essa propriedade, chamada *diferenciabilidade*, no caso unidimensional reduz-se a existência da derivada usual.

Antes de dar a respetiva definição recordemos que, se $f :]\alpha,\beta[\subset \mathbb{R} \to \mathbb{R}$ é uma função diferenciável (isto é tem a derivada finita) num ponto $a \in]\alpha,\beta[$, então, pela definição da derivada $f'(a)$, para h suficientemente pequeno (tal que $a + h \in]\alpha,\beta[$) tem-se

$$\Delta f := f(a+h) - f(a) = f'(a) h + o(h), \tag{2.3}$$

onde a função $o(h)$ verifica a condição $\lim_{h \to 0} \frac{o(h)}{h} = 0$ (diz-se um *o-pequeno quando* $h \to 0$). Note-se que a função

$$y(h) = f'(a) h$$

é linear, pois o seu gráfico é uma recta que passa pela origem. A fórmula (2.3) diz-nos que, para h pequeno, $\Delta f - y(h)$ tende para 0 *"mais depressa"* do que h quando $h \to 0$. Logo o acréscimo de f em torno de a pode ser substituido por uma forma linear (em relação ao acréscimo da variável independente) com erro que tende para zero mais depressa que h (nomeadamente, com ordem de convergência a zero superior a 1).

Passemos agora para funções de várias variáveis e por analogia com a representação (2.3) chegamos à seguinte definição.

Definição 2.9 *Diz-se que a função $f : D \subset \mathbb{R}^n \to \mathbb{R}$ é* diferenciável *num ponto $\mathbf{a} \in \operatorname{int} D$ se existe uma forma linear $\mathbf{h} \mapsto \langle \boldsymbol{\xi}, \mathbf{h} \rangle$, $\mathbf{a} \in \mathbb{R}^n$, tal que*

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \langle \boldsymbol{\xi}, \mathbf{h} \rangle + o(\mathbf{h}) \tag{2.4}$$

para $\mathbf{h}$ com a norma suficientemente pequena (pelo menos tal que $\mathbf{a}+\mathbf{h} \in D$). Aqui a função $o(\mathbf{h})$ satisfaz a igualdade

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

A forma linear $\langle \boldsymbol{\xi}, \mathbf{h} \rangle$ chama-se *diferencial* (ou *diferencial total*) de f no ponto $\mathbf{a}$ e denota-se por $df(\mathbf{a})$.

Observação 2.10 *É muito importante perceber bem o que é o diferencial de uma função escalar e não fazer confusão com a derivada de uma função real de variável real. Na Análise Real, a derivada e o diferencial de uma função num ponto existem simultaneamente. Isto quer dizer que a existência de uma implica a existência da outra e vice-versa. Se, por exemplo, $f(x) = x^2+x+1$, então a derivada de f no ponto 2 é o número 5, enquanto que o diferencial de f no ponto 2 é a função (forma) linear $df(a)(h) = 5h$.*

Resumindo: *O diferencial $df(\mathbf{a})$ é uma forma linear, não um número. O valor $df(\mathbf{a})(\mathbf{h})$ é um número real, definido para cada vetor $\mathbf{h} \in \mathbb{R}^n$.*

No caso $n = 2$ pode ser dada interpretação geométrica muito clara do conceito de diferenciabilidade apenas introduzido de mesma forma que a diferenciabilidade de uma função real de variável real. Lembramos que a diferenciabilidade de $f : [a,b] \to \mathbb{R}$ num ponto $x_0 \in]a,b[$ significa possibilidade de passar a *reta tangente* não vertical ao gráfico de f no ponto $(x_0, f(x_0))$. Analogamente (ver a fórmula (2.4)), a diferenciabilidade de uma função escalar $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, em $(x_0, y_0) \in \operatorname{int} D$ significa que existe um plano não vertical

$$z - z_0 = \langle \boldsymbol{\xi}, \mathbf{x} - \mathbf{x}_0 \rangle \tag{2.5}$$

cujos pontos são muito próximos aos pontos da superfície $z = f(x,y)$ em torno de $M(x_0, y_0, z_0)$ (a ordem de tal proximidade é superior de $\|\mathbf{x} - \mathbf{x}_0\|$). Aqui $z_0 = f(x_0, y_0)$ e $\mathbf{x} - \mathbf{x}_0 = (x - x_0, y - y_0)$. Não verticalidade desse plano (chamado *plano tangente*) significa que o coefficiente do termo $z - z_0$ em fórmula (2.5) é diferente de 0.

Uma vez que qualquer forma linear está associada com um certo vetor (no caso de diferenciabilidade é o vetor $\boldsymbol{\xi}$), é importante saber qual esse vetor. Acontece que $\boldsymbol{\xi}$ coincide com o vetor das derivadas parciais de f. Daqui sai, em particular, unicidade do diferencial total. Demonstremos seguinte teorema.

Teorema 2.11 *Supomos que $f : D \subset \mathbb{R}^n \to \mathbb{R}$ é uma função diferenciável em $\mathbf{a} \in \operatorname{int} D$. Então*

(i) a derivada direccional $f'(\mathbf{a}; \mathbf{v})$ existe para qualquer $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$, e tem-se

$$f'(\mathbf{a}; \mathbf{v}) = df(\mathbf{a})(\mathbf{v})\,; \tag{2.6}$$

(ii) todas as derivadas parciais $\frac{\partial f}{\partial x_i}(\mathbf{a})$, $i = 1, 2, ..., n$, *existem, e*

$$df(\mathbf{a})(\mathbf{h}) = \frac{\partial f}{\partial x_1}(\mathbf{a})h_1 + \frac{\partial f}{\partial x_2}(\mathbf{a})h_2 + ... + \frac{\partial f}{\partial x_n}(\mathbf{a})h_n \tag{2.7}$$

para todo $\mathbf{v} = (v_1, ..., v_n) \in \mathbb{R}^n$;

(iii) f *é contínua em* $\mathbf{a}$.

Demonstração:

(i) Como f é diferenciável em $\mathbf{a}$, temos

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + df(\mathbf{a})(\mathbf{h}) + o(\mathbf{h}), \tag{2.8}$$

sendo $\mathbf{h}$ tal que $\mathbf{a}+\mathbf{h} \in D$ e $\frac{o(\mathbf{h})}{\|\mathbf{h}\|} \to 0$ quando $\|\mathbf{h}\| \to 0$. Pomos $\mathbf{h} = \lambda\mathbf{v}$ com $\lambda > 0$ suficientemente pequeno. Então por linearidade (homogeneidade da forma linear) temos

$$df(\mathbf{a})(\mathbf{h}) = df(\mathbf{a})(\lambda\mathbf{v}) = \lambda df(\mathbf{a})(\mathbf{v}),$$

e por (2.8)

$$\frac{f(\mathbf{a}+\lambda\mathbf{v}) - f(\mathbf{a})}{\lambda} = df(\mathbf{a})(\mathbf{v}) + \frac{o(\mathbf{h})}{\lambda}. \tag{2.9}$$

Como

$$\frac{o(\mathbf{h})}{\lambda} = \frac{o(\mathbf{h})}{\|\mathbf{h}\|}\|\mathbf{v}\| \to 0 \quad \text{quando } \lambda \to 0^+,$$

o segundo membro de (2.9) tende para $df(\mathbf{a})(\mathbf{v})$ quando $\lambda \to 0^+$. Portanto o primeiro membro tende para o mesmo limite, o que prova (2.6).

(ii) Por linearidade do diferencial $df(\mathbf{a})$ imediatamente sai que

$$df(\mathbf{a})(-\mathbf{e}_i) = -df(\mathbf{a})(\mathbf{e}_i), \quad i = 1, 2, ..., n,$$

e usando (2.6) temos

$$f'(\mathbf{a}; -\mathbf{e}_i) = -f'(\mathbf{a}; \mathbf{e}_i).$$

Logo existem as derivadas parciais $\frac{\partial f}{\partial x_i}(\mathbf{a}) = df(\mathbf{a})(\mathbf{e}_i) = \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$, $i = 1, 2, ..., n$. Representando $\boldsymbol{\xi}$ na forma coordenada $(\xi_1, ..., \xi_n)$, temos que $\frac{\partial f}{\partial x_i}(\mathbf{a}) = \xi_i$, e a igualdade (2.7) sai.

(iii) Escrevemos a fórmula (2.4) na forma

$$f(\mathbf{x}) = f(\mathbf{a}) + \langle \boldsymbol{\xi}, \mathbf{x}-\mathbf{a} \rangle + o(\mathbf{x}-\mathbf{a}), \tag{2.10}$$

onde $\frac{o(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} \to 0$ e, em particular, $o(\mathbf{x}-\mathbf{a}) \to 0$ quando $\mathbf{x} \to \mathbf{a}$.

De (2.10) pela *desigualdade de Cauchy-Schwarz*

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{a})| &\leq |\langle \boldsymbol{\xi}, \mathbf{x}-\mathbf{a} \rangle| + |o(\mathbf{x}-\mathbf{a})| \\ &\leq \|\boldsymbol{\xi}\| \|\mathbf{x}-\mathbf{a}\| + |o(\mathbf{x}-\mathbf{a})|. \end{aligned} \tag{2.11}$$

Tendo em conta que

$$\|\boldsymbol{\xi}\| = \sqrt{\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(\mathbf{a})\right)^2},$$

de (2.11) concluimos que $f(\mathbf{x}) \to f(\mathbf{a})$ quando $\mathbf{x} \to \mathbf{a}$, pelo que f é contínua em $\mathbf{a}$. ■

O vetor $\boldsymbol{\xi} = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), ..., \frac{\partial f}{\partial x_n}(\mathbf{a})\right)$ associado com o diferencial da função diferenciável tem papel importante em Matemática e aplicações dela e chama-se *gradiente* da função f no ponto $\mathbf{a}$ (denota-se por $\operatorname{grad} f(\mathbf{a})$ ou por $\nabla f(\mathbf{a})$). Em términos de gradiente podemos escrever a fórmula para derivada direccional (ver Teorema 2.11 (i)-(ii)):

$$f'(\mathbf{a}; \mathbf{v}) = \langle \nabla f(\mathbf{a}), \mathbf{v} \rangle, \ \mathbf{v} \in \mathbb{R}^n, \ \mathbf{v} \neq \mathbf{0}. \tag{2.12}$$

O símbolo ∇ (lê-se "*nabla*") frequentemente usa-se sozinho para denotar o *operador diferencial formal* $\left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}\right)$.

Voltando ao sentido geométrico da diferenciabilidade podemos escrever agora a equação exacta do plano tangente ao gráfico da função $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ num ponto $\mathbf{a} \in \operatorname{int} D$ (onde f é diferenciável):

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0). \tag{2.13}$$

O vetor $\mathbf{n}(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\mathbf{i} + \frac{\partial f}{\partial y}(x_0, y_0)\mathbf{j} - \mathbf{k}$ é ortogonal ao plano tangente, e a reta que passa através o ponto $M(x_0, y_0, z_0)$ e que contem o vetor $\mathbf{n}(x_0, y_0)$ diz-se a *reta normal* á superfície $z = f(x, y)$ no ponto M. Ela pode ser definida com as equações paramétricas

$$\begin{cases} x &= x_0 + t\frac{\partial f}{\partial x}(x_0, y_0); \\ y &= y_0 + t\frac{\partial f}{\partial y}(x_0, y_0); \\ z &= z_0 - t, \qquad t \in \mathbb{R}. \end{cases}$$

As propriedades algébricas de funções diferenciáveis e diferenciais saem logo da definição e da respetivas propriedades das derivadas parciais (ver Proposição 2.5).

Proposição 2.12 *Sejam $f, g : D \subset \mathbb{R}^n \to \mathbb{R}$ funções diferenciáveis em $\mathbf{a} \in \operatorname{int} D$, e $\alpha \in \mathbb{R}$. Então*

(i) $f + g$ e αf são diferenciáveis em $\mathbf{a}$ e

$$d(f+g)(\mathbf{a}) = d(f)(\mathbf{a}) + d(g)(\mathbf{a}) \quad \text{e} \quad d(\alpha f)(\mathbf{a}) = \alpha d(f)(\mathbf{a});$$

(ii) fg é diferenciável em $\mathbf{a}$ e

$$d(fg)(\mathbf{a}) = d(f)(\mathbf{a})\, g(\mathbf{a}) + f(\mathbf{a})\, d(g)(\mathbf{a});$$

(iii) $\frac{f}{g}$ é diferenciável em $\mathbf{a}$ se $g(\mathbf{a}) \neq 0$ e

$$d\left(\frac{f}{g}\right)(\mathbf{a}) = \frac{d(f)(\mathbf{a})\, g(\mathbf{a}) - f(\mathbf{a})\, d(g)(\mathbf{a})}{g^2(\mathbf{a})}.$$

Conforme Teorema 2.11 existência das derivadas parciais é uma condição necessária para diferenciabilidade mas não é suficiente como podemos ver do seguinte simples exemplo.

Exemplo 2.13 *Seja* $f : \mathbb{R}^2 \to \mathbb{R}$,

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

Então pela definição

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0$$

e analogamente $\frac{\partial f}{\partial y}(0,0) = 0$. *No entanto,* f *não é diferenciável no ponto* $(0,0)$ *porque não é contínua (ver Capítulo 1).*

Uma função pode não ser diferenciável se até todas as derivadas direccionais (e parciais também) existem, ver Exemplo 2.8. Por isso o cálculo das derivadas parciais é só a primeira etapa da demonstração de diferenciabilidade. Depois temos de considerar a diferença

$$o(\mathbf{h}) := f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \langle \boldsymbol{\xi}, \mathbf{h} \rangle$$

e mostrar que $o(\mathbf{h}) / \|\mathbf{h}\| \to 0$ quando $\|\mathbf{h}\| \to 0$.

Exemplo 2.14 *Seja* $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} \left(x^2+y^2\right) \sin \frac{1}{\sqrt{x^2+y^2}} & se\ (x,y) \neq (0,0) \\ 0 & se\ (x,y) = (0,0). \end{cases}$$

Mostremos que f *é diferenciável em* $(0,0)$.

Calculemos primeiro as derivadas parciais de f *em* $(0,0)$. *Como*

$$\begin{aligned} \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} &= \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} \\ &= \lim_{h\to 0} \frac{h^2 \sin \frac{1}{|h|}}{h} = 0, \end{aligned}$$

temos $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$. *Assim, por (2.7),* $df(\mathbf{a})(\mathbf{h}) = 0$, *qualquer que seja o vector* $\mathbf{h} \in \mathbb{R}^2$. *Basta agora verificar que se tem* $f(x,y) = o(x,y)$ *com* $\lim\limits_{(x,y)\to(0,0)} \frac{o(x,y)}{\|(x,y)\|} = 0$*: com efeito*

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y)}{\|(x,y)\|} = \lim_{(x,y)\to(0,0)} \frac{\left(x^2+y^2\right) \sin \frac{1}{\sqrt{x^2+y^2}}}{\sqrt{x^2+y^2}}.$$

Substituindo $\sqrt{x^2+y^2} = u$, *o cálculo do limite em questão reduz-se ao limite de uma função real de variável real e é igual a*

$$\lim_{u\to 0} \left(u \sin \frac{1}{u} \right) = 0,$$

como se sabe da Análise Matemática I (o produto de um infinitésimo por uma função limitada é um infinitésimo).

Para não averiguar cada vez por definição (como fizemos no exemplo anterior) é preferível de ter um critério suficiente de diferenciabilidade que poderia ser facilmente aplicado. Tal critério formula-se também em términos das derivadas parciais.

Teorema 2.15 *Sejam $f : D \subset \mathbb{R}^n \to \mathbb{R}$ e $\mathbf{a} \in \operatorname{int} D$. Se todas as derivadas parciais de f existem nalguma vizinhança de $\mathbf{a}$ e são contínuas em $\mathbf{a}$, então f é diferenciável em $\mathbf{a}$.*

Demonstração:
Para simplesidade vamos mostrar o teorema no caso $n = 2$. O caso geral trata-se analogamente. Seja $\mathbf{a} = (a, b)$. Dado $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$ com h_1 e h_2 suficientemente pequenos tais que $(a, b + h_2)$, $(a + h_1, b + h_2) \in D$, representemos o acréscimo

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) &= f(a+h_1, b+h_2) - f(a,b) \\ &= [f(a+h_1, b+h_2) - f(a+h_1, b)] \\ &\quad + [f(a+h_1, b) - f(a,b)]. \end{aligned} \tag{2.14}$$

Fixando $y = b$ a função $f(x, b)$ de variável real x, por condição do Teorema, tem derivada no ponto a, que é a derivada parcial $f'_x(a, b)$. Por definição

$$f'_x(a,b) = \lim_{h_1 \to 0} \frac{f(a+h_1, b) - f(a,b)}{h_1}$$

donde (compare com (2.3))

$$f(a+h_1, b) - f(a,b) = f'_x(a,b)\, h_1 + \alpha(h_1)\, h_1 \tag{2.15}$$

com uma função $\alpha(h_1) \to 0$ se $h_1 \to 0$. Fixamos agora $x = a + h_1$ com h_1 suficientemente pequeno e consideramos a função $y \mapsto f(a + h_1, y)$ de variável y. Como numa vizinhança de b existe e contínua a derivada desta função, igual a $f'_y(a + h_1, y)$ (ver condições do Teorema), pelo *Teorema de Lagrange de acrescimos finitos* (ver Análise Matemática I) existe um número entre b e $b + h_2$ (seja $b + \theta h_2$ com $0 < \theta < 1$) tal que

$$f(a+h_1, b+h_2) - f(a+h_1, b) = f'_y(a+h_1, b+\theta h_2)\, h_2. \tag{2.16}$$

Observamos que θ depende de h_1 e h_2 (logo de $\mathbf{h}$) mas apesar disso $(a + h_1, b + \theta h_2) \to (a, b)$ se $\|\mathbf{h}\| \to 0$ (**Justifique!**) Pela continuidade da derivada parcial f'_y então temos

$$\lim_{\mathbf{h} \to \mathbf{0}} f'_y(a+h_1, b+\theta h_2) = f'_y(a,b).$$

A última igualdade pode ser escrita como

$$f'_y(a+h_1, b+\theta h_2) = f'_y(a,b) + \beta(\mathbf{h}) \tag{2.17}$$

para alguma função $\beta(\mathbf{h}) \to 0$, $\mathbf{h} \to 0$. Substituindo agora (2.17) em (2.16) e depois (junto com (2.15)) em (2.14) temos

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) &= f'_x(a,b)\, h_1 + \alpha(h_1)\, h_1 + \left(f'_y(a,b) + \beta(\mathbf{h})\right) h_2 \\ &= f'_x(a,b)\, h_1 + f'_y(a,b)\, h_2 + \alpha(h_1)\, h_1 + \beta(\mathbf{h})\, h_2. \end{aligned}$$

Uma vez que

$$\frac{|h_i|}{\|\boldsymbol{h}\|} = \frac{|h_i|}{\sqrt{h_1^2 + h_2^2}} \leq 1, \; i = 1, 2,$$

para $o(\mathbf{h}) := \alpha(h_1) h_1 + \beta(\mathbf{h}) h_2$ temos

$$\frac{o(\mathbf{h})}{\|\mathbf{h}\|} = \alpha(h_1) \frac{h_1}{\|\boldsymbol{h}\|} + \beta(\mathbf{h}) \frac{h_2}{\|\boldsymbol{h}\|} \to 0 \text{ se } \mathbf{h} \to 0.$$

Daqui por definição sai a diferenciabilidade de f no ponto $\mathbf{a} = (a, b)$. ■

Observação 2.16 *Demonstrando o teorema anterior nos usamos de facto a continuidade de uma derivada parcial f'_y (supondo só que a derivada f'_x existe no ponto em causa). Analogamente, no caso de funções de n variáveis para diferenciabilidade basta supôr continuidade no ponto* $\mathbf{a}$ *de $n - 1$ derivadas parciais quaisquer enquanto a derivada restante pode apenas existir em* $\mathbf{a}$. *No entanto, para objectivos prácticos não vale pena de distinguir as variáveis (em relação a quais há continuidade da derivada e em relação a qual não há).*

Definição 2.17 *Sejam $f : D \subset \mathbb{R}^n \to \mathbb{R}$,* $\mathbf{a} \in \operatorname{int} D$. *Diz-se que f é* continuamente diferenciável em $\mathbf{a}$ *se existem todas as derivadas parciais de f numa vizinhança de* $\mathbf{a}$ *e são continuas em* $\mathbf{a}$. *Dado um conjunto aberto $U \subset D$ diz-se que f é* continuamente diferenciável em U *(de* classe $\mathcal{C}^1$ em U, *simbolicamente $f \in \mathcal{C}^1(U)$) se f é continuamente diferenciável em todos os pontos de U.*

Assim, do Teorema 2.15 sai que qualquer função $f : D \subset \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^1$ num conjunto aberto $U \subset D$ é diferenciável em todos os pontos de U.

Exemplo 2.18 *Dado $i = 1, 2, ..., n$ consideremos a função simples $f(\mathbf{x}) = f(x_1, ..., x_n) = x_i$,* $\mathbf{x} \in \mathbb{R}^n$. *Uma vez que as derivadas parciais dela $\frac{\partial f}{\partial x_j} \equiv 0$ se $j \neq i$ e $\frac{\partial f}{\partial x_j} \equiv 1$ se $j = i$ são contínuas, esta função é continuamente diferenciável (em todo espaço $\mathbb{R}^n$) e o seu diferencial $df = dx_i$ associa a cada vetor* $\mathbf{h}$ *a sua i-esima coordenada h_i (ver (2.7)). Conforme isso tradicionalmente acrescimo h_i denota-se por dx_i e chama-se* ***diferencial da variável independente*** *x_i, enquanto para o acrescimo vetorial* $\mathbf{h}$ *é natural admetir o símbolo* $\mathbf{dx} = (dx_1, ..., dx_n)$. *Em tudo que está a seguir vamos usar essa simbologia. Notamos ainda que o gradiente da função acima $\nabla f(\mathbf{x}) = \mathbf{e}_i$ (i-esimo vetor básico). No caso $n = 2$ ou 3 naturalmente em vez de dx_i vamos escrever dx, dy ou dz.*

Exemplo 2.19 *Vamos mostrar que a função $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = ye^x + x^3 + 1$, é diferenciável em todo $(x, y) \in \mathbb{R}^2$.*

Encontremos primeiro as derivadas parciais por x (considerando variável y fixa) e por y (supondo que x é fixo):

$$\begin{aligned} \frac{\partial f}{\partial x} &= ye^x + 3x^2; \\ \frac{\partial f}{\partial y} &= e^x. \end{aligned}$$

Sendo funções acima elementares e definidos em todo $\mathbb{R}^2$ concluimos que $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ são contínuas, e conforme Teorema 2.15 f é (continuamente) diferenciável. O seu diferencial é a função linear (em relação às variáveis dx e dy):

$$df(x, y) = \left(ye^x + 3x^2\right) dx + e^x dy,$$

enquanto o gradiente é o vetor em $\mathbb{R}^2$ *seguinte (depende de* x *e* y*):*

$$\nabla f(x,y) = \left(ye^x + 3x^2\right)\mathbf{i} + e^x\mathbf{j}.$$

Observamos que o gradiente (como vetor das derivadas parciais) pode ser definido para qualquer função que tinha tais derivadas (não necessariamente diferenciável). Em vez o diferencial faz sentido só no caso de diferenciabilidade de função. Assim, no Exemplo 2.13 $\nabla f(0,0) = 0$ enquanto o diferencial $df(0,0)$ não faz sentido sendo a função f não diferenciável.

Fazemos mais um exemplo.

Exemplo 2.20 *O gradiente da função*

$$f(x,y,z) = xy + yz + zx - 4$$

num ponto qualquer $(x,y,z) \in \mathbb{R}^3$ *é*

$$\nabla f(x,y,z) = \left(\frac{\partial f}{\partial x}(x,y,z), \frac{\partial f}{\partial y}(x,y,z), \frac{\partial f}{\partial z}(x,y,z)\right) = (y+z, x+z, y+x).$$

Como todas as derivadas parciais são contínuas, a função f *é (continuamente) diferenciável em todo* $\mathbb{R}^3$*, e o diferencial* df *faz sentido:*

$$df(x,y,z) = (y+z)\,dx + (x+z)\,dy + (y+x)\,dz.$$

Calculado no ponto concreto $\mathbf{a} = (1,2,0)$ *temos*

$$\nabla f(1,2,0) = (2,1,3) = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k},$$

e o diferencial de f *em* $\mathbf{a}$ *é a aplicação linear*

$$df(\mathbf{a})(\mathbf{dx}) = \langle \mathbf{a}, \mathbf{dx}\rangle = 2dx + dy + 3dz.$$

Exercício 2.21 *Em que domínio do espaço* $\mathbb{R}^3$ *a função* $f : \mathbb{R}^3 \to \mathbb{R}$*,* $f(x,y,z) = \operatorname{arctg}\frac{xy}{z}$*, é diferenciável ? Determine o diferencial e o gradiente dela.*

Terminamos esta secção com uma propriedade do gradiente (de facto o sentido físico dele). De outra propriedade importante (o sentido geométrico) vamos falar mais tarde quando aprendemos funções implicitas.

Apesar que para funções de várias variáveis não é aplicável o conceito de monotonia (o espaço $\mathbb{R}^n$ não é *linearmente ordenado* para $n > 1$), é possível falar de crescimento (ou decrescimento) de uma função $f : \mathbb{R}^n \to \mathbb{R}$ ao longo de uma direcção. Nomeadamente, dado $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$, a função f diz-se *crescente* (*decrescente*) *na direcção* $\mathbf{v}$ perto de um ponto $\mathbf{a} \in \mathbb{R}^n$ se $f(\mathbf{a}) \le f(\mathbf{a} + t\mathbf{v})$ (resp., $f(\mathbf{a}) \ge f(\mathbf{a} + t\mathbf{v})$) para todo suficientemente pequeno $t > 0$.

Teorema 2.22 *Seja* $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *uma função diferenciável em* $\mathbf{a} \in \operatorname{int} D$ *e suponhamos que* $\nabla f(\mathbf{a}) \neq 0$*. Então, perto de* $\mathbf{a}$

(i) f *é crescente na direcção de qualquer vetor que faça um ângulo agudo com* $\nabla f(\mathbf{a})$*;*

(ii) f *é decrescente na direcção de qualquer vetor que faça um ângulo obtuso com* $\nabla f(\mathbf{a})$*;*

(iii) *na direcção do vector $\nabla f(\mathbf{a})$ a função f cresce mais rapidamente do que em qualquer outra direcção, isto é $f'(\mathbf{a};\mathbf{v})$ (que conforme Observação 2.2 apresenta a taxa de variação ou a velocidade da função f na direcção do vetor $\mathbf{v}$) admete o valor máximo quando $\mathbf{v} = \nabla f(\mathbf{a})$.*

Demonstração:
Interpretando a derivada direccional $f'(\mathbf{a};\mathbf{v})$ como derivada a direita da função de variável real $t \mapsto f(\mathbf{a} + t\mathbf{v})$ no ponto 0, imediatamente vemos que f é crescente ou decrescente perto do ponto $\mathbf{a}$ na direcção de $\mathbf{v}$ dependendo se $f'(\mathbf{a};\mathbf{v})$ é positiva ou negativa, respetivamente. Então, pela fórmula (2.12) para todos $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$, temos

$$f'(\mathbf{a};\mathbf{v}) = \langle \nabla f(\mathbf{a}), \mathbf{v}\rangle = \|\nabla f(\mathbf{a})\| \, \|\mathbf{v}\| \cos\theta = \|\nabla f(\mathbf{a})\| \cos\theta,$$

onde $\theta \in [0,\pi]$ é o ângulo entre os vectores $\mathbf{v}$ e $\nabla f(\mathbf{a})$. Logo $f'(\mathbf{a};\mathbf{v}) > 0$ sse θ é um ângulo agudo enquanto $f'(\mathbf{a};\mathbf{v}) < 0$ sse θ é obtuso. Daqui saem afirmações (i) e (ii). O valor máximo de $f'(\mathbf{a};\mathbf{v})$ é atingido quando $\theta = 0$ (valor máximo para $\cos\theta$), isto é, quando $\mathbf{v}$ tem a direcção do vector $\nabla f(\mathbf{a})$, ou seja $\mathbf{v} = \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$. ■

Exemplo 2.23 *Suponhamos que a função $f : \mathbb{R}^2 \to \mathbb{R}$ do Exemplo 2.19 representa a temperatura em cada ponto (x, y) do plano. Assim, no ponto $(0, 3)$ está a temperatura $f(0, 3) = 4$. Como $\nabla f(0, 3) = (3, 1)$ então, perto de $(0, 3)$ e sobre a recta que passa por $(0, 3)$ com a direcção do vector $(3, 1)$, a temperatura sobe. É também nesta direcção a temperatura sobe mais rapidamente. Nas direcções que fazem um ângulo agudo com o vetor $(3, 1)$ a temperatura também sobe, mas não tão rapidamente. Em vez nas direcções que fazem um ângulo obtuso com o vector $(3, 1)$ a temperatura desce sendo descesa maior na direcção $-\nabla f(0, 3) = (-3, -1)$.*

Exercício 2.24 *Suponha que a função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x, y) = xe^y$$

representa a temperatura em cada ponto (x, y) do plano.

(a) *Determine a taxa de variação de f no ponto $P = (2, 0)$ na direcção do vector $\overrightarrow{PQ} = Q - P$ onde $Q = \left(\frac{1}{2}, 2\right)$.*

(b) *Partindo do ponto P qual é a direcção em que a temperatura aumenta mais rapidamente?*

(c) *Qual é a máxima taxa de variação de f no ponto P (entre todas as direcções)?*

2.3 Regra de cadeia

Além das regras de derivação anunciadas em Proposições 2.5 e 2.12 há uma regra de derivação de funções compostas de várias variáveis que costuma ser útil quando a função em causa é demasiado complicada. Lembramos primeiro que para duas funções reais de variável real f e g a derivada de composição delas $h = f \circ g$ é dada por

$$h'(t) = f'(g(t))\, g'(t),$$

sempre que g é derivável em t e f é derivável em $g(t)$.

Generalizando esta fórmula temos seguinte resultado.

Teorema 2.25 *Sejam* $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *uma função escalar diferenciável em* $\mathbf{a} = (a_1, ..., a_n) \in \operatorname{int} D$ *e* $g_i : [\alpha, \beta] \subset \mathbb{R} \to \mathbb{R}$, $\alpha < \beta$, $i = 1, ..., n$, *funções diferenciáveis num ponto* $t_0 \in]\alpha, \beta[$, $a_i = g_i(t_0)$, $i = 1, ..., n$. *Então a função composta* $\Phi(t) = f(g_1(t), ..., g_n(t))$ *é diferenciável em* t_0 *e tem-se*

$$\Phi'(t_0) = \frac{\partial f}{\partial x_1}(x_0)\, g_1'(t_0) + ... + \frac{\partial f}{\partial x_n}(x_0)\, g_n'(t_0). \tag{2.18}$$

Demonstração.
Sendo a função f diferenciável em $\mathbf{a}$ temos, para $\mathbf{h} = (h_1, ..., h_n) \in \mathbb{R}^n$ com $\|\mathbf{h}\|$ suficientemente pequeno (tal que $\mathbf{a} + \mathbf{h} \in D$),

$$\begin{aligned} & f(a_1 + h_1, ..., a_n + h_n) - f(a_1, ..., a_n) \\ = & \frac{\partial f}{\partial x_1}(\mathbf{a})\, h_1 + ... + \frac{\partial f}{\partial x_n}(\mathbf{a})\, h_n + o(\mathbf{h}). \end{aligned} \tag{2.19}$$

Atribuímos a h_i, $i = 1, ..., n$, o valor concreto $g_i(t_0 + \tau) - g_i(t_0)$ para algum τ pequeno. Como as funções g_i são contínuas (uma vez que são diferenciáveis) temos que $g_i(t_0 + \tau) - g_i(t_0) \to 0$ quando $\tau \to 0$. Logo estas diferenças podem ser consideradas como acréscimos das variáveis x_i (no ponto $\mathbf{a}$). Tendo ainda em conta que $g_i(t_0) = a_i$, escrevemos a igualdade (2.19) na forma:

$$\begin{aligned} \Phi(t_0 + \tau) - \Phi(t_0) &= f(g_1(t_0 + \tau), ..., g_n(t_0 + \tau)) - f(g_1(t_0), ..., g_n(t_0)) \\ &= f(g_1(t_0) + h_1, ..., g_n(t_0) + h_n) - f(g_1(t_0), ..., g_n(t_0)) \\ &= f(a_1 + h_1, ..., a_n + h_n) - f(a_1, ..., a_n) \\ &= \frac{\partial f}{\partial x_1}(\mathbf{a})(g_1(t_0 + \tau) - g_1(t_0)) + ... \\ &\quad + \frac{\partial f}{\partial x_n}(\mathbf{a})(g_n(t_0 + \tau) - g_n(t_0)) + o(\mathbf{h}). \end{aligned}$$

Depois de dividir ambos os membros da igualdade obtida por $\tau \neq 0$ temos

$$\begin{aligned} \frac{\Phi(t_0 + \tau) - \Phi(t_0)}{\tau} &= \frac{\partial f}{\partial x_1}(\mathbf{a}) \frac{g_1(t_0 + \tau) - g_1(t_0)}{\tau} + ... \\ &\quad + \frac{\partial f}{\partial x_n}(\mathbf{a}) \frac{g_n(t_0 + \tau) - g_n(t_0)}{\tau} + \frac{o(\mathbf{h})}{\tau}. \end{aligned} \tag{2.20}$$

Como

$$\lim_{\tau \to 0} \frac{g_i(t_0 + \tau) - g_i(t_0)}{\tau} = g_i'(t_0), \; i = 1, 2, ..., n, \tag{2.21}$$

temos primeiro que as funções $\tau \mapsto \frac{g_i(t_0+\tau) - g_i(t_0)}{\tau}$ são limitados em torno de 0, isto é para todos $\tau \neq 0$ com $|\tau|$ suficientemente pequeno tem-se

$$\left| \frac{g_i(t_0 + \tau) - g_i(t_0)}{\tau} \right| \leq M, \; i = 1, 2, ..., n,$$

com algum número $M > 0$. Daqui sai que

$$\begin{aligned} \frac{\|\mathbf{h}\|}{|\tau|} &= \frac{\sqrt{(g_1(t_0 + \tau) - g_1(t_0))^2 + ... + (g_n(t_0 + \tau) - g_n(t_0))^2}}{|\tau|} \\ &= \sqrt{\left(\frac{g_1(t_0 + \tau) - g_1(t_0)}{\tau}\right)^2 + ... + \left(\frac{g_n(t_0 + \tau) - g_n(t_0)}{\tau}\right)^2} \end{aligned}$$

é limitado também ($\leq M\sqrt{n}$). Logo

$$\lim_{\tau\to 0}\frac{o(\mathbf{h})}{\tau}=\lim_{\tau\to 0}\frac{o(\mathbf{h})}{\|\mathbf{h}\|}\cdot\frac{\|\mathbf{h}\|}{\tau}=0$$

donde (tendo em conta (2.21)) concluímos que o limite

$$\lim_{\tau\to 0}\frac{\Phi(t_0+\tau)-\Phi(t_0)}{\tau}=\Phi'(t_0)$$

existe. Passando em (2.20) ao limite mostramos a fórmula (2.18). ■

Observação 2.26 *Recordando a definição de gradiente, podemos escrever a fórmula (2.18) na forma vetorial*

$$\Phi'(t_0)=\left\langle\nabla f(\mathbf{x}_0),g'(t_0)\right\rangle, \tag{2.22}$$

onde

$$g'(t_0)=\left(g_1'(t_0),...,g_n'(t_0)\right).$$

Na fórmula (2.22) temos que substituir $\mathbf{x}_0$ *pelo vector* $(g_1(t_0),...,g_n(t_0))$.

Exemplo 2.27 *Sejam* $f:\mathbb{R}^3\to\mathbb{R}$ *e* $g:\mathbb{R}\to\mathbb{R}$, $g(t)=(g_1(t),g_2(t),g_3(t))$, *definidas, respetivamente, por*

$$f(x,y,z)=z^2-x^2-y^2\quad e\quad g_1(t)=\cos t,\quad g_2(t)=\sin t,\quad g_3(t)=t^3.$$

Temos

$$\nabla f(x,y,z)=(-2x,-2y,2z)$$

e

$$g'(t)=\left(-\sin t,\cos t,3t^2\right).$$

Então a derivada da função composta $f\circ g$ *num ponto* t *qualquer é*

$$\begin{aligned}(f\circ g)'(t)&=\left\langle\nabla f(g(t)),g'(t)\right\rangle=\left\langle\left(-2\cos t,-2\sin t,2t^3\right),\left(-\sin t,\cos t,3t^2\right)\right\rangle\\&=2\cos t\sin t-2\sin t\cos t+6t^5=6t^5.\end{aligned}$$

No caso mais complicado quando as variáveis $x_1,...,x_n$ no teorema anterior dependem a sua vez de várias variáveis, sejam $u_1,...,u_k$, para encontrar a derivada parcial $\frac{\partial f}{\partial u_j}$ basta fixar outras variáveis independentes u_s, $s\neq j$, e referir a Teorema 2.25 (nos assinalamos de mesma letra a função que depende de x_i e a função composta dependente de u_j). Assim podemos escrever a fórmula que exprime as derivadas parciais em ordem as *variáveis independentes* u_j, $j=1,...,k$, através as derivadas em ordem as *variáveis dependentes* x_i:

$$\frac{\partial f}{\partial u_j}=\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_j}+...+\frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_j}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial u_j}. \tag{2.23}$$

Esta fórmula é bem conhecida por *regra da cadeia*. Tendo em conta a diferenciabilidade de funções $x_i(u_1,...,u_k)$ (que no entanto não foi usada para dedução da fórmula (2.23)) podemos também demonstrar a diferenciabilidade da função composta.

Consideramos dois exemplos onde a aplicação da regra de cadeia (2.23) essencialmente simplifica os cálculos.

Exemplo 2.28 *Seja* $f : \mathbb{R} \to \mathbb{R}$, $f(x) = (\sin x)^{\cos^2 x}$, *a função de uma variável definida no conjunto*

$$D(f) = \{x \in \mathbb{R} : \sin x > 0\}.$$

Apesar que f *é função de uma variável, para encontrar a derivada dela* $f'(x)$ *pode ser útil a regra de cadeia (2.23). Pois, fazemos a substituição de variáveis denotando*

$$\begin{cases} u = u(x) = \sin x; \\ v = v(x) = \cos^2 x \end{cases} \tag{2.24}$$

e representamos a função $f(x)$ *como função composta* $g(u(x), v(x))$ *onde*

$$g(u,v) = u^v.$$

Assim, pela regra de cadeia temos

$$f'(x) = \frac{df}{dx} = \frac{\partial g}{\partial u}\frac{du}{dx} + \frac{\partial g}{\partial v}\frac{dv}{dx} \tag{2.25}$$

(aqui as derivadas parciais em relação a única variável estão assinaladas com símbolos $\frac{df}{dx}$, $\frac{du}{dx}$ *e* $\frac{dv}{dx}$ *tradicionais em Análise Matemática I). Encontrando depois todas as derivadas envolvidas:*

$$\begin{aligned} \frac{\partial g}{\partial u} &= vu^{v-1}, \quad \frac{\partial g}{\partial v} = u^v \ln u, \\ \frac{du}{dx} &= \cos x, \quad \frac{dv}{dx} = -2\cos x \sin x = -\sin 2x, \end{aligned}$$

e substituindo em (2.25) concluimos que

$$\begin{aligned} f'(x) &= vu^{v-1}\cos x - u^v \sin 2x \\ &= \cos^2 x\,(\sin x)^{\cos^2 x - 1}\cos x - (\sin x)^{\cos^2 x}\sin 2x \\ &= \cos^3 x\,(\sin x)^{-\sin^2 x} - \sin 2x\,(\sin x)^{\cos^2 x}. \end{aligned}$$

Aqui nos tomamos em conta a definição de u *e* v *(ver (2.24)) e aplicamos algumas fórmulas trigonométricas.*

Exemplo 2.29 *Consideremos a função* f *de três variáveis*

$$f(x,y,z) = \frac{\sin(xy) + z}{\ln(x^2 - yz)}$$

e calculemos as derivadas parciais $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ *e* $\frac{\partial f}{\partial z}$ *em todos pontos onde existam. Aqui é conviniente fazer uma substituição*

$$\begin{cases} u = u(x,y,z) = \sin(xy) + z; \\ v = v(x,y,z) = \ln\left(x^2 - yz\right). \end{cases} \tag{2.26}$$

Então em términos das variáveis (dependentes) u *e* v *a função* $f(x,y,z)$ *pode ser escrita como* $g(u,v) = \frac{u}{v}$, *isto é*

$$f(x,y,z) = g(u(x,y,z), v(x,y,z)).$$

Aplicando a regra de cadeia (trés vezes), temos

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial g}{\partial v}\frac{\partial v}{\partial x};\\ \frac{\partial f}{\partial y} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial g}{\partial v}\frac{\partial v}{\partial y};\\ \frac{\partial f}{\partial z} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial z}+\frac{\partial g}{\partial v}\frac{\partial v}{\partial z}.\end{aligned} \tag{2.27}$$

As derivadas $\frac{\partial g}{\partial u}$ *e* $\frac{\partial g}{\partial u}$ *encontram-se imediatamente:*

$$\begin{aligned}\frac{\partial g}{\partial u} &= \frac{1}{v}=\frac{1}{\ln\left(x^2-yz\right)};\\ \frac{\partial g}{\partial v} &= -\frac{u}{v^2}=-\frac{\sin\left(xy\right)+z}{\ln^2\left(x^2-yz\right)},\end{aligned} \tag{2.28}$$

enquanto derivando as funções (2.26) temos:

$$\begin{aligned}\frac{\partial u}{\partial x} &= y\cos\left(xy\right), \quad \frac{\partial v}{\partial x}=\frac{2x}{x^2-yz},\\ \frac{\partial u}{\partial y} &= x\cos\left(xy\right), \quad \frac{\partial v}{\partial y}=-\frac{z}{x^2-yz},\\ \frac{\partial u}{\partial z} &= 1, \qquad\qquad\quad \frac{\partial v}{\partial z}=-\frac{y}{x^2-yz}.\end{aligned} \tag{2.29}$$

Agora só basta substituir os valores das derivadas (2.28) e (2.29) em (2.27). Por exemplo,

$$\frac{\partial f}{\partial x}=\frac{y\cos\left(xy\right)}{\ln\left(x^2-yz\right)}-\frac{\sin\left(xy\right)+z}{\ln^2\left(x^2-yz\right)}\frac{2x}{x^2-yz}.$$

Dê a representação das derivadas parciais $\frac{\partial f}{\partial y}$ ***e*** $\frac{\partial f}{\partial z}$***!***

Exercício 2.30 *Aplique a regra de cadeia para cálculo das derivadas parciais* $\frac{\partial f}{\partial x}$ *e* $\frac{\partial f}{\partial y}$ *da função* $f:\mathbb{R}^2\to\mathbb{R},\ f\left(x,y\right)=\sin\left(x+y\right)^{\sin(xy)}$.

Usando a regra de cadeia demonstramos a seguinte propriedade importante, chamada *invariância da forma do primeiro diferencial.*

Teorema 2.31 *A fórmula para o diferencial de uma função diferenciável (num ponto* $\mathbf{x}_0$*)*

$$df=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}dx_i \tag{2.30}$$

é também válida quando x_i*,* $i=1,...,n$*, dependem de outras variáveis (*dx_i *denota o diferencial da respectiva função).*

Demonstração.
Se $x_i = x_i(u_1, ..., u_k)$ são funções (de várias variáveis $u_1, ..., u_k$) então o diferencial da função composta (denotada pela mesma letra f) será, por definição,

$$df = \sum_{j=1}^{k} \frac{\partial f}{\partial u_j} du_j.$$

Substituindo nesta fórmula as derivadas $\frac{\partial f}{\partial u_j}$ calculadas usando da regra de cadeia (ver (2.23)) escrevemos

$$df = \sum_{j=1}^{k} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial u_j} du_j.$$

Alterando a ordem do somatório e fazendo algumas trasformações óbvias, encontramos

$$df = \sum_{i=1}^{n} \sum_{j=1}^{k} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial u_j} du_j = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \sum_{j=1}^{k} \frac{\partial x_i}{\partial u_j} du_j.$$

Da última igualdade, e recordando que

$$dx_i = \sum_{j=1}^{k} \frac{\partial x_i}{\partial u_j} du_j,$$

sai a representação (2.30). ■

2.4 Teorema da função implícita

Uma das várias maneiras de definir uma função real de variável real (como sabemos da Análise Matemática I) é através da equação

$$F(x, y) = 0. \tag{2.31}$$

Nomeadamente, se dado $D \subset \mathbb{R}$ e a cada $x \in D$ está associado um (único) y satisfazendo a equação (2.31) diz-se que em D está definida *implicitamente* uma função $y = f(x)$ tal que

$$F(x, f(x)) = 0, \ \ x \in D.$$

Por exemplo, para $F(x, y) = x^2 + y^2 - 1$ existem várias funções definidas no segmento $[-1, 1]$ que satisfazem a equação (2.31):

$$\begin{aligned}
f_1(x) &= \sqrt{1 - x^2}; \\
f_2(x) &= -\sqrt{1 - x^2}; \\
f_3(x) &= \begin{cases} \sqrt{1 - x^2} & \text{se} \quad x \in [-1, 0[\,, \\ -\sqrt{1 - x^2} & \text{se} \quad x \in [0, 1]\,; \end{cases} \\
f_4(x) &= \begin{cases} -\sqrt{1 - x^2} & \text{se} \quad x \in \left[-1, \frac{1}{2}\right], \\ \sqrt{1 - x^2} & \text{se} \quad x \in \left]\frac{1}{2}, 1\right]. \end{cases}
\end{aligned}$$

De facto, existe uma infinidade de funções, sendo apenas duas delas (f_1 e f_2) contínuas. Para definir uma função à custa da equação (2.31) temos, no entanto, de ter a certeza que para cada x a tal raiz y é única. No exemplo anterior, para termos unicidade precisamos de uma condição extra ($y \geq 0$ ou $y \leq 0$).

Por outro lado, a (única) função implícita pode ser definida localmente. Fixemos um ponto (x_0, y_0) com $y_0 \neq 0$ tal que $F(x_0, y_0) = 0$ (isto é, (x_0, y_0) está na circunferência $x^2 + y^2 = 1$ mas é diferente dos pontos $(\pm 1, 0)$). Então existe uma vizinhança rectangular deste ponto (seja $U =]x_0 - \delta, x_0 + \delta[\times]y_0 - \varepsilon, y_0 + \varepsilon[$) que não intersecta a recta $y = 0$. Se em (2.31) tivermos a restrição de $F(x, y) = x^2 + y^2 - 1$ ao conjunto U, então existe uma (única) função $y = f(x)$ definida implicitamente no intervalo $]x_0 - \delta, x_0 + \delta[$ tal que $f(x_0) = y_0$. Além disso, tal função é contínua e diferenciável em $]x_0 - \delta, x_0 + \delta[$. Em torno dos pontos $(\pm 1, 0)$ (isto é, quando $y_0 = 0$) não existe tal função implícita (por exemplo, à direita do ponto $(1, 0)$ a equação (2.31) não tem soluções, enquanto que para qualquer ponto x à esquerda de $(1, 0)$ há duas soluções).

Agora o nosso objectivo é encontrar condições que nos permitam definir coretamente uma função implícita na vizinhança de um ponto dado. Observamos, no entanto, que apenas em situações bastante raras (como é o caso do exemplo acima), é possível representar tal função através de alguma expressão analítica usando funções elementares. Por exemplo, se

$$F(x, y) = xy + \sin(x + y) + \cos(x + y) - 5$$

a equação (2.31) não pode ser resolvida relativamente a y (nem a x) na vizinhança de nenhum ponto.

Teorema 2.32 (de função implicita) *Seja $F : D \to \mathbb{R}$, $D \subset \mathbb{R}^2$, uma função contínua numa vizinhança de um ponto $(x_0, y_0) \in \operatorname{int} D$ tal que a derivada parcial $F'_y(x_0, y_0)$ existe nessa vizinhança e é contínua em (x_0, y_0). Se $F(x_0, y_0) = 0$ e $F'_y(x_0, y_0) \neq 0$ então existem $\delta > 0$ e $\varepsilon > 0$ tais que para cada $x \in]x_0 - \delta, x_0 + \delta[$ a equação (2.31) tem uma única solução (seja $y = f(x) \in]y_0 - \varepsilon, y_0 + \varepsilon[$), e a função $f :]x_0 - \delta, x_0 + \delta[\to]y_0 - \varepsilon, y_0 + \varepsilon[$, $f(x_0) = y_0$, é contínua. Se a derivada $F'_x(x_0, y_0)$ existe então a função f é diferenciável no ponto x_0, e a sua derivada pode ser calculada pela fórmula:*

$$f'(x_0) = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}. \tag{2.32}$$

Além disso, se a função F é continuamente diferenciável numa vizinhança de (x_0, y_0) e $F'_y(x, y) \neq 0$, então existe a derivada $f'(x)$ e é contínua nalguma vizinhança de x_0, e verifica-se a fórmula (2.32) em todos os pontos dessa vizinhança.

Demonstração.
Supomos que $F'_y(x_0, y_0) > 0$, o outro caso é análogo. Seja $\eta > 0$ tal que a função F seja contínua e exista a derivada parcial $F'_y(x, y) > 0$ na vizinhança rectangular (aliás quadrada)

$$U =]x_0 - \eta, x_0 + \eta[\times]y_0 - \eta, y_0 + \eta[.$$

Uma vez que $F'_y(x_0, y_0)$ é a derivada usual da função $y \mapsto F(x_0, y)$, a condição $F'_y(x, y) > 0$ diz-nos que esta função de uma variável é estritamente crescente próximo de y_0. Assim, existe $0 < \varepsilon < \frac{\eta}{2}$ tal que

$$F(x_0, y_0 - \varepsilon) < F(x_0, y_0) = 0 < F(x_0, y_0 + \varepsilon) .$$

Pela continuidade de F nos pontos $(x_0, y_0 \pm \varepsilon) \in U$ conclui-se que a função F mantém os mesmos sinais nalgumas vizinhanças deles. Por outras palavras, existe $0 < \delta < \varepsilon$, tal que

$$F(x,y) < 0 \quad \text{para todos} \quad (x,y) \in]x_0 - \delta, x_0 + \delta[\times]y_0 - \varepsilon - \delta, y_0 - \varepsilon + \delta[$$

e

$$F(x,y) > 0 \quad \text{para todos} \quad (x,y) \in]x_0 - \delta, x_0 + \delta[\times]y_0 + \varepsilon - \delta, y_0 + \varepsilon + \delta[.$$

Fixamos agora $x \in]x_0 - \delta, x_0 + \delta[$. Pela escolha anterior temos $F(x, y_0 - \varepsilon) < 0$ e $F(x, y_0 + \varepsilon) > 0$. Aplicando o *Teorema de Bolzano do valor intermédio* para funções de variável real encontramos um $y \in]y_0 - \varepsilon, y_0 + \varepsilon[$ tal que $F(x,y) = 0$. Tal y é único por causa da monotonia estrita de F (recordemos que $F'_y(x,y) > 0$). Associamos este y a x e definimos assim a função $y = f(x)$. Pela unicidade sai imediatamente que $f(x_0) = y_0$. A continuidade de f sai directamente da construção. Pois dado $\varepsilon > 0$ escolhemos $\delta > 0$ tal que $f(x) \in]y_0 - \varepsilon, y_0 + \varepsilon[$ (ou, equivalentemente, $|f(x) - f(x_0)| < \varepsilon$) para todos $x \in]x_0 - \delta, x_0 + \delta[$.

Verifiquemos agora a diferenciabilidade. Dados $\varepsilon > 0$ e $\delta > 0$ como acima escolhemos um acréscimo τ, $|\tau| < \delta$, da variável independente x e o respectivo acréscimo h, $|h| < \varepsilon$, de y e escrevemos a condição de diferenciabilidade da função de duas variáveis F no ponto (x_0, y_0):

$$F(x_0 + \tau, y_0 + h) - F(x_0, y_0) = F'_x(x_0, y_0)\,\tau + F'_y(x_0, y_0)\,h + o(\tau, h) \tag{2.33}$$

onde a função $o(\tau, h)$ é tal que

$$\lim_{(\tau,h)\to(0,0)} \frac{o(\tau,h)}{\sqrt{\tau^2 + h^2}} = 0.$$

Denotamos por $\alpha(\tau,h)$ a infinitésima $\frac{o(\tau,h)}{\sqrt{\tau^2+h^2}}$ (isto é $\alpha(\tau,h) \to 0$ se $(\tau,h) \to (0,0)$) e representamos $o(\tau,h)$ na forma

$$\begin{aligned} o(\tau,h) &= \alpha(\tau,h)\sqrt{\tau^2 + h^2} \\ &= \alpha(\tau,h)\left(\frac{\tau^2}{\sqrt{\tau^2+h^2}} + \frac{h^2}{\sqrt{\tau^2+h^2}}\right) \\ &= \alpha_1(\tau,h)\,\tau + \alpha_2(\tau,h)\,h \end{aligned} \tag{2.34}$$

onde

$$\alpha_1(\tau,h) := \alpha(\tau,h)\frac{\tau}{\sqrt{\tau^2+h^2}}$$

e

$$\alpha_2(\tau,h) := \alpha(\tau,h)\frac{h}{\sqrt{\tau^2+h^2}}$$

são infinitésimos quando $(\tau,h) \to (0,0)$ (porque as fracções $\frac{\tau}{\sqrt{\tau^2+h^2}}$ e $\frac{h}{\sqrt{\tau^2+h^2}}$ são limitadas). Tendo em conta a representação (2.34) podemos rescrever (2.33) como

$$\begin{aligned} & F(x_0+\tau, y_0+h) - F(x_0,y_0) \\ = \; & \big(F'_x(x_0,y_0) + \alpha_1(\tau,h)\big)\tau + \big(F'_y(x_0,y_0) + \alpha_2(\tau,h)\big)h. \end{aligned} \tag{2.35}$$

Consideramos o acrescimo de variável y na forma $h = f(x_0 + \tau) - f(x_0) \to 0$ se $\tau \to 0$. Então obviamente

$$F(x_0, y_0) = 0, \quad F(x_0 + \tau, y_0 + h) = F(x_0 + \tau, f(x_0 + \tau)) = 0,$$

e por isso a fórmula (2.35) dá

$$\left(F'_x(x_0,y_0)+\alpha_1(\tau,h)\right)\tau+\left(F'_y(x_0,y_0)+\alpha_2(\tau,h)\right)\left(f(x_0+\tau)-f(x_0)\right)=0,$$

ou equivalentemente

$$\frac{f(x_0+\tau)-f(x_0)}{\tau}=-\frac{F'_x(x_0,y_0)+\alpha_1(\tau,h)}{F'_y(x_0,y_0)+\alpha_2(\tau,h)}. \tag{2.36}$$

Passando ao limite em (2.36) quando $\tau\to0$ (logo também $h\to0$), imediatamente obtemos (2.32) (lembramos que as funções α_1 e α_2 são infinitésimas) ■

Notemos que a fórmula (2.32) dá o valor exacto da derivada $f'(x)$ (explicitamente através de x e y) apesar de não se conseguir obter o valor da função $f(x)$ de forma explicita.

Exemplo 2.33 *Seja*

$$F(x,y)=xy+\sin(x+y)+\cos(x+y)-1$$

uma função contínua e diferenciável em todo o plano $\mathbb{R}^2$*.* ***(Justifique!)*** *Temos* $F(0,0)=0$ *e a derivada parcial*

$$F'_y(x,y)=x+\cos(x+y)-\sin(x+y)$$

(também contínua) satisfaz $F'_y(0,0)=1\neq0$*. Logo, pelo Teorema da função implícita, existe uma função* $y=f(x)$ *definida numa vizinhança de zero,* $f(0)=0$*, que satisfaz a igualdade*

$$F(x,f(x))=0,$$

e com derivada

$$f'(x)=-\frac{y+\cos(x+y)-\sin(x+y)}{x+\cos(x+y)-\sin(x+y)}.$$

Em particular, $f'(0)=-1$ *(isto é, a recta* $y=-x$ *é tangente ao gráfico da função* f *no ponto* $(0,0)$*).*

A fórmula (2.32) permite escrever a equação da reta tangente ao gráfico da função definida implicitamente. Pois, da Análise Matemática I, sabemos que a equação da reta tangente ao gráfico de uma função $y=f(x)$ num ponto (x_0,y_0), $y_0=f(x_0)$, é dada por

$$y-y_0=f'(x_0)(x-x_0). \tag{2.37}$$

Se $y=f(x)$ é uma função definida implicitamente numa vizinhança do ponto (x_0,y_0) através da equação $F(x,y)=0$, supondo que $F'_y(x_0,y_0)\neq0$, então substituindo $f'(x_0)$ em (2.37) pela fórmula (2.32) obtemos a equação da reta tangente na forma simétrica:

$$\frac{\partial F}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial F}{\partial y}(x_0,y_0)(y-y_0)=0. \tag{2.38}$$

Observação 2.34 *Suponhamos que* F *é continuamente diferenciável numa vizinhança de* (x_0,y_0) *com* $F'_y(x_0,y_0)=0$ *mas* $F'_x(x_0,y_0)\neq0$*. Então, aplicando o teorema anterior, existe uma única*

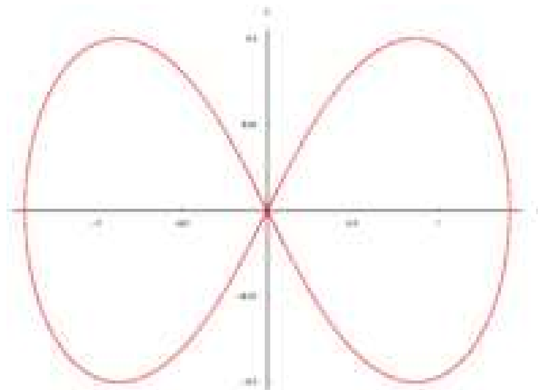

Figura 2.1:

função $x = g(y)$, contínua e diferenciável, definida implicitamente em torno de y_0, cuja derivada pode ser calculada pela fórmula (análoga a (2.32)):

$$g'(y_0) = -\frac{F'_y(x_0, y_0)}{F'_x(x_0, y_0)}.$$

Sendo $x - x_0 = g'(y_0)(y - y_0)$ a equação da reta tangente ao gráfico da função $x = g(y)$ no ponto (x_0, y_0), facilmente chegamos à forma simétrica dessa equação escrita em função das derivadas parciais de F, ou seja, à equação (2.38).

Se ambas as derivadas $F'_x(x_0, y_0)$ e $F'_y(x_0, y_0)$ são iguais a zero (isto é $\nabla F(x_0, y_0) = \mathbf{0}$) então no ponto (x_0, y_0) não existe uma reta tangente bem definida (tal situação acontece por exemplo no caso de autointersecções da curva definida pela equação $F(x, y) = 0$). Neste caso o ponto (x_0, y_0) diz-se *crítico.*

Exemplo 2.35 *Seja $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = \left(x^2 + y^2\right)^2 - 2\left(x^2 - y^2\right)$. A função F e as derivadas parciais dela são contínuas em todo plano, $F(0, 0) = 0$ mas*

$$\begin{aligned}
\frac{\partial F}{\partial x} &= 4x\left(x^2 + y^2 - 1\right), \ \frac{\partial F}{\partial x}(0, 0) = 0; \\
\frac{\partial F}{\partial y} &= 4y\left(x^2 + y^2 + 1\right), \ \frac{\partial F}{\partial y}(0, 0) = 0.
\end{aligned}$$

*Assim, $(0, 0)$ é o ponto crítico da curva definida a custa da equação (implicita) $F(x, y) = 0$. Esta curva chama-se **lemniscata de Bernoulli** e tem a forma apresentada na Fig. 2.1.*

O Teorema da função implícita (Teorema 2.32) pode ser extendido a funções de várias variáveis.

Teorema 2.36 *Seja $F : D \times \mathbb{R} \to \mathbb{R}$ ($D \subset \mathbb{R}^n$ um conjunto aberto) uma função contínua nalguma vizinhança do ponto $(\mathbf{a}, z_0) \in D \times \mathbb{R}$ tal que $F(\mathbf{a}, z_0) = 0$. Sem perda de generalidade, supomos que esta vizinhança tem a forma*

$$U = B_\eta(\mathbf{a}) \times \left]z_0 - \eta, z_0 + \eta\right[$$

para algum $\eta > 0$. Supomos também que a derivada parcial $F'_z(\mathbf{x}, z)$ existe e contínua em U e $F'_z(\mathbf{a}, z_0) \neq 0$. Então existem $\delta > 0$ e $\varepsilon > 0$ tais que $B_\delta(\mathbf{a}) \times]z_0 - \varepsilon, z_0 + \varepsilon[\subset U$ e para todo $\mathbf{x} \in B_\delta(\mathbf{a})$ existe um único $z \in]z_0 - \varepsilon, z_0 + \varepsilon[$ (seja $z = f(\mathbf{x})$, $f(\mathbf{a}) = z_0$) que satisfaz a igualdade $F(\mathbf{x}, f(\mathbf{x})) = 0$, isto é a função implícita $f(\mathbf{x})$ está bem definida na vizinhança $B_\delta(\mathbf{a})$. Esta função (de n variáveis) é contínua em $\mathbf{a}$. Além disso, se F é diferenciável em $(\mathbf{a}, z_0)$ então f é diferenciável em $\mathbf{a}$, e as derivadas parciais em ordem de x_i, $i = 1, ..., n$, são dadas por

$$f'_{x_i}(\mathbf{a}) = -\frac{F'_{x_i}(\mathbf{a}, z_0)}{F'_z(\mathbf{a}, z_0)}. \tag{2.39}$$

Usando o a fórmula (2.39) no caso $n = 2$ e $\mathbf{a} = (x_0, y_0)$ podemos obter a equação do plano tangente ao gráfico da função $z = f(x, y)$ definida implicitamente a custa da equação $F(x, y, z) = 0$ no ponto (x_0, y_0, z_0), $F(x_0, y_0, z_0) = 0$ (analogamente como fizemos acima no caso $n = 1$ para reta tangente a uma curva). Pois, recordando a definição do plano tangente na forma explicita (2.13) e substituindo os valores das derivadas, de (2.39) chegamos à equação simmétrica (compare com (2.38)):

$$\frac{\partial F}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0)(z - z_0) = 0. \tag{2.40}$$

Analogamente ao caso $n = 1$ se $F'_z(x_0, y_0, z_0) = 0$ mas uma das outras derivadas parciais ($F'_x(x_0, y_0, z_0)$ ou $F'_y(x_0, y_0, z_0)$) é diferente de zero, a equação $F(x, y, z) = 0$ pode ser escrita explicitamente em relação a x ou a y, respetivamente, e a equação do plano tangente (2.40) fica a mesma. Assim, (x_0, y_0, z_0) é um *ponto crítico* (isto é, é um ponto onde o plano tangente não está bem definido) apenas quando $\nabla F(x_0, y_0, z_0) = \mathbf{0}$.

As equações (2.38) e (2.40), obtidas acima, permitem justificar o sentido geométrico do vetor gradiente de uma função de duas ou três variáveis de que ainda falámos no final de Secção 2.2.

Recordemos que o *conjunto de nível* da função $f : D \subset \mathbb{R}^n \to \mathbb{R}$ é o conjunto

$$L(c) = \{\mathbf{x} \in D : f(\mathbf{x}) = c\},$$

onde c é um dos valores que f toma (por outras palavras, c pertence ao contradomínio de f). No caso $n = 2$ ou $n = 3$ os conjuntos de nível chamam-se, como sabemos, respetivamente, *linhas* (ou curvas) *de nível* e *superfícies de nível*.

Teorema 2.37 *Seja $f : D \subset \mathbb{R}^n \to \mathbb{R}$ ($n = 2$ ou 3) uma função contínua numa vizinhança de $\mathbf{a} \in \operatorname{int} D$, diferenciável nesse ponto e tal que $\nabla f(\mathbf{a}) \neq \mathbf{0}$. Então $\nabla f(\mathbf{a})$ é um vector ortogonal à reta (ou ao plano tangente) à linha (resp., superfície) de nível $L(c)$ onde $c = f(\mathbf{a})$, no ponto $\mathbf{a}$.*

Demonstração:
Consideramos o caso $n = 2$. O caso três-dimensional considera-se analogamente. Seja $\mathbf{a} = (x_0, y_0) \in \operatorname{int} D$. A linha de nível $c := f(x_0, y_0)$ tem a forma

$$L(c) = \left\{(x, y) \in \mathbb{R}^2 : f(x, y) = c\right\}.$$

A função $F(x, y) = f(x, y) - c$ é contínua numa vizinhança de $\mathbf{a}$, diferenciável em $\mathbf{a}$ e pelo menos uma das suas derivadas parciais $F'_x(x_0, y_0)$ ou $F'_y(x_0, y_0)$ é diferente de zero (**justifique!**).

Conforme Teorema 2.32 e Observação 2.34 a linha de nível $L(c)$ pode ser considerada localmente (numa vizinhança do ponto (x_0, y_0)) como o gráfico de uma função implícita g de uma variável (x ou y). Em qualquer dos casos (2.38) é a equação da reta tangente ao gráfico da função g no ponto $\mathbf{a} = (x_0, y_0)$. Na forma vetorial esta equação pode ser escrita como

$$\langle \nabla F(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle = 0,$$

isto é $\nabla F(\mathbf{a}) = \nabla f(\mathbf{a})$ é ortogonal aos vetores $\mathbf{x} - \mathbf{a}$ que pertencem à reta tangente. ∎

Exemplo 2.38 *Seja $f : \mathbb{R}^3 \to \mathbb{R}$ a função definida por*

$$f(x, y, z) = x^2 + y^2 + z^2.$$

Determinemos um vetor ortogonal à superfície de nível de f que passa pelo ponto $(1, 2, 3)$, a equação dessa superfície e a equação do plano tangente a essa superfície no referido ponto.

Como $f(1, 2, 3) = 14$, a superfície pedida é definida pela equação $x^2 + y^2 + z^2 = 14$ sendo uma esfera centrada na origem de raio $\sqrt{14}$. Assim, um vetor ortogonal a esta esfera no ponto $(1, 2, 3)$ é $\nabla f(1, 2, 3) = (2, 4, 6)$, e a equação do plano tangente pedida é

$$2(x - 1) + 4(y - 2) + 6(z - 3) = 0,$$

ou seja, $x + 2y + 3z = 14$.

2.5 Aplicações vetoriais diferenciáveis

A teoria da derivação para funções vetoriais é uma extensão direta da das funções escalares. Por exemplo, dados uma aplicação $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ e um ponto $\mathbf{a} \in \operatorname{int} D$ a *derivada direccional* $f'(\mathbf{a}; \mathbf{v})$ *segundo um vetor* $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$, calcula-se pela mesma fórmula (2.1) se existir. Só limite nela é o limite de uma aplicação vetorial (que admete valores no espaço $\mathbb{R}^m$) de uma variável escalar.

Da definição logo sai que as derivadas direccionais de uma aplicação $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ podem ser consideradas por coordenadas. Nomeadamente, dados $\mathbf{a} \in \operatorname{int} D$ e $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$, tem-se

$$f'(\mathbf{a}; \mathbf{v}) = \left(f_1'(\mathbf{a}; \mathbf{v}), f_2'(\mathbf{a}; \mathbf{v}), ..., f_m'(\mathbf{a}; \mathbf{v})\right)$$

onde $f_1, f_2, ..., f_m$ são as funções-coordenadas da aplicação f. Assim a derivada $f'(\mathbf{a}; \mathbf{v})$ é um vetor no espaço $\mathbb{R}^m$.

Como vemos a seguir o conceito da derivada parcial para aplicações vetoriais não é muito útil. Por isso, passamos logo à definição de *aplicação diferenciável*.

Definição 2.39 *Diz-se que a aplicação $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ é* diferenciável *num ponto $\mathbf{a} \in \operatorname{int} D$ se existe uma aplicação linear $\mathbf{h} \mapsto \mathbf{A}\mathbf{h}$ ($\mathbf{A}$ é uma $m \times n$-matriz) tal que*

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \mathbf{A}\mathbf{h} + o(\mathbf{h}), \tag{2.41}$$

para $\mathbf{h}$ com a norma suficientemente pequena (pelo menos tal que $\mathbf{a} + \mathbf{h} \in D$). Aqui $o(\mathbf{h})$ é uma aplicação $\mathbb{R}^n \to \mathbb{R}^m$ definida numa vizinhança de 0 e tal que

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

Assim vemos que a definição de diferenciabilidade no caso de aplicações vetoriais é absolutamente mesma que para função escalar (acrescimo da aplicação representa-se como soma de uma parte linear e um infinitésimo de ordem superior a 1). A diferença só é na interpretação da linearidade. No caso de uma função escalar trata-se de uma *forma linear* (uma função linear $\mathbb{R}^n \to \mathbb{R}$ que está associada com um vetor fixo $\boldsymbol{\xi} \in \mathbb{R}^n$). Em vez no caso $m > 1$ temos já a parte linear na forma de uma aplicação $\mathbb{R}^n \to \mathbb{R}^m$ que representa-se a custa de uma matriz $\mathbf{A}$ de ordem $m \times n$, e em vez do *produto interno* $\langle \boldsymbol{\xi}, \mathbf{h} \rangle$ temos aqui o *produto de matrizes* $\mathbf{Ah}$ (ver Capítulo 1).

Escrevendo a igualdade (2.41) por coordenadas:

$$f_i(\mathbf{a}+\mathbf{h}) - f_i(\mathbf{a}) = \sum_{j=1}^{n} \xi_{ij} h_j + o_i(\mathbf{h})$$

(aqui $\mathbf{A} = [\xi_{ij}]_{ij}$ e $o_i(\mathbf{h}) / \|\mathbf{h}\| \to 0$ se $\mathbf{h} \to \mathbf{0}$) observamos que uma aplicação $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ é *diferenciável* num ponto $\mathbf{a} \in \operatorname{int} D$ sse todas as funções-coordenadas $f_1, f_2, ..., f_m$ são diferenciáveis em $\mathbf{a}$. Além disso (ver Teorema 2.11 (ii)) $\xi_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{a})$. Por outras palavras, as linhas da matriz $\mathbf{A}$ são gradientes das funções f_i $(i = 1, 2, ..., n)$ em ponto em causa. Tal matriz $\mathbf{A}$ chama-se *matriz jacobiana* (ou *matriz de Jacobi*) de f em $\mathbf{a}$ e denota-se por $\operatorname{Jac} f(\mathbf{a})$:

$$\operatorname{Jac} f(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{bmatrix}.$$

Se $n = m$ a matriz de Jacobi torna-se quadrática e o determinante dela chama-se *jacobiano* (denota-se por $J(f)(\mathbf{a})$ ou $\frac{\partial(f_1, f_2, ..., f_n)}{\partial(x_1, x_2, ..., x_n)}(\mathbf{a})$).

De mesma forma como para funções escalares (ver Teorema 2.15) uma aplicação $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ é diferenciável em $\mathbf{a} \in \operatorname{int} D$ se todas as derivadas parciais $\frac{\partial f_i}{\partial x_j}$ $(i = 1, ..., n; j = 1, ..., m)$ existem em torno do ponto $\mathbf{a}$ e são contínuas em $\mathbf{a}$. Neste caso dizemos também que f é aplicação *continuamente diferenciável* em $\mathbf{a}$.

2.6 Derivada da aplicação composta

Usando a *regra de cadeia* obtemos uma fórmula de derivação de aplicações diferenciáveis que generaliza a fórmula de derivada da função composta no caso de uma variável.

Teorema 2.40 *Sejam $D \subset \mathbb{R}^p$ e $W \subset \mathbb{R}^n$ conjuntos abertos, $f : D \to \mathbb{R}^m$ e $g : W \to \mathbb{R}^p$ aplicações tais que $g(W) \subset D$. Se g é diferenciável em $\mathbf{a} \in W$ e f é diferenciável no ponto $g(\mathbf{a})$, então a aplicação composta $\varphi : W \to \mathbb{R}^m$, $\varphi(\mathbf{x}) = f(g(\mathbf{x})) = (f \circ g)(\mathbf{x})$ é diferenciável em $\mathbf{a}$ e tem-se*

$$\operatorname{Jac} \varphi(\mathbf{a}) = \operatorname{Jac} f(g(\mathbf{a})) \cdot \operatorname{Jac} g(\mathbf{a}), \tag{2.42}$$

onde a parte direita da igualdade representa o produto de matrizes.

Demonstração:

Representamos a função vetorial φ pelas suas coordenadas $(\varphi_1, ..., \varphi_m)$. A diferenciabilidade de todas as funções

$$\varphi_i(x_1, ..., x_n) = f_i(g_1(x_1, ..., x_n), ..., g_p(x_1, ..., x_n))$$

sai da diferenciabilidade da função composta (ver Secção 2.3). Para calcular uma das derivadas parciais de φ_i usamos a regra da cadeia (ver (2.23)):

$$\frac{\partial \varphi_i}{\partial x_j}(\mathbf{a}) = \sum_{k=1}^{p} \frac{\partial f_i}{\partial g_k}(g(\mathbf{a})) \frac{\partial g_k}{\partial x_j}(\mathbf{a})$$

e vemos que na parte direita desta fórmula está o produto da i-ésima linha da matriz Jacobiana de f calculada em $g(\mathbf{a})$ pela j-ésima coluna da matriz Jacobiana de g calculada em $\mathbf{a}$. Daqui sai a fórmula (2.42). ■

Exemplo 2.41 *Consideremos as funções* $f(u,v) = (\ln u, ve^{-u})$ *e* $g(x,y,z) = (x^2+1, y+z)$. *Temos*

$$D_f =]0, +\infty[\times \mathbb{R}, \quad D_g = \mathbb{R}^3 \quad e \quad D_{f \circ g} = \mathbb{R}^3.$$

A matriz jacobiana de f *num ponto* $(u,v) \in D_f$ *é*

$$\operatorname{Jac} f(u,v) = \begin{bmatrix} \frac{1}{u} & 0 \\ -ve^{-u} & e^{-u} \end{bmatrix}.$$

Substituindo $(u,v) = g(x,y,z)$ *temos*

$$\operatorname{Jac} f(g(x,y,z)) = \operatorname{Jac} f(x^2+1, y+z) = \begin{bmatrix} \frac{1}{x^2+1} & 0 \\ -(y+z)e^{-(x^2+1)} & e^{-(x^2+1)} \end{bmatrix}.$$

Derivando agora a aplicação $g : \mathbb{R}^3 \to \mathbb{R}^2$ *encontramos*

$$\operatorname{Jac} g(x,y,z) = \begin{bmatrix} 2x & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Então a matriz jacobiana de $f \circ g$ *num ponto* (x,y,z) *é o produto das duas últimas matrizes:*

$$\begin{aligned} \operatorname{Jac} f(g(x,y,z)) &= \begin{bmatrix} \frac{1}{x^2+1} & 0 \\ -(y+z)e^{-(x^2+1)} & e^{-(x^2+1)} \end{bmatrix} \begin{bmatrix} 2x & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{2x}{x^2+1} & 0 & 0 \\ -2x(y+z)e^{-(x^2+1)} & e^{-(x^2+1)} & e^{-(x^2+1)} \end{bmatrix}. \end{aligned}$$

2.7 Teorema da aplicação implícita

O Teorema da função implícita (Teorema 2.32) pode ser extendido a aplicações vetoriais. Nomeadamente, aqui trata-se não **uma** equação $F(x,y)=0$ ou $F(x_1,...,x_n;y)=0$ que implicitamente define **uma** função y (dependente de uma ou de várias variáveis) mas um sistema de equações

$$\left\{\begin{array}{l} F_1(x_1,...,x_n;y_1,...,y_m)=0; \\ F_2(x_1,...,x_n;y_1,...,y_m)=0; \\ \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ F_m(x_1,...,x_n;y_1,...,y_m)=0, \end{array}\right. \tag{2.43}$$

que (também implicitamente) define m funções escalares y_j $((j=1,...,m)$ dependentes das variáveis $x_1,...,x_n$. Nota-se que o sistema acima pode ser considerado como uma equação $F(\mathbf{x};\mathbf{y})=0$ onde $F:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$ é uma aplicação vetorial com as funções-coordenadas $F_1,...,F_2$; $\mathbf{x}=(x_1,...,x_n)$ e $\mathbf{y}=(y_1,...,y_m)$. Assim pretendemos definir localmente (nalguma vizinhança de um ponto dado $\mathbf{a}$) uma aplicação vetorial $\mathbf{y}=\mathbf{y}(\mathbf{x})$ (com m coordenadas) tal que $F(\mathbf{x};\mathbf{y}(\mathbf{x}))=0$. Observamos que o número das equações em sistema (2.43) tem regorosamente de coincidir com o número de variáveis dependentes.

Demonstração do seguinte teorema pode ser encontrada em qualquer livro do Cálculo (por exemplo, em [4] ou em **??**).

Teorema 2.42 *Seja $F:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$ uma aplicação vetorial definida no conjunto $\mathbf{X}\times\mathbf{Y}$ onde $\mathbf{X}\subset\mathbb{R}^n$ e $\mathbf{Y}\subset\mathbb{R}^n$ são abertos e continuamente diferenciáveis numa vizinhança de um ponto $(\mathbf{x}^0,\mathbf{y}^0)\in\mathbf{X}\times\mathbf{Y}$, tal que $F(\mathbf{x}^0,\mathbf{y}^0)=0$. Supomos também que*

$$\frac{\partial(F_1,...,F_m)}{\partial(y_1,...,y_m)}\left(\mathbf{x}^0,\mathbf{y}^0\right)\neq 0$$

*onde $F_1,...,F_m$ são funções-coordenadas de F. Então existem vizinhanças $U(\mathbf{x}^0)$ e $V(\mathbf{y}^0)$ nos espaços $\mathbb{R}^n$ e $\mathbb{R}^m$, respetivamente, tais que para a cada $\mathbf{x}\in U(\mathbf{x}^0)$ pode ser associado **um e só um** $\mathbf{y}\in V(\mathbf{y}^0)$ com $F(\mathbf{x},\mathbf{y})=0$. A função implicita $f:U(\mathbf{x}^0)\to V(\mathbf{y}^0)$, $\mathbf{y}^0=f(\mathbf{x}^0)$, assim definida é continuamente diferenciável, e a matriz jacobiana de f em $\mathbf{x}^0$ determina-se pela fórmula seguinte:*

$$\operatorname{Jac} f\left(\mathbf{x}^0\right)=-\left(\operatorname{Jac}_{\mathbf{y}} F\left(\mathbf{x}^0,\mathbf{y}^0\right)\right)^{-1}\cdot\operatorname{Jac}_{\mathbf{x}} F\left(\mathbf{x}^0,\mathbf{y}^0\right) \tag{2.44}$$

onde $\operatorname{Jac}_{\mathbf{x}} F(\mathbf{x}^0,\mathbf{y}^0)$ e $\operatorname{Jac}_{\mathbf{y}} F(\mathbf{x}^0,\mathbf{y}^0)$ são matrizes jaconianas da aplicação $F(\mathbf{x},\mathbf{y})$ em ordem as variáveis x_i $(i=1,...,n)$ e y_j $(j=1,...,m)$, respetivamente.

Exemplo 2.43 $(n=m=2)$ *Mostremos que as equações*

$$\left\{\begin{array}{l} x^3u+yv^3=2; \\ xu^3+y^3v=2 \end{array}\right. \tag{2.45}$$

definem u e v implicitamente como funções de x e y numa vizinhança do ponto $(1,1;1,1)$, isto é que a cada par (x,y) próximo de $(1,1)$ corresponde um único (u,v),

$$\begin{array}{rcl} u &=& u(x,y); \\ v &=& v(x,y), \end{array}$$

na vizinhança de $(1,1)$ *tal que o quadruplo* $(x,y;u,v)$ *satisfaz às equações (2.45). Além disso, calculemos* $\operatorname{Jac} f(x,y)$ *no ponto* $(x,y)=(1,1)$ *onde* $f(x,y)=(u(x,y),v(x,y))$.

Seja $F:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2$ *definida por*

$$F(x,y;u,v)=\left(x^3u+yv^3-2,\,xu^3+y^3v-2\right),$$

e sejam

$$\begin{aligned} F_1(x,y;u,v) &= x^3u+yv^3-2;\\ F_2(x,y;u,v) &= xu^3+y^3v-2. \end{aligned}$$

A aplicação F *é continuamente diferenciável (**de classe** $\mathcal{C}^1$) em* $\mathbb{R}^2\times\mathbb{R}^2$, $F(1,1;1,1)=0$. *Determinemos o Jacobiano da aplicação* F *em relação às variáveis* u *e* v:

$$\frac{\partial(F_1,F_2)}{\partial(u,v)}(1,1,1,1)=\left|\begin{array}{cc} \frac{\partial F_1}{\partial u}(x,y;u,v) & \frac{\partial F_1}{\partial v}(x,y;u,v)\\ \frac{\partial F_2}{\partial u}(x,y;u,v) & \frac{\partial F_2}{\partial v}(x,y;u,v)\end{array}\right|_{(1,1,1,1)}$$

$$=\left|\begin{array}{cc} x^3 & 3yv^2\\ 3xu^2 & y^3\end{array}\right|_{(1,1,1,1)}=\left|\begin{array}{cc} 1 & 3\\ 3 & 1\end{array}\right|=-8\neq 0.$$

Portanto, pelo Teorema 2.42 a aplicação implicita $(u,v)=f(x,y)$ *é bem definida e continuamente diferenciável nalguma vizinhança de* $(x,y,u,v)=(1,1,1,1)$. *Determinemos agora* $\operatorname{Jac} f(x,y)$ *no ponto* $(x,y)=(1,1)$ *pela fórmula (2.44). Nomeadamente,*

$$\begin{aligned}\operatorname{Jac} f(x,y) &= -\left[\begin{array}{cc} \frac{\partial F_1}{\partial u}(x,y;u,v) & \frac{\partial F_1}{\partial v}(x,y;u,v)\\ \frac{\partial F_2}{\partial u}(x,y;u,v) & \frac{\partial F_2}{\partial v}(x,y;u,v)\end{array}\right]^{-1}\times\\ &\quad\times\left[\begin{array}{cc} \frac{\partial F_1}{\partial x}(x,y;u,v) & \frac{\partial F_1}{\partial y}(x,y;u,v)\\ \frac{\partial F_2}{\partial x}(x,y;u,v) & \frac{\partial F_2}{\partial y}(x,y;u,v)\end{array}\right]. \end{aligned}\tag{2.46}$$

Acima nos já vemos que no ponto $(1,1)$

$$\left[\begin{array}{cc} \frac{\partial F_1}{\partial u}(x,y;u,v) & \frac{\partial F_1}{\partial v}(x,y;u,v)\\ \frac{\partial F_2}{\partial u}(x,y;u,v) & \frac{\partial F_2}{\partial v}(x,y;u,v)\end{array}\right]=\left[\begin{array}{cc} 1 & 3\\ 3 & 1\end{array}\right],$$

e a matriz inversa

$$\left[\begin{array}{cc} 1 & 3\\ 3 & 1\end{array}\right]^{-1}=\left[\begin{array}{cc} -\frac{1}{8} & \frac{3}{8}\\ \frac{3}{8} & -\frac{1}{8}\end{array}\right].$$

De outro lado,

$$\left[\begin{array}{cc} \frac{\partial F_1}{\partial x}(x,y;u,v) & \frac{\partial F_1}{\partial y}(x,y;u,v)\\ \frac{\partial F_2}{\partial x}(x,y;u,v) & \frac{\partial F_2}{\partial y}(x,y;u,v)\end{array}\right]_{(1,1,1,1)}$$

$$=\left[\begin{array}{cc} 3x^2u & v^3\\ u^3 & 3y^2v\end{array}\right]_{(1,1,1,1)}=\left[\begin{array}{cc} 3 & 1\\ 1 & 3\end{array}\right].$$

Substituindo em (2.46) obtemos

$$\begin{aligned}\operatorname{Jac} f(1,1) &= -\left[\begin{array}{cc} -\frac{1}{8} & \frac{3}{8}\\ \frac{3}{8} & -\frac{1}{8}\end{array}\right]\cdot\left[\begin{array}{cc} 3 & 1\\ 1 & 3\end{array}\right]\\ &= \left[\begin{array}{cc} 0 & -1\\ -1 & 0\end{array}\right]. \end{aligned}$$

Matriz jacobiana pode ser encontrada de outra maneira (como alternativa do cálculo de matrizes). Nomeadamente, aplicamos a regra da cadeia às equações

$$\begin{aligned} F_1(x,y,u,v) &= x^3u + yv^3 - 2 = 0; \\ F_2(x,y,u,v) &= xu^3 + y^3v - 2 = 0, \end{aligned}$$

onde $u = u(x,y)$ *e* $v = v(x,y)$, *ou seja, às equações*

$$\begin{aligned} x^3u(x,y) + yv^3(x,y) - 2 &= 0; \\ xu^3(x,y) + y^3v(x,y) - 2 &= 0. \end{aligned} \tag{2.47}$$

Derivando em ordem a x *obtemos o seguinte sistema de duas equações lineares (em relação às derivadas parciais* $\frac{\partial u}{\partial x}$ *e* $\frac{\partial v}{\partial x}$*):*

$$\begin{cases} 3x^2u + x^3\frac{\partial u}{\partial x} + 3yv^2\frac{\partial v}{\partial x} = 0; \\ u^3 + 3xu^2\frac{\partial u}{\partial x} + y^3\frac{\partial v}{\partial x} = 0. \end{cases}$$

Tendo em conta que $u(1,1) = v(1,1) = 1$, *resulta*

$$\begin{cases} \frac{\partial u}{\partial x}(1,1) + 3\frac{\partial v}{\partial x}(1,1) = -3; \\ 3\frac{\partial u}{\partial x}(1,1) + \frac{\partial v}{\partial x}(1,1) = -1. \end{cases}$$

Resolvendo este sistema, obtemos

$$\begin{cases} \frac{\partial u}{\partial x}(1,1) = 0; \\ \frac{\partial v}{\partial x}(1,1) = -1. \end{cases}$$

Analogamente, derivando as equações (2.47) em ordem a y *e procedendo de forma análoga, obtemos*

$$\begin{cases} \frac{\partial u}{\partial y}(1,1) = -1; \\ \frac{\partial v}{\partial y}(1,1) = 0. \end{cases}$$

Assim,

$$\operatorname{Jac} f(1,1) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

2.8 Aplicação inversa

Seja $f : D \to \mathbb{R}^n$ uma aplicação vetorial onde $D \subset \mathbb{R}^n$ e $E := f(D) \subset \mathbb{R}^n$ são conjuntos abertos. A aplicação inversa f^{-1} (se existir) associa a qualquer $\mathbf{y} \in E$ um ponto $\mathbf{x} \in D$ tal que $f(\mathbf{x}) = \mathbf{y}$ ou

$$F(\mathbf{x},\mathbf{y}) := f(\mathbf{x}) - \mathbf{y} = \mathbf{0}.$$

Logo, f^{-1} pode ser vista como uma função implicita. Dados $\mathbf{x}_0 \in D$ e $\mathbf{y}_0 = f(\mathbf{x}_0)$ supomos que f é continuamente diferenciável em torno de $\mathbf{x}_0$. Então pelo Teorema 2.42 de aplicações implicitas f^{-1} **existe e continuamente diferenciável** numa vizinhança de $\mathbf{y}_0$ se o jacobiano $J(f)(\mathbf{x}_0) \neq 0$. Se f é de classe $\mathcal{C}^1$ em todo domínio D e $J(f) \neq 0$ em D, então f^{-1} é de classe

$\mathcal{C}^1$ em E. Neste caso a aplicação f diz-se um *difeomorfismo* em D. A matriz jacobiana dela encontra-se diretamente da igualdade

$$f \circ f^{-1} = \mathcal{I} \tag{2.48}$$

onde $\mathcal{I}$ é a aplicação identidade em $\mathbb{R}^n$, isto é $\mathcal{I}(\mathbf{x}) = \mathbf{x} = (x_1, ..., x_n)$, $\mathbf{x} \in \mathbb{R}^n$. Aplicando a fórmula (2.42) e tomando em conta que $\operatorname{Jac} \mathcal{I}(\mathbf{x}) = I_n$ é a matriz unitária de ordem n,

$$I_n = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ .. & .. & .. & .. \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

da (2.48) obtemos

$$\operatorname{Jac}(f) \operatorname{Jac}\left(f^{-1}\right) = I_n$$

donde

$$\operatorname{Jac} f^{-1} = (\operatorname{Jac} f)^{-1}$$

(matrizes jacobianas das aplicações inversas são inversas no sentido de âlgebra matricial).

Bibliografia

[1] J.Campos Ferreira, *Introdução à Análise em* $\mathbb{R}^n$, DMIST, 2003.

[2] R. Courant, *Differential and Integral Calculus*, Vol. 2, Chap. 1-5, Wiley-Interscience, 1988.

[3] B. Demidovich, Problemas e exercícios de Análise Matemática, Editora McGraw-Hill de Portugal, 1993.

[4] F. R. Dias Agudo, *Análise Real*, vol. 3, Escolar Editora, 1990.

[5] M. Ferreira e I. Amaral, *Matemática, Cálculo diferencial em* $\mathbb{R}^n$, Edição Sílabo, 1991.

[6] M. Krasnov, A. Kiselev, G. makarenko, E. Shikin, *Mathematical Analysis for Engineers*, Vol. 2, Chap. 22-24, Mir Publisher, Moscow, 1990.

[7] E. Lages Lima, *Curso de Análise*, vol. 2, Projecto Euclides, 1995.

[8] S. Lan, *Calculus of Several Variables*, Springer-Verlag, 1978.

[9] L. T. Magalhães, *Álgebra Linear como Introdução à Matemática Aplicada*, Texto Editora, 1997.

[10] J. Marsden, A. Tromba, *Vector Calculus*, W. H. Freeman and co, 1976.

[11] A. Ostrowski, *Lições de Cálculo Diferencial e Integral*, vol. 2, 4^a edição, Fundação Caloustre Gilbenkian,1987.

[12] G. Pires, *Cálculo Diferencial e Integral em* $\mathbb{R}^n$, IST Press, 2012.

[13] C. Sarrico, *Cálculo Diferencial e Integral para Funções de Várias Variáveis*, Esfera do caos, 2003.

[14] J. Stewart, *Calculus*, Concepts and Contexts, 2nd Edition, Thomson Learning, 2005.

[15] E. Swokowski, *Cálculo com Geometria Analítica*, vol. 2, 2^a edição, Makron Books do Brasil Editora.