

Índice

Índice	1
1 Funções de várias variáveis	2
1.1 O espaço vectorial \mathbb{R}^n	2
1.2 Noções topológicas em \mathbb{R}^n	10
1.3 Sucessões em \mathbb{R}^n	16
1.4 Funções. Limite e continuidade	17
2 Derivadas de funções escalares	31
2.1 Derivadas direccionais e derivadas parciais	31
2.2 Diferenciabilidade e diferencial	34
2.3 Gradiente	42
2.4 Derivada da função composta	43
2.5 Teorema da função implícita	46
2.6 Interpretação geométrica do gradiente	50
Bibliografia	52

Capítulo 1

Funções de várias variáveis

Antes de começar com o estudo das funções propriamente ditas vamos recordar alguns conceitos de Álgebra Linear (Secção 1.1), introduzir o conceito de vizinhança e a partir deste outras noções topológicas em \mathbb{R}^n (Secção 1.2). Estas duas primeiras secções não são só fundamentais para o estudo das funções mas como para toda a Análise Matemática II.

1.1 O espaço vectorial \mathbb{R}^n

Sendo n um número natural, o conjunto

$$\mathbb{R}^n := \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R},$$

ou seja, o produto cartesiano de n factores iguais a \mathbb{R} , é constituído por todos os elementos da forma (x_1, x_2, \dots, x_n) , onde cada x_i , com $i = 1, 2, \dots, n$, é um número real. Representamos os elementos de \mathbb{R}^n por uma única letra, a negrito, por exemplo $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Dados $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ em \mathbb{R}^n dizemos que $\mathbf{x} = \mathbf{y}$ sse $x_i = y_i$ para $i = 1, 2, \dots, n$. Assim, cada elemento $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ determina, de forma unívoca, cada uma das suas coordenadas x_1, x_2, \dots, x_n (designadas, respectivamente, por 1ª, 2ª, ..., n-ésima coordenada de \mathbf{x}).

Seguindo a notação usual, quando $n = 2$ ou $n = 3$, usaremos, na maioria das vezes, (x, y) e (x, y, z) em vez de (x_1, x_2) e (x_1, x_2, x_3) , respectivamente.

Dados $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ em \mathbb{R}^n , e α um número real, definimos a soma $\mathbf{x} + \mathbf{y}$ e o produto $\alpha\mathbf{x}$ por

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).$$

O conjunto \mathbb{R}^n quando munido destas operações, é um *espaço vectorial real*. Os elementos de \mathbb{R}^n chamam-se *pontos* ou *vectores* e os números reais chamam-se *escalares*. O elemento com as n coordenadas iguais a zero, $\mathbf{0} = (0, 0, \dots, 0)$, chama-se *zero* de \mathbb{R}^n ou *vector nulo*. Geometricamente, considerar $\mathbf{x} \in \mathbb{R}^n$, $n = 2$ ou $n = 3$, como um vector significa pensar na seta que tem origem no ponto $\mathbf{0}$ e extremidade no ponto \mathbf{x} .

O espaço vectorial \mathbb{R}^n tem dimensão n , sendo a sua *base canónica* dada por $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$, onde $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, \dots, 0)$, ..., $\mathbf{e}_n = (0, 0, \dots, 1)$. Dado $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ tem-se

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n.$$

No caso $n = 2$ ou $n = 3$ denotamos $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ por $\mathbf{i}, \mathbf{j}, \mathbf{k}$, respectivamente.

Introduzimos agora uma outra operação de "multiplicação".

Produto interno. Norma

Definição 1.1 Sendo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$, chama-se produto interno (ou produto escalar) de \mathbf{x} e \mathbf{y} , e denota-se por $\langle \mathbf{x}, \mathbf{y} \rangle$, o número real

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + \dots + x_ny_n.$$

As propriedades principais do produto interno são:

Proposição 1.2 Sejam $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Então tem-se

- I1) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$;
- I2) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ e $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$;
- I3) $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha\mathbf{y} \rangle$;
- I4) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ e $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ sse $\mathbf{x} = \mathbf{0}$.

Demonstração: Exercício.

Com base na noção de produto interno introduzimos um outro conceito, fundamental em tudo o que segue - a norma de um vector em \mathbb{R}^n - e que generaliza a noção de módulo de um número real.

Definição 1.3 Sendo $\mathbf{x} \in \mathbb{R}^n$ chamamos norma euclídeana de \mathbf{x} , e designamos pelo símbolo $\|\mathbf{x}\|_n$, o número real

$$\|\mathbf{x}\|_n = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Observação 1.4 Quando não houver perigo de confusão em relação ao conjunto a que pertence o vector \mathbf{x} pode-se escrever, simplesmente, $\|\mathbf{x}\|$ em vez de $\|\mathbf{x}\|_n$

Proposição 1.5 Sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Então tem-se

- N1) $\|\mathbf{x}\| \geq 0$;
- N2) $\|\mathbf{x}\| = 0$ sse $\mathbf{x} = \mathbf{0}$;
- N3) $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$;
- N4) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (desigualdade triangular).

Demonstração: Ver [13, p. 4].

Observação 1.6 *É possível definir outras normas em \mathbb{R}^n com as propriedades acabadas de referir como, por exemplo, a norma da soma e a norma do máximo, definidas, respectivamente, por*

$$\|\mathbf{x}\|_+ = |x_1| + \dots + |x_n| \quad e \quad \|\mathbf{x}\|_M = \max\{|x_1|, \dots, |x_n|\}.$$

Temos a seguinte relação

$$\|\mathbf{x}\|_M \leq \|\mathbf{x}\| \leq \|\mathbf{x}\|_+ \quad (1.1)$$

qualquer que seja $\mathbf{x} \in \mathbb{R}^n$ (justifique!).

Nestes apontamentos iremos usar apenas a norma euclideana, à qual chamaremos simplesmente *norma*.

A norma verifica também as seguintes propriedades, sendo a segunda uma consequência imediata da desigualdade triangular.

Proposição 1.7 *Sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Então tem-se*

(i) $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ (desigualdade de Cauchy-Schwarz);

(ii) $||\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$.

Demonstração: Ver [13, p. 4].

Definição 1.8 *Sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ dois vectores não nulos. O número $\theta \in [0, \pi]$ definido por*

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

diz-se o ângulo entre os vectores \mathbf{x} e \mathbf{y} .

Observação 1.9 *Assim, o produto interno entre dois vectores \mathbf{x} e \mathbf{y} de \mathbb{R}^n , $n = 2, 3$, pode ser interpretado geometricamente da seguinte forma*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta,$$

sendo $\theta \in [0, \pi]$ o ângulo entre os vectores \mathbf{x} e \mathbf{y} .

Definição 1.10 *Diz-se que os vectores \mathbf{x} e \mathbf{y} de \mathbb{R}^n são ortogonais (ou perpendiculares), e escreve-se $\mathbf{x} \perp \mathbf{y}$, se $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.*

Distância

Definição 1.11 *Sejam $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ pontos de \mathbb{R}^n . A distância euclideana (ou simplesmente distância) entre \mathbf{x} e \mathbf{y} , denotada $d(\mathbf{x}, \mathbf{y})$, é dada por*

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

Observação 1.12 Tendo em conta a definição de norma de um vector, a distância entre \mathbf{x} e \mathbf{y} é dada por

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}.$$

Proposição 1.13 Sejam $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. Então tem-se

D1) $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (desigualdade triangular);

D2) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;

D3) $d(\mathbf{x}, \mathbf{y}) \geq 0$ e $d(\mathbf{x}, \mathbf{y}) = 0$ sse $\mathbf{x} = \mathbf{y}$.

Demonstração: Exercício.

Observação 1.14 Assim, a norma de um vector $\mathbf{x} \in \mathbb{R}^n$ é a distância de \mathbf{x} à origem.

Em \mathbb{R}^3 podemos definir mais dois produtos entre vectores, nomeadamente, o produto externo e o produto misto.

Produto externo

Definição 1.15 O produto externo (também chamado produto vectorial) de dois vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, denotado $\mathbf{x} \times \mathbf{y}$, é um vector $\mathbf{z} \in \mathbb{R}^3$ tal que

- (i) $\langle \mathbf{z}, \mathbf{x} \rangle = 0$ e $\langle \mathbf{z}, \mathbf{y} \rangle = 0$ (isto é, \mathbf{z} é ortogonal a \mathbf{x} e a \mathbf{y});
- (ii) $\|\mathbf{z}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$, onde $\theta \in [0, \pi]$ é o ângulo entre \mathbf{x} e \mathbf{y} ;
- (iii) um observador colocado na posição terminal do vector \mathbf{z} observa a rotação mais curta de \mathbf{x} para \mathbf{y} no sentido contrário dos ponteiros do relógio.

Exemplo 1.16 Em particular, temos $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

O produto externo goza das seguintes propriedades.

Proposição 1.17 Sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. Então tem-se

- E1) $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$ (anti-simetria);
- E2) para cada \mathbf{y} fixo a função $\mathbf{x} \mapsto \mathbf{x} \times \mathbf{y}$ é linear;
- E3) $\mathbf{x} \times \mathbf{y} = \mathbf{0}$ sse \mathbf{x} e \mathbf{y} são colineares (isto é, $\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}$ para alguns $\alpha, \beta \in \mathbb{R}$ não nulos simultaneamente);
- E4) $\|\mathbf{x} \times \mathbf{y}\|$ é igual à área do paralelogramo que tem como lados os vectores \mathbf{x} e \mathbf{y} ;
- E5) $\|\mathbf{x} \times \mathbf{y}\| = \sqrt{\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle^2}$.

Demonstração: Ver, por exemplo, [9, p. 217].

Na prática, para $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ o vector $\mathbf{x} \times \mathbf{y}$ pode obter-se calculando o seguinte determinante (simbólico):

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}.$$

Ou seja,

$$\mathbf{x} \times \mathbf{y} = (x_2y_3 - x_3y_2)\mathbf{i} + (x_3y_1 - x_1y_3)\mathbf{j} + (x_1y_2 - x_2y_1)\mathbf{k}.$$

Não se trata de um verdadeiro determinante porque na primeira linha temos vectores e não números.

Exemplo 1.18 Consideremos os vectores $\mathbf{x} = (-1, 1, 2)$, $\mathbf{y} = (0, -1, 2)$, $\mathbf{z} = (2, -1, -3) \in \mathbb{R}^3$. Temos

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 1 & 2 \\ 0 & -1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ -1 & 2 \end{vmatrix} \mathbf{i} - \begin{vmatrix} -1 & 2 \\ 0 & 2 \end{vmatrix} \mathbf{j} + \begin{vmatrix} -1 & 1 \\ 0 & -1 \end{vmatrix} \mathbf{k} = 4\mathbf{i} + 2\mathbf{j} + \mathbf{k}.$$

Exercício 1.19 Para os mesmos vectores do exemplo anterior determine $\mathbf{x} \times \mathbf{z}$, $\mathbf{y} \times \mathbf{z}$ e $\mathbf{y} \times \mathbf{x}$.

Produto misto

Definição 1.20 O produto misto de três vectores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$, denotado por $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle$ é o escalar $\langle \mathbf{x} \times \mathbf{y}, \mathbf{z} \rangle$.

Proposição 1.21 M1) Se colocamos os vectores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ numa sucessão da forma

$$\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$$

então o produto misto de quaisquer três elementos consecutivos dá o mesmo valor. Em particular, $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{z}, \mathbf{x} \rangle$. Por outro lado, se trocarmos a ordem de dois vectores quaisquer então o produto misto muda o sinal. Em particular, $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = -\langle \mathbf{x}, \mathbf{z}, \mathbf{y} \rangle$.

M2) As aplicações $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle$, $\mathbf{y} \mapsto \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle$, $\mathbf{z} \mapsto \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle$ são lineares.

M3) $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = 0$ sse $\mathbf{x}, \mathbf{y}, \mathbf{z}$ são linearmente dependentes, isto é pertencem ao mesmo plano. Em particular, $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = 0$ se dois dos vectores $\mathbf{x}, \mathbf{y}, \mathbf{z}$ coincidem.

Demonstração: Exercício.

Proposição 1.22 O produto misto dos vectores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ é dado por

$$\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

Demonstração: Exercício.

Exemplo 1.23 Nas condições do Exemplo 1.18 e utilizando a Proposição 1.21 temos

$$\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = \langle (4, 2, 1), (2, -1, -3) \rangle = 3,$$

o que também pode ser obtido calculando directamente o determinante $\begin{vmatrix} -1 & 1 & 2 \\ 0 & -1 & 2 \\ 2 & -1 & -3 \end{vmatrix}$.

Observação 1.24 Do ponto de vista geométrico, o módulo do produto misto representa o volume do paralelepípedo construído sobre os vectores considerados.

Formas quadráticas

Definição 1.25 Uma aplicação $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ é uma aplicação bilinear quando verifica as condições seguintes, para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}', \mathbf{y}' \in \mathbb{R}^m$ e qualquer $\alpha \in \mathbb{R}$:

- (i) $f(\mathbf{x} + \mathbf{y}, \mathbf{x}') = f(\mathbf{x}, \mathbf{x}') + f(\mathbf{y}, \mathbf{x}')$;
- (ii) $f(\mathbf{x}, \mathbf{x}' + \mathbf{y}') = f(\mathbf{x}, \mathbf{x}') + f(\mathbf{x}, \mathbf{y}')$;
- (iii) $f(\alpha \mathbf{x}, \mathbf{x}') = f(\mathbf{x}, \alpha \mathbf{x}') = \alpha f(\mathbf{x}, \mathbf{x}')$.

Em particular, quando $n = m$ e $p = 1$ a aplicação linear f toma o nome de forma bilinear.

Proposição 1.26 Seja $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ uma forma bilinear e A a matriz $n \times n$ cujos elementos são $a_{ij} = f(\mathbf{e}_i, \mathbf{e}_j)$, com $i, j = 1, \dots, n$. A forma bilinear f fica completamente determinada pela matriz A , uma vez que, para $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ quaisquer, se tem

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^n a_{ij} x_i y_j = \mathbf{x}^T \mathbf{A} \mathbf{y},$$

onde \mathbf{x}^T é a transposta da matriz coluna formada pelas coordenadas do vector \mathbf{x} .

Reciprocamente, dada uma matriz B qualquer, de ordem n , existe uma e uma só forma bilinear $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $g(\mathbf{e}_i, \mathbf{e}_j) = b_{ij}$, para quaisquer $i, j = 1, \dots, n$.

Definição 1.27 Seja $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ uma forma bilinear. A forma quadrática associada a f é a aplicação $Q_f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $Q_f(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Observação 1.28 Dada uma matriz $A = [a_{ij}]$ de ordem n , tendo em conta a Proposição 1.26, podemos dizer que a forma quadrática associada a A é a função $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.

Todas as formas quadráticas podem ser expressas em termos de matrizes simétricas¹.

¹Uma matriz quadrada diz-se simétrica quando é igual á sua transposta.

Proposição 1.29 *Se A é uma matriz $n \times n$, X é uma matriz $n \times 1$ e B é a parte simétrica de A , $B = (A + A^T)/2$, tem-se $X^T A X = X^T B X$.*

Demonstração:

Como $X^T A X$ é escalar, tem-se $X^T A X = (X^T A X)^T = X^T A^T X$. Portanto,

$$X^T A X = \frac{X^T A X + X^T A^T X}{2} = X^T \frac{(A + A^T)}{2} X = X^T B X.$$

■

Definição 1.30 *Diz-se que uma forma quadrática Q , ou uma matriz simétrica que lhe esteja associada, é*

- (i) *definida positiva se $Q(\mathbf{x}) > 0$ para $\mathbf{x} \neq 0$;*
- (ii) *definida negativa se $Q(\mathbf{x}) < 0$ para $\mathbf{x} \neq 0$;*
- (iii) *semidefinida positiva se $Q(\mathbf{x}) \geq 0$ para todo \mathbf{x} ;*
- (iv) *semidefinida negativa se $Q(\mathbf{x}) \leq 0$ para todo \mathbf{x} ;*
- (v) *indefinida se existem pontos onde Q é positiva e pontos onde Q é negativa.*

Definição 1.31 *Os valores próprios de uma matriz A de ordem n são as soluções da equação*

$$\det(A - \lambda I_n) = 0, \text{ sendo } I_n = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix} \text{ a matriz identidade de ordem } n.$$

Proposição 1.32 *Sejam A uma matriz simétrica real $n \times n$ e $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ a forma quadrática associada a A . Tem-se*

- (i) *Q é definida positiva sse todos os valores próprios de A são positivos;*
- (ii) *Q é definida negativa sse todos os valores próprios de A são negativos;*
- (iii) *Q é semidefinida positiva sse todos os valores próprios de A são ≥ 0 ;*
- (iv) *Q é semidefinida negativa sse todos os valores próprios de A são ≤ 0 ;*
- (v) *Q é indefinida sse A tem valores próprios positivos e valores próprios negativos.*

Demonstração: Ver [9, p. 296].

Na prática há uma forma mais simples de classificar uma forma bilinear, de acordo com a proposição seguinte:

Proposição 1.33 (Critério de Silvestre) *Sejam $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ uma forma quadrática associada a uma matriz simétrica A de ordem n . Consideremos os números seguintes (designados por menores principais de A):*

$$\Delta_1 = |a_{11}| = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \dots, \quad \Delta_n = |A|.$$

A forma Q é:

- (i) definida positiva sse $\Delta_k > 0$ ($k = 1, 2, \dots, n$);
- (ii) definida negativa sse $(-1)^k \Delta_k > 0$ ($k = 1, 2, \dots, n$);
- (iii) semidefinida positiva sse $\Delta_k \geq 0$ ($k = 1, 2, \dots, n$);
- (iv) semidefinida negativa sse $(-1)^k \Delta_k \geq 0$ ($k = 1, 2, \dots, n$).

Demonstração: Ver [9, p. 296].

Como é óbvio, em qualquer outro caso a forma quadrática é indefinida.

Exemplos 1.34 1. A forma quadrática $q : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $q(x, y) = x^2 + xy + y^2$ é definida positiva. De facto, está associada à matriz simétrica $A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$ e tem-se $\Delta_1 = |a_{11}| = 1 > 0$ e $\Delta_2 = 1 - \frac{1}{4} = \frac{3}{4} > 0$.

2. A forma quadrática $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = -2x^2 + 2xy - y^2$ é definida negativa. A matriz simétrica que lhe está associada é $\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}$. Tem-se $\Delta_1 = |a_{11}| = -2 < 0$ e $\Delta_2 = 2 - 1 = 1 > 0$.

3. Enquanto que a forma quadrática $q : \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $q(x, y, z) = x^2 + y^2 - z^2 - 3xz + xy + 4yz$ é indefinida. A matriz simétrica associada a q é $\begin{bmatrix} 1 & \frac{1}{2} & -\frac{3}{2} \\ \frac{1}{2} & 1 & 2 \\ -\frac{3}{2} & 2 & -1 \end{bmatrix}$, e tem-se

$$\Delta_1 = 1, \quad \Delta_2 = 1 - \frac{1}{4} = \frac{3}{4}, \quad \Delta_3 = -10.$$

Exercício 1.35 Classifique as formas quadráticas seguintes

- (i) $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $f(x, y, z) = 6x^2 + 8xy + 4y^2 + 3z^2 + 4yz$;
- (ii) $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $g(x, y, z) = 4x^2 + 4xy + 5y^2$;
- (iii) $q : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por $q(x, y, z) = -2x^2 - 2xy + 4xz + 2yz - 2y^2 - 2z^2$.

Recorrendo à noção de norma (tal como no caso de \mathbb{R} recorremos à de módulo) ou, se preferirmos, à de distância, podemos agora introduzir vários conceitos fundamentais para o estudo do cálculo diferencial em \mathbb{R}^n .

1.2 Noções topológicas em \mathbb{R}^n

A teoria da derivação no caso unidimensional trata com funções definidas em intervalos abertos. Para generalizar esta teoria a \mathbb{R}^n , vamos considerar generalizações de intervalos abertos chamados *conjuntos abertos*, para isso precisamos da noção de bola aberta em \mathbb{R}^n .

Definição 1.36 *Sejam $\mathbf{a} \in \mathbb{R}^n$ e $\varepsilon > 0$. O conjunto de todos os pontos $\mathbf{x} \in \mathbb{R}^n$ tais que $\|\mathbf{x} - \mathbf{a}\| < \varepsilon$ chama-se bola aberta de raio ε e centro \mathbf{a} e representa-se por $B_\varepsilon(\mathbf{a})$ ou $B(\mathbf{a}; \varepsilon)$.*

Em \mathbb{R} é, simplesmente, o intervalo aberto centrado em a e de raio ε (também conhecido por vizinhança de a de raio ε): $]a - \varepsilon, a + \varepsilon[$. Em \mathbb{R}^2 é o círculo "aberto" de centro \mathbf{a} e raio ε , isto é, o conjunto de todos os pontos do plano cuja a distância ao ponto \mathbf{a} é menor do que ε .

Na definição de alguns conceitos fundamentais da Análise em \mathbb{R}^n , como por exemplo o de limite, as "bolas" acabadas de definir desempenham naturalmente o papel que coube às "vizinhanças", no caso de \mathbb{R} (ver Secção 1.4). Mas convém deixar bem claro que, em \mathbb{R}^n , o termo *vizinhança* é usado num sentido muito mais geral do que acabámos de referir (ver Definição 1.48).

Vamos agora definir mais algumas noções importantes, chamadas *noções topológicas* por se exprimirem através do conceito de bola aberta.

Definição 1.37 *Sejam $S \subset \mathbb{R}^n$ e $\mathbf{a} \in \mathbb{R}^n$.*

- (i) *Diz-se que \mathbf{a} é ponto interior de S se existe uma bola aberta com centro em \mathbf{a} , cujos pontos pertencem todos a S , i.e., se existe $\varepsilon > 0$ tal que $B_\varepsilon(\mathbf{a}) \subset S$.*

O conjunto de todos os pontos interiores de S diz-se o interior de S e representa-se por $\text{int } S$.

- (ii) *Diz-se que \mathbf{a} é ponto de fronteira a S se para qualquer $\varepsilon > 0$ se tem $B_\varepsilon(\mathbf{a}) \cap S \neq \emptyset$ e $B_\varepsilon(\mathbf{a}) \cap (\mathbb{R}^n \setminus S) \neq \emptyset$.*

O conjunto de todos os pontos fronteiros de S chama-se fronteira de S e representa-se por $\text{fr } S$ ou ∂S .

- (iii) *Diz-se que \mathbf{a} é ponto aderente a S se qualquer bola aberta centrada em \mathbf{a} intersecta S , ou seja, se para qualquer $\varepsilon > 0$ se tem $B_\varepsilon(\mathbf{a}) \cap S \neq \emptyset$.*

O conjunto de todos os pontos aderentes a S chama-se aderência ou fecho de S , e representa-se por \overline{S} .

Observação 1.38 *Chama-se complementar de S ao conjunto $\mathbb{R}^n \setminus S := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \notin S\}$.*

Definição 1.39 *Seja $S \subset \mathbb{R}^n$.*

- (i) *Diz-se que S é aberto se $S = \text{int } S$.*

- (ii) *Diz-se que S é fechado se $S = \overline{S}$.*

Proposição 1.40 A bola aberta $B_\varepsilon(\mathbf{a})$ é um conjunto aberto ($\forall \mathbf{a} \in \mathbb{R}^n, \varepsilon > 0$).

Demonstração: Seja $\mathbf{y} \in B_\varepsilon(\mathbf{a})$, logo $\|\mathbf{y} - \mathbf{a}\| < \varepsilon$, o que implica que existe $\delta > 0$ tal que $\|\mathbf{y} - \mathbf{a}\| < \varepsilon - \delta$. Então $B_\delta(\mathbf{y}) \subset B_\varepsilon(\mathbf{a})$. De facto, se $\mathbf{z} \in B_\delta(\mathbf{y})$ tem-se $\|\mathbf{z} - \mathbf{y}\| < \delta$, logo

$$\|\mathbf{z} - \mathbf{a}\| \leq \|\mathbf{z} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{a}\| < \delta + \varepsilon - \delta = \varepsilon,$$

isto é, $\mathbf{z} \in B_\varepsilon(\mathbf{a})$. Donde $B_\varepsilon(\mathbf{a}) \subset \text{int } B_\varepsilon(\mathbf{a}) \subset B_\varepsilon(\mathbf{a})$. ■

Proposição 1.41 A bola fechada $\overline{B}_\varepsilon(\mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{a}\| \leq \varepsilon\}$ é um conjunto fechado ($\forall \mathbf{a} \in \mathbb{R}^n, \varepsilon > 0$).

Todas as propriedades dos conjuntos abertos, fechados são válidas também no espaço \mathbb{R}^n . Em particular, para quaisquer $A, B \subset \mathbb{R}^n$ tem-se

1. A é aberto sse $\mathbb{R}^n \setminus A$ é fechado;
2. $\overline{A \cup B} = \overline{A} \cup \overline{B}$;
3. $\text{int}(A \cap B) = \text{int } A \cap \text{int } B$.

Exemplos 1.42 (1) Em \mathbb{R} o exemplo mais simples de um conjunto aberto é um intervalo aberto.

Um intervalo fechado $[a, b]$ não é um conjunto aberto porque para nenhum dos pontos extremos do intervalo existe uma bola inteiramente contida no intervalo dado.

(2) A bola $B_1(\mathbf{0})$ em \mathbb{R}^2 centrada na origem e raio 1 é um exemplo de um conjunto aberto de \mathbb{R}^2 . Todo o ponto $\mathbf{a} \in B_1(\mathbf{0})$ é o centro de um círculo contido em $B_1(\mathbf{0})$, embora para os pontos próximos da fronteira o raio desse círculo seja muito pequeno.

(3) Em \mathbb{R}^2 podem-se construir conjuntos abertos através do produto cartesiano de conjuntos abertos de \mathbb{R} , como se pode ver no seguinte resultado.

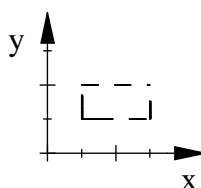
Proposição 1.43 Sejam A_1 e A_2 subconjuntos abertos (resp. fechados) de \mathbb{R} . O seu produto cartesiano:

$$A_1 \times A_2 = \{(a_1, a_2) : a_1 \in A_1 \text{ e } a_2 \in A_2\}$$

é um conjunto aberto (fechado).

Demonstração:

Demonstremos apenas no caso em que A_1 e A_2 são intervalos abertos. O produto cartesiano de dois intervalos abertos é um rectângulo aberto:



Para demonstrar que $A_1 \times A_2$ é um subconjunto aberto de \mathbb{R}^2 , escolhamos $\mathbf{a} = (a_1, a_2) \in A_1 \times A_2$ qualquer e mostremos que \mathbf{a} é ponto interior de $A_1 \times A_2$. Como A_1 e A_2 são abertos em \mathbb{R} existe uma bola aberta $B(a_1; \varepsilon_1)$ contida em A_1 e uma bola aberta $B(a_2; \varepsilon_2)$ contida em A_2 . Seja $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Prova-se que $B(\mathbf{a}; \varepsilon) \subseteq A_1 \times A_2$. De facto, se $\mathbf{x} = (x_1, x_2)$ é um ponto qualquer de $B(\mathbf{a}; \varepsilon)$ então $\|\mathbf{x} - \mathbf{a}\| < \varepsilon$, pelo que $\|x_1 - a_1\| < \varepsilon \leq \varepsilon_1$ e $\|x_2 - a_2\| < \varepsilon \leq \varepsilon_2$. Logo $x_1 \in B(a_1; \varepsilon_1)$ e $x_2 \in B(a_2; \varepsilon_2)$. Deste modo $x_1 \in A_1$ e $x_2 \in A_2$, pelo que $(x_1, x_2) \in A_1 \times A_2$. Isto mostra que cada ponto de $B(\mathbf{a}; \varepsilon)$ está em $A_1 \times A_2$. Por conseguinte todo o ponto de $A_1 \times A_2$ é ponto interior, pelo que $A_1 \times A_2$ é aberto. ■

Observação 1.44 Note que, na demonstração anterior, as bolas $B(a_1; \varepsilon_1)$ e $B(a_2; \varepsilon_2)$ estão em \mathbb{R} e a bola $B(\mathbf{a}; \varepsilon)$ está em \mathbb{R}^2 .

Observação 1.45 Um subconjunto aberto de \mathbb{R} não é um conjunto aberto quando visto como um subconjunto de \mathbb{R}^2 , pois não contém nenhuma bola de \mathbb{R}^2 .

Mas se $A \subset \mathbb{R}$ é fechado então A visto como subconjunto de \mathbb{R}^2 é fechado (em \mathbb{R}^2). Pois o conjunto correspondente a A em \mathbb{R}^2 é $A \times \{0\}$ que, pela Proposição 1.43, é fechado.

Observação 1.46 Em \mathbb{R}^n os conjuntos \emptyset e \mathbb{R}^n são simultaneamente abertos e fechados.

Proposição 1.47 (i) A intersecção de um número finito de conjuntos abertos é um conjunto aberto.

(ii) A reunião de um número finito de conjuntos fechados é um conjunto fechado.

(iii) A reunião de conjuntos abertos é um conjunto aberto.

(iv) A intersecção de conjuntos fechados é um conjunto fechado.

Demonstração: Ver [2].

Definição 1.48 Diz-se que um conjunto $V \subset \mathbb{R}^n$ é uma vizinhança de $\mathbf{a} \in \mathbb{R}^n$ se existe $\varepsilon > 0$ tal que $V \supset B_\varepsilon(\mathbf{a})$ (isto é, V contém \mathbf{a} e uma bola aberta centrada em \mathbf{a}). Se V é um conjunto aberto diz-se que V é uma vizinhança aberta.

Definição 1.49 Seja S um subconjunto de \mathbb{R}^n .

(i) Diz-se que $\mathbf{a} \in \mathbb{R}^n$ é ponto de acumulação de S se qualquer bola aberta centrada em \mathbf{a} tem pelo menos um ponto de S distinto de \mathbf{a} , isto é, para qualquer $\varepsilon > 0$ tem-se $S \cap (B_\varepsilon(\mathbf{a}) \setminus \{\mathbf{a}\}) \neq \emptyset$.

O conjunto de todos os pontos de acumulação de S chama-se derivado de S e representa-se por S' .

(ii) Um ponto $\mathbf{a} \in S$ que não seja ponto de acumulação diz-se ponto isolado de S , isto é, \mathbf{a} é ponto isolado de S se existe $\varepsilon > 0$ tal que $B_\varepsilon(\mathbf{a}) \cap S = \{\mathbf{a}\}$.

Observação 1.50 Da definição anterior resulta que um ponto isolado de S pertence a S e que um ponto de acumulação de S pode não pertencer a S .

Observação 1.51 Para ver a diferença entre S' e \overline{S} consideremos, por exemplo, $S = [0, 1] \cup \{2\}$. Temos $\overline{S} = [0, 1] \cup \{2\}$ e $S' = [0, 1]$ e o ponto 2 é um ponto isolado de S .

Observação 1.52 Seja $S \subset \mathbb{R}^n$. As afirmações seguintes são consequência das definições anteriores:

- (i) $\text{fr } S = \text{fr } (\mathbb{R}^n \setminus S)$;
- (ii) $\text{int } S \subset S$, e $S \subset \overline{S}$;
- (iii) $\overline{S} = \text{int } S \cup \text{fr } S = S \cup \text{fr } S = S \cup S'$.

Exemplo 1.53 Se S é a bola $B_1(\mathbf{0})$ em \mathbb{R}^2 temos $\text{int } S = S$, $\mathbb{R}^n \setminus S = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| > 1\}$, $\text{fr } S = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1\}$, $\overline{S} = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$.

O conjunto $S_\varepsilon(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{a}\| = \varepsilon\}$ é a fronteira das bolas $B_\varepsilon(\mathbf{a})$ e $\overline{B}_\varepsilon(\mathbf{a})$ e chama-se esfera.

Observação 1.54 Note que, em geral, $\text{int } S \neq \text{int } \overline{S}$. Para verificar esta afirmação considere, por exemplo, o conjunto $S = \{(x, y) : x^2 + y^2 \leq 1, x, y \in \mathbb{Q}\}$. De facto, tem-se $\text{int } S = \emptyset$ e $\text{int } \overline{S} = \text{int } \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$

Exercício 1.55 Considere o conjunto $S \subset \mathbb{R}^2$ dado por $S = B_1(\mathbf{0}) \cup \{(2, 2)\}$.

- a) Determine $\text{int } S$, \overline{S} , S' .
- b) Diga, justificando, se S é aberto e/ou fechado.
- c) Verifique se $\overline{\text{int } S} = \overline{S}$.

Definição 1.56 Um conjunto $S \subset \mathbb{R}^n$ diz-se limitado se existe um número real $L > 0$ tal que para qualquer $\mathbf{x} \in S$ se tenha $\|\mathbf{x}\| \leq L$.

Exemplo 1.57 O conjunto $\{(x, y) \in \mathbb{R}^2 : |x| \leq 1, y \geq 0\}$ não é limitado. (Justifique!)

Definição 1.58 Sejam $A, B \subset \mathbb{R}^n$ tais que $A \subset B$. Diz-se que A é denso em B se $B \subset \overline{A}$.

Exemplo 1.59 O conjunto $\mathbb{Q}^n := \{(r_1, \dots, r_n) : r_i \in \mathbb{Q}, i = 1, \dots, n\}$ é denso em \mathbb{R}^n .

Conjuntos conexos. Curvas

Definição 1.60 Um conjunto $X \subset \mathbb{R}^n$ chama-se topologicamente conexo se não existem dois conjuntos abertos (ou fechados) disjuntos A e B (em \mathbb{R}^n) tais que $X \cap A \neq \emptyset$, $X \cap B \neq \emptyset$ e $X \subset A \cup B$.

Na Análise usa-se também outro conceito de conexidade, mas para o apresentar precisamos primeiro de introduzir o conceito de curva.

Definição 1.61 Um caminho em \mathbb{R}^n definido num intervalo $[a, b] \subset \mathbb{R}$ é uma função contínua $r : [a, b] \rightarrow \mathbb{R}^n$.

Observação 1.62 Dados $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, o caminho $r : [0, 1] \rightarrow \mathbb{R}^n$, definido por $r(t) = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$ chama-se o caminho rectilíneo que liga \mathbf{x} a \mathbf{y} .

Exemplo 1.63 O caminho rectilíneo que liga $(1, 2, 3)$ a $(2, 2, 1)$ é $r : [0, 1] \rightarrow \mathbb{R}^3$ dado por

$$r(t) = (1, 2, 3) + t(1, 0, -2) = (1 + t, 2, 3 - 2t).$$

Definição 1.64 Uma curva C é a imagem de um caminho. Isto é, existem um intervalo $[a, b] \subset \mathbb{R}$ e uma função contínua $r : [a, b] \rightarrow \mathbb{R}^n$ tais que $C = r([a, b])$.

Se $r(a) = r(b)$ a curva diz-se fechada.

Observação 1.65 Por vezes confunde-se a curva com o caminho (que é sua parametrização). Mas uma mesma curva pode ser percorrida utilizando diferentes caminhos ou parametrizações, por exemplo

$$\begin{array}{ll} r_1 : [0, 1] \rightarrow \mathbb{R}^2 & e \quad r_2 : [0, 1] \rightarrow \mathbb{R}^2 \\ t \rightarrow (t, t) & t \rightarrow (1 - t, 1 - t) \end{array}$$

definem a mesma curva:

$$r_1([0, 1]) = r_2([0, 1]) = \{(x, x) : x \in [0, 1]\}.$$

Definição 1.66 O caminho $r : [a, b] \rightarrow \mathbb{R}^n$ diz-se:

- (i) regular se for de classe \mathcal{C}^1 e $r'(t) \neq \mathbf{0}$ para qualquer $t \in]a, b[$;
- (ii) seccionalmente regular se existe uma partição finita do intervalo $[a, b]$ tal que r é regular em cada um dos seus subintervalos.

Exemplo 1.67 A função $r : [-1, 1] \rightarrow \mathbb{R}^2$ definida por $r(t) = (t, |t|)$ é contínua em $[-1, 1]$ logo é um caminho em \mathbb{R}^2 . A derivada existe, é contínua e diferente de $\mathbf{0}$ para qualquer $t \in [-1, 1] \setminus \{0\}$, logo trata-se de um caminho regular em $[-1, 0]$ e $[0, 1]$, ou seja, seccionalmente regular em $[-1, 1]$.

Definição 1.68 Dois caminhos $r : [a, b] \rightarrow \mathbb{R}^n$ e $\tilde{r} : [c, d] \rightarrow \mathbb{R}^n$ dizem-se equivalentes se existe uma aplicação bijectiva continuamente diferenciável $\varphi : [c, d] \rightarrow [a, b]$ tal que $\varphi'(t) \neq 0$ para qualquer $t \in]c, d[$ e $\tilde{r}(t) = r \circ \varphi(t)$ para qualquer $t \in [c, d]$.

Se $\varphi'(t) > 0$ diz-se que r e \tilde{r} têm o mesmo sentido; se $\varphi'(t) < 0$ diz-se que r e \tilde{r} têm sentidos opostos; no primeiro caso diz-se que a função φ preserva o sentido, e no segundo caso que inverte o sentido.

Observação 1.69 Note que quaisquer dois caminhos equivalentes definem a mesma curva.

Exemplo 1.70 A elipse $4x^2 + 9y^2 = 36$ percorrida no sentido anti-horário é parametrizada por

$$r(t) = (3 \cos t, 2 \sin t), \quad t \in [0, 2\pi].$$

A parametrização

$$r_1(t) = (3 \cos(2t), 2 \sin(2t)), \quad t \in [0, \pi]$$

também descreve a mesma elipse. De facto, $r_1(t) = r(\varphi(t))$ onde $\varphi : [0, \pi] \rightarrow [0, 2\pi]$ é dada por $\varphi(t) = 2t$.

Já a parametrização

$$r_2(t) = (3 \cos(2t), 2 \sin(2t)), \quad t \in [0, 2\pi]$$

percorre a elipse duas vezes, mas não é equivalente a r nem a r_1 . Neste caso dizemos que r_2 é um caminho com multiplicidade 2.

Definição 1.71 Um conjunto $X \subset \mathbb{R}^n$ diz-se conexo por arcos (ou apenas conexo) se para cada par de pontos $\mathbf{x}, \mathbf{y} \in X$ existe um caminho $r : [\alpha, \beta] \rightarrow \mathbb{R}^n$ tal que $r(t) \in X$ para qualquer $t \in [\alpha, \beta]$, $r(\alpha) = \mathbf{x}$ e $r(\beta) = \mathbf{y}$.

Proposição 1.72 Se $X \subset \mathbb{R}^n$ é conexo por arcos então X é topologicamente conexo.

O exemplo seguinte mostra que o contrário não é verdade

Exemplo 1.73 Consideremos o conjunto

$$X = \left\{ (x, y) \in \mathbb{R}^2 : y = \sin \frac{1}{x}, x > 0 \right\} \cup \{(0, 0)\}.$$

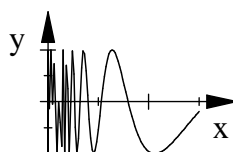


gráfico da função $y = \sin \frac{1}{x}$, $x > 0$

Claro que X é topologicamente conexo porque $(0, 0)$ é ponto de acumulação de X , mas não é simplesmente conexo porque não se pode ligar nenhum ponto da forma $(x, \sin \frac{1}{x})$ à origem por uma curva verificando as condições da Definição 1.71.

Definição 1.74 Um subconjunto S de \mathbb{R}^n diz-se convexo se dados dois pontos em S o segmento que os une também está em S .

Nota 1.75 Todo o conjunto convexo é conexo. Mas nem todo o conexo é convexo.

Todos estes conceitos topológicos podem ser caracterizados em termos de sucessões, tal como em \mathbb{R} .

1.3 Sucessões em \mathbb{R}^n

Definição 1.76 Uma sucessão em \mathbb{R}^n é uma aplicação $f : \mathbb{N} \rightarrow \mathbb{R}^n$ que a cada $m \in \mathbb{N}$ associa o ponto $f(m)$ denotado por $\mathbf{x}^m = (x_1^m, \dots, x_n^m) \in \mathbb{R}^n$, e chamado m -ésimo termo da sucessão. Em vez de f escrevemos $\{\mathbf{x}^m\}$ para denotar a sucessão. Se $\{k_m\}$ é uma sucessão de números naturais tal $m_1 > m_2$ implica $k_{m_1} > k_{m_2}$ (i.e., a ordem de k coincide com a ordem dos índices m) então a sucessão $\{\mathbf{x}^{k_m}\}$ chama-se subsucessão de $\{\mathbf{x}^m\}$.

Definição 1.77 Diz-se que a sucessão $\{\mathbf{x}^m\} \subset \mathbb{R}^n$ converge para $\mathbf{x} \in \mathbb{R}^n$, e escreve-se $\mathbf{x}^m \rightarrow \mathbf{x}$ quando $m \rightarrow \infty$, se para cada $\varepsilon > 0$ existe $m^* \in \mathbb{N}$ tal que para qualquer $m \geq m^*$ se tem $\|\mathbf{x}^m - \mathbf{x}\| \leq \varepsilon$.

Observação 1.78 Da definição anterior é imediato que $\mathbf{x}^m \rightarrow \mathbf{x}$ quando $m \rightarrow \infty$ é equivalente a $\|\mathbf{x}^m - \mathbf{x}\| \rightarrow 0$, quando $m \rightarrow \infty$. Isto reduz a convergência em \mathbb{R}^n à convergência de números reais ≥ 0 .

Exemplo 1.79 A sucessão em \mathbb{R}^3 com termo geral

$$\mathbf{x}^m = \left(\frac{m^2 - 1}{m^2}, 0, \frac{1}{3^m} \right)$$

converge para o vector $\mathbf{e}_1 = (1, 0, 0)$. Para mostrar basta ter em conta que

$$\|\mathbf{x}^m - \mathbf{e}_1\| = \sqrt{\frac{1}{m^2} + \frac{1}{3^{2m}}} \rightarrow 0 \quad \text{quando } m \rightarrow \infty.$$

Observação 1.80 Naturalmente, quando a sucessão $\{\mathbf{x}^m\}$ é convergente, chamamos limite de $\{\mathbf{x}^m\}$ ao único (Provar!) vector \mathbf{x} tal que $\mathbf{x}^m \rightarrow \mathbf{x}$, e escreveremos $\lim_{m \rightarrow \infty} \mathbf{x}^m = \mathbf{x}$.

O limite de qualquer subsucessão $\{\mathbf{x}^{k_m}\}$ da sucessão $\{\mathbf{x}^m\}$ chama-se sublimite da sucessão $\{\mathbf{x}^m\}$. Ao contrário do limite de uma sucessão, o sublimite não é necessariamente único.

A convergência de uma sucessão é equivalente à convergência da sucessão das suas coordenadas:

Proposição 1.81 A sucessão $\{\mathbf{x}^m\} \subset \mathbb{R}^n$, $\mathbf{x}^m = (x_1^m, \dots, x_n^m)$, $m = 1, 2, \dots$, converge para um ponto $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ sse para cada $i = 1, 2, \dots, n$ tem-se $\lim_{m \rightarrow \infty} x_i^m = x_i$.

Demonstração: Exercício.

Teorema 1.82 (Bolzano-Weierstrass) Toda a sucessão limitada em \mathbb{R}^n possui uma sub-sucessão convergente (ou, por outras palavras, possui pelo menos um sublimite).

Demonstração: Ver, por exemplo, [7, Teorema 5, p.16].

Estendemos agora a \mathbb{R}^n o critério de Cauchy para a convergência de sucessões de números reais.

Definição 1.83 Uma sucessão $\{\mathbf{x}^m\} \subset \mathbb{R}^n$ diz-se uma sucessão de Cauchy quando para cada $\varepsilon > 0$ existe um número $m^* \in \mathbb{N}$ tal que

$$m, p \geq m^* \Rightarrow \|\mathbf{x}^m - \mathbf{x}^p\| \leq \varepsilon.$$

Usando a Proposição 1.81 e o critério de Cauchy para sucessões numéricas podemos demonstrar a critério de Cauchy para sucessões em \mathbb{R}^n .

Teorema 1.84 Uma sucessão $\{\mathbf{x}^m\} \subset \mathbb{R}^n$ é convergente sse é de Cauchy.

Demonstração: Ver, por exemplo, [7, Teorema 7, p.17].

Definição 1.85 Diz-se que uma sucessão $\{\mathbf{x}^m\} \subset \mathbb{R}^n$ é limitada se existe $M > 0$ tal que $\|\mathbf{x}^m\| \leq M$, $\forall m = 1, 2, \dots$ (i.e., o conjunto dos termos da sucessão é limitado).

Proposição 1.86 Um conjunto $S \subset \mathbb{R}^n$ é fechado sse o limite de qualquer sucessão convergente de elementos de S pertence a S .

Definição 1.87 Um conjunto $S \subset \mathbb{R}^n$ diz-se compacto se cada sucessão $\{\mathbf{x}^m\} \subset S$ tem uma subsucessão $\{\mathbf{x}^{k_m}\}$ que converge para um ponto $\mathbf{x} \in S$.

Usando esta propriedade e a Proposição 1.86 podemos formular o Teorema de Bolzano-Weierstrass da seguinte forma:

Teorema 1.88 (Bolzano-Weierstrass) Um conjunto $S \subset \mathbb{R}^n$ é compacto sse S é limitado e fechado.

Exemplo 1.89 A esfera $\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 3\}$ é um conjunto compacto. (Justifique!)

1.4 Funções. Limite e continuidade

Funções de várias variáveis

Até agora estudaram-se apenas funções reais de variável real. No entanto, existem situações em diferentes áreas da ciência em que são necessárias funções de várias variáveis, umas com valores reais (também chamadas *funções escalares*) como, por exemplo, no cálculo de áreas e volumes, outras com valores vectoriais (chamadas *funções vectoriais*) como, por exemplo, no cálculo da velocidade de um ponto num fluido em movimento. São estas funções, escalares e vectoriais, que vamos estudar a seguir.

Iniciamos o estudo das funções de \mathbb{R}^n em \mathbb{R}^m com a introdução de alguns conceitos já conhecidos no caso das funções reais de variável real, nomeadamente, os conceitos de função, domínio, contradomínio entre outros.

Definição 1.90 *Sejam A e B conjuntos. Chama-se função definida em A com valores em B , a toda a correspondência entre A e B que a cada elemento $x \in A$ faz corresponder um único elemento $y \in B$ e representa-se por $y = f(x)$, onde x é a variável independente e y é a variável dependente.*

O conjunto A é o conjunto dos pontos onde a função está definida, chama-se domínio de f e representa-se por D_f ou $D(f)$.

Chama-se contradomínio de f e representa-se por CD_f ou $CD(f)$ ao subconjunto de B que contém todos os valores que a função toma.

O gráfico de uma função $f : A \rightarrow B$ é o conjunto

$$Gr(f) = \{(x, y) \in A \times B : x \in A \text{ e } y = f(x)\}.$$

Diz-se que $f : A \rightarrow B$ é uma função real de variável real quando $A \subset \mathbb{R}$ e $B \subset \mathbb{R}$.

Diz-se que $f : A \rightarrow B$ é uma função real de n variáveis reais ou uma função escalar quando $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}$, com $n > 1$.

Diz-se que $f : A \rightarrow B$ é uma função vectorial (ou aplicação) quando $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}^m$, com $n, m > 1$.

Observação 1.91 *No último caso da definição anterior temos uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, com $n, m > 1$, definida por m funções reais de n variáveis reais $f_i : D_i \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, \dots, m$, onde D_i é o domínio da função f_i . Ou seja,*

$$f(x_1, x_2, \dots, x_n) = (y_1, y_2, \dots, y_m),$$

onde

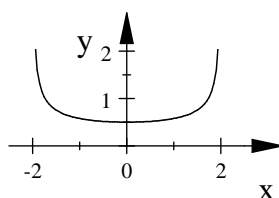
$$\begin{aligned} y_1 &= f_1(x_1, x_2, \dots, x_n) \\ y_2 &= f_2(x_1, x_2, \dots, x_n) \\ &\vdots \\ y_m &= f_m(x_1, x_2, \dots, x_n), \end{aligned}$$

Cada função f_i diz-se uma coordenada de f . Tem-se $D_f = D_1 \cap D_2 \cap \dots \cap D_m$, ou seja, o domínio de f é dado pela intersecção dos domínios das funções coordenadas.

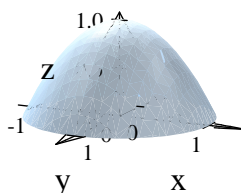
Exemplos 1.92 (1) *Consideremos a função real de variável real $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = \frac{1}{\sqrt{4-x^2}}$. Esta função tem por domínio o conjunto*

$$D_f = \{x \in \mathbb{R} : 4 - x^2 > 0\} =]-2, 2[$$

e contradomínio o conjunto $[\frac{1}{2}, +\infty)$, sendo o gráfico



- (2) Consideremos a função escalar $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = \sqrt{1 - x^2 - y^2}$, temos $D_f = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$, $CD_f = [0, 1]$ e



- (3) Consideremos a função vectorial $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por $f(x, y) = \left(\ln x, \sqrt{y}, \frac{x^2}{x^2 + y^2} \right)$. Esta função é constituída por 3 funções coordenadas $f_i : D_i \subset \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, 2, 3$, tais que

$$\begin{aligned} f_1(x, y) &= \ln x, & D_1 &= \{(x, y) \in \mathbb{R}^2 : x > 0\} \\ f_2(x, y) &= \sqrt{y}, & D_2 &= \{(x, y) \in \mathbb{R}^2 : y \geq 0\} \\ f_3(x, y) &= \frac{x^2}{x^2 + y^2}, & D_3 &= \{(x, y) \in \mathbb{R}^2 : x \neq 0 \text{ ou } y \neq 0\}. \end{aligned}$$

O domínio da função f é igual à intersecção dos domínios das suas funções coordenadas:

$$D_f = D_1 \cap D_2 \cap D_3 = \{(x, y) \in \mathbb{R}^2 : x > 0, y \geq 0\}.$$

A representação gráfica de uma função real de variável real faz-se num espaço de dimensão 2, e a de uma função escalar definida em $D \subset \mathbb{R}^2$ faz-se num espaço de dimensão 3. Em casos como o do Exemplo 1.92 costuma-se recorrer à representação de *curvas de nível*.

Definição 1.93 Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Chama-se conjunto de nível c ao conjunto

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

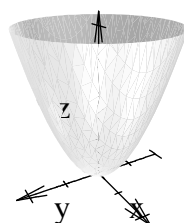
onde c é um dos valores que f toma. Quando $n = 2$ o conjunto $L(c)$ diz-se uma curva de nível e quando $n = 3$ diz-se uma superfície de nível.

Exemplo 1.94 Consideremos a função escalar $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = \sqrt{1 - x^2 - y^2}$. As curvas nível são circunferências centradas em $(0, 0)$ e raio $\sqrt{1 - c^2}$. Por exemplo, a curva de nível $L(0)$ é a circunferência de centro $(0, 0)$ e raio 1 e é dada por

$$L(0) = \{(x, y) \in D : \sqrt{1 - x^2 - y^2} = 0\} = \{(x, y) \in D : x^2 + y^2 = 1\}.$$

Tal como acontece com as funções reais de variável real, também as funções de várias variáveis podem ser definidas de diferentes maneiras. No entanto nem sempre é possível passar de umas para outras, por exemplo, em geral, não é possível escrever de forma explícita uma função que é dada implicitamente, como veremos mais à frente. Temos então as seguintes formas de definir uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$,

- (i) *explícita* quando é dada a expressão analítica da função f , isto é, $z = f(x_1, \dots, x_n)$. Exemplo (em \mathbb{R}^2) $z = f(x, y) = x^2 \sin(2y)$;
- (ii) *implícita* quando não há uma expressão analítica f mas é dada uma função $F : D \times \mathbb{R} \rightarrow \mathbb{R}$ tal que para qualquer $\mathbf{x} \in D$ existe um único $z \in \mathbb{R}$ que satisfaz a equação $F(\mathbf{x}, z) = 0$. Neste caso diz-se que z é uma função de \mathbf{x} definida implicitamente. Por exemplo z como função de x e y (variáveis independentes em \mathbb{R}^2) pode ser dada através da equação $F(x, y, z) = xe^z \cos(xy) + \ln(y^2 + 1) = 0$. Nalguns casos, como no exemplo anterior, a função z pode ser escrita explicitamente em função de x e y . Mas isto nem sempre é possível, como acontece, por exemplo, com a equação $xz = \sin(y + z)$.
- (iii) *gráfica* quando é dado o gráfico da função, isto é, o conjunto $\{(\mathbf{x}, z) \in D \times \mathbb{R} : z = f(\mathbf{x})\}$. Esta representação está apenas definida para $n = 1$ ou $n = 2$. Por exemplo



representa o conjunto $\{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$ que é o gráfico da função $f(x, y) = x^2 + y^2$, esta superfície chama-se parabolóide de revolução;

- (iv) *paramétrica* quando o gráfico da função f se representa como o conjunto

$$\{(\mathbf{x}, z) : \mathbf{x} = X(\mathbf{u}), z = Z(\mathbf{u}), \mathbf{u} \in G\},$$

onde $G \subset \mathbb{R}^n$ é um conjunto de parâmetros e $X : G \rightarrow D$, $Z : G \rightarrow \mathbb{R}$ são funções. Por exemplo,

- (a) para cada $r > 0$ fixado, as equações

$$x = r \cos \theta \sin \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \varphi,$$

para $(\theta, \varphi) \in [0, 2\pi] \times [0, \pi]$, representam uma esfera de raio r e centro na origem;

(b) para cada $r > 0$, as equações

$$x = r \cos \theta, \quad y = r \sin \theta \quad z = \pm \sqrt{1 + r^2},$$

para $\theta \in [0, 2\pi]$ representam um hiperbolóide de revolução de duas folhas,

(c) para cada $r > 1$, as equações

$$x = r \cos \theta, \quad y = r \sin \theta \quad z = \pm \sqrt{r^2 - 1},$$

para $\theta \in [0, 2\pi]$ representam um hiperbolóide de revolução de uma folhas.

Limite

Os conceitos de limite e de continuidade, assim como as suas propriedades, introduzidos na Análise Matemática I, são facilmente generalizáveis a funções de várias variáveis.

Tal como para funções reais de variável real, o conceito de limite num ponto $\mathbf{a} \in \mathbb{R}^n$, para uma função definida numa parte D de \mathbb{R}^n , exige que a variável independente $\mathbf{x} \in D$ se possa aproximar arbitrariamente do ponto \mathbf{a} . Assim, sendo irrelevante que \mathbf{a} pertença ou não a D , é porém necessário que \mathbf{a} seja ponto de acumulação de D , para que nos possamos aproximar de \mathbf{a} através de pontos em $D \setminus \{\mathbf{a}\}$.

Definição 1.95 (Heine) *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 1$, uma aplicação e \mathbf{a} um ponto de acumulação de D . Diz-se que o ponto $\mathbf{b} \in \mathbb{R}^m$ é o limite de f quando \mathbf{x} tende para \mathbf{a} , com $\mathbf{x} \in D$, se para cada sucessão $\{\mathbf{x}^k\} \subset D$, $\{\mathbf{x}^k\} \rightarrow \mathbf{a}$, $k \rightarrow \infty$, tem-se*

$$\lim_{k \rightarrow \infty} f(\mathbf{x}^k) = \mathbf{b}.$$

Neste caso escreve-se $\lim_{\mathbf{x} \rightarrow \mathbf{a}, \mathbf{x} \in D} f(\mathbf{x}) = \mathbf{b}$ ou, simplesmente, $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \mathbf{b}$ (se D está fixado).

Definição 1.96 (Cauchy) *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 1$, e $\mathbf{a} \in D'$. Diz-se que $\mathbf{b} \in \mathbb{R}^m$ é o limite de f quando \mathbf{x} tende para \mathbf{a} , com $\mathbf{x} \in D$, se para cada $\varepsilon > 0$ existe $\delta > 0$ tal que*

$$\mathbf{x} \in D \quad e \quad \|\mathbf{x} - \mathbf{a}\|_n \leq \delta \quad \text{implica} \quad \|f(\mathbf{x}) - \mathbf{b}\|_m \leq \varepsilon. \quad (1.2)$$

Recorde que $\|\cdot\|_n$ e $\|\cdot\|_m$ denotam a norma euclideana em \mathbb{R}^n e \mathbb{R}^m , respectivamente. Quando não houver perigo de confusão escrevemos simplesmente $\|\cdot\|$ para denotar qualquer uma delas.

Demonstra-se, analogamente como em \mathbb{R} , que as definições de Heine e de Cauchy são equivalentes, mas não vamos fazer isso aqui.

Observação 1.97 *O limite em (1.2) é equivalente ao limite nulo da função escalar $\|f(\mathbf{x}) - \mathbf{b}\|$:*

$$\lim_{\|\mathbf{x} - \mathbf{a}\| \rightarrow 0} \|f(\mathbf{x}) - \mathbf{b}\| = 0.$$

Denotando por f_j e b_j ($j = 1, 2, \dots, m$) a coordenada j de f e de \mathbf{b} , respectivamente, temos, por (1.1),

$$\max_{j=1,2,\dots,m} |f_j(\mathbf{x}) - b_j| \leq \|f(\mathbf{x}) - \mathbf{b}\| \leq \sum_{j=1}^m |f_j(\mathbf{x}) - b_j|. \quad (1.3)$$

A ultima desigualdade permite-nos estabelecer o seguinte resultado

Proposição 1.98 *Sejam $D \subset \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}^m$, $\mathbf{a} \in D'$ e $\mathbf{b} \in \mathbb{R}^m$. Temos*

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \mathbf{b} \Leftrightarrow \lim_{\mathbf{x} \rightarrow \mathbf{a}} f_j(\mathbf{x}) = b_j, \quad \forall j = 1, 2, \dots, m.$$

Observação 1.99 *Para perceber melhor a Definição 1.96 suponhamos que $m = 1$ e sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{a} = (a_1, \dots, a_n) \in D'$ e $b \in \mathbb{R}$. Pela definição de norma euclideana dizer que $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = b$ significa que o valor da função f pode ser arbitrariamente próximo de b se a distância de \mathbf{x} ao ponto \mathbf{a} for suficientemente pequena. Observemos que a distância de \mathbf{x} a \mathbf{a} pode ser medida no sentido da norma euclideana ou através das coordenadas (isto é, calculando a distância de cada coordenada de \mathbf{x} à respectiva coordenada de \mathbf{a} - ver (1.3)).*

Exemplo 1.100 *Seja $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ a função definida por $f(x, y) = \frac{xy^2}{x^2 + y^2}$.*

Provemos que $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$. Temos então que mostrar que

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tal que se } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\} \text{ e } \sqrt{x^2 + y^2} < \delta \text{ então } \left| \frac{xy^2}{x^2 + y^2} - 0 \right| < \varepsilon.$$

Seja $\varepsilon > 0$ arbitrário. Queremos encontrar $\delta > 0$ tal que para $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ se tem

$$\sqrt{x^2 + y^2} < \delta \Rightarrow \left| \frac{xy^2}{x^2 + y^2} - 0 \right| < \varepsilon. \quad (1.4)$$

Como $|x| \leq \sqrt{x^2 + y^2}$ então

$$\left| \frac{xy^2}{x^2 + y^2} - 0 \right| = \frac{|x| y^2}{x^2 + y^2} \leq \frac{\sqrt{x^2 + y^2} (x^2 + y^2)}{x^2 + y^2} = \sqrt{x^2 + y^2}.$$

Assim, escolhendo $0 < \delta \leq \varepsilon$ tem-se (1.4).

Proposição 1.101 (Unicidade do limite) *O limite de uma função num ponto, quando existe, é único.*

Demonstração:

Sai directamente da definição de Heine e da unicidade do limite de uma sucessão. ■

Proposição 1.102 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 1$, $\mathbf{a} \in D'$. Se $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \mathbf{b} \in \mathbb{R}^m$ então a função f é limitada em torno do ponto \mathbf{a} , isto é, existem $\delta > 0$ e $M > 0$ tais que $\|f(\mathbf{x})\| \leq M$ para todo o $\mathbf{x} \in D \cap B_\delta(\mathbf{a})$.*

Demonstração:

Por definição de limite para qualquer $\varepsilon > 0$ existe $\delta > 0$ tal que, para todo $\mathbf{x} \in D$

$$\|\mathbf{x} - \mathbf{a}\| \leq \delta \Rightarrow \|f(\mathbf{x}) - \mathbf{b}\| \leq \varepsilon.$$

Tomando $\varepsilon = 1$ na definição de limite, obtemos $\delta > 0$ tal que

$$(\mathbf{x} \in D, \|\mathbf{x} - \mathbf{a}\| \leq \delta) \Rightarrow \|f(\mathbf{x}) - \mathbf{b}\| \leq 1.$$

Por outro lado, por uma propriedade das normas (ver Proposição 1.7 (ii)) sabemos que

$$\|f(\mathbf{x})\| - \|\mathbf{b}\| \leq \|f(\mathbf{x}) - \mathbf{b}\| \leq 1.$$

Portanto, se $\mathbf{x} \in D \cap B_\delta(\mathbf{a})$ tem-se que

$$\|f(\mathbf{x})\| \leq \|\mathbf{b}\| + 1.$$

Tomando $M = \|\mathbf{b}\| + 1$ concluímos que $\|f(\mathbf{x})\| \leq M$, para todo $\mathbf{x} \in D \cap B_\delta(\mathbf{a})$, para algum $\delta > 0$. ■

Proposição 1.103 *Sejam $f, g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 1$, $\mathbf{a} \in D'$, $\mathbf{b}, \mathbf{c} \in \mathbb{R}^m$ e $\lambda \in \mathbb{R}$. Se $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \mathbf{b}$ e $\lim_{\mathbf{x} \rightarrow \mathbf{a}} g(\mathbf{x}) = \mathbf{c}$, então tem-se*

- (a) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} (f(\mathbf{x}) + g(\mathbf{x})) = \mathbf{b} + \mathbf{c}$;
- (b) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \lambda f(\mathbf{x}) = \lambda \mathbf{b}$;
- (c) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \langle f(\mathbf{x}), g(\mathbf{x}) \rangle = \langle \mathbf{b}, \mathbf{c} \rangle$;
- (d) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) \times g(\mathbf{x}) = \mathbf{b} \times \mathbf{c}$, se $m = 3$;
- (e) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \|f(\mathbf{x})\| = \|\mathbf{b}\|$;
- (f) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{\mathbf{b}}{\mathbf{c}}$, se $m = 1$ e $\mathbf{c} \neq 0$.

Demonstração:

Provaremos apenas a propriedade (c). A (d) demonstra-se de forma análoga, e as outras provam-se da mesma maneira que as propriedades correspondentes para as funções reais de variável real.

Suponhamos que $\mathbf{b} \neq \mathbf{0}$. Aplicando a desigualdade triangular do módulo e a desigualdade de Cauchy-Schwarz (ver Proposição 1.7 (i)) obtemos, respectivamente,

$$\begin{aligned} |\langle f(\mathbf{x}), g(\mathbf{x}) \rangle - \langle \mathbf{b}, \mathbf{c} \rangle| &= |\langle f(\mathbf{x}), g(\mathbf{x}) \rangle - \langle \mathbf{b}, g(\mathbf{x}) \rangle| + |\langle \mathbf{b}, g(\mathbf{x}) \rangle - \langle \mathbf{b}, \mathbf{c} \rangle| \\ &\leq |\langle f(\mathbf{x}) - \mathbf{b}, g(\mathbf{x}) \rangle| + |\langle \mathbf{b}, g(\mathbf{x}) - \mathbf{c} \rangle| \\ &\leq \|f(\mathbf{x}) - \mathbf{b}\| \|g(\mathbf{x})\| + \|\mathbf{b}\| \|g(\mathbf{x}) - \mathbf{c}\|. \end{aligned}$$

Fixemos $\varepsilon > 0$ qualquer. Usando a definição de \mathbf{b} e de \mathbf{c} e a proposição anterior encontramos $\delta_1 > 0$ e $M > 0$ tais que

$$\|g(\mathbf{x})\| \leq M \quad \forall \mathbf{x} \in D, \|\mathbf{x} - \mathbf{a}\| \leq \delta_1;$$

$\delta_2 > 0$ tal que

$$\|f(\mathbf{x}) - \mathbf{b}\| \leq \frac{\varepsilon}{2M} \quad \forall \mathbf{x} \in D, \|\mathbf{x} - \mathbf{a}\| \leq \delta_2$$

e $\delta_3 > 0$ tal que

$$\|g(\mathbf{x}) - \mathbf{c}\| \leq \frac{\varepsilon}{2\|\mathbf{b}\|} \quad \forall \mathbf{x} \in D, \|\mathbf{x} - \mathbf{a}\| \leq \delta_3.$$

Pondo $\delta = \min\{\delta_1, \delta_2, \delta_3\}$ vemos que para todo $\mathbf{x} \in D$ com $\|\mathbf{x} - \mathbf{a}\| \leq \delta$ tem-se

$$|\langle f(\mathbf{x}), g(\mathbf{x}) \rangle - \langle \mathbf{b}, \mathbf{c} \rangle| \leq \frac{\varepsilon}{2M} M + \|\mathbf{b}\| \frac{\varepsilon}{2\|\mathbf{b}\|} = \varepsilon.$$

O caso $\mathbf{b} = \mathbf{0}$ é óbvio e é deixado como exercício. ■

Quando estudámos limites de funções reais de variável real vimos que existem dois limites particularmente importantes: o *limite lateral direito* de f em a e o *limite lateral esquerdo* de f em a . Vimos também que uma função real de variável real tem limite num ponto a sse existem, forem finitos e iguais os dois limites laterais.

No caso de funções de várias variáveis a situação é mais complexa. Não existem apenas dois caminhos possíveis para nos aproximarmos de \mathbf{a} , mas sim uma infinidade deles, e não é possível calcular o limite ao longo de todos eles. No entanto, tal como anteriormente, se existirem dois caminhos diferentes ao longo dos quais o limite é diferente, podemos concluir que não há limite da função no ponto \mathbf{a} .

Para explicar melhor o que acabámos de afirmar precisamos do seguinte

Definição 1.104 Se $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma função e A é um subconjunto de D , chama-se restrição de f ao conjunto A e representa-se por $f|_A$ à função que tem por domínio A e verifica a condição $f|_A(\mathbf{x}) = f(\mathbf{x})$ para todo $\mathbf{x} \in A$.

Definição 1.105 Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função e \mathbf{a} um ponto de acumulação de D . Se A é um subconjunto de D ao qual \mathbf{a} ainda é ponto de acumulação chama-se limite de f no ponto \mathbf{a} relativo ao conjunto A ou restrito ao conjunto A ao limite em \mathbf{a} da restrição de $f|_A$, caso ele exista, e escreve-se

$$\lim_{\substack{\mathbf{x} \rightarrow \mathbf{a} \\ \mathbf{x} \in A}} f(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} f|_A(\mathbf{x}).$$

Observação 1.106 Das definições de limite e de restrição de uma função resulta que, caso exista limite de f em \mathbf{a} , todos os limites restritos de f em \mathbf{a} existem e são iguais ao limite de f em \mathbf{a} . No entanto, pode haver limite restrito a certos conjuntos sem que haja o limite em \mathbf{a} . De facto, se for possível determinar dois conjuntos $A, B \subset D$ com $\mathbf{a} \in A' \cap B'$ para os quais se tenha

$$\lim_{\substack{\mathbf{x} \rightarrow \mathbf{a} \\ \mathbf{x} \in A}} f(\mathbf{x}) \neq \lim_{\substack{\mathbf{x} \rightarrow \mathbf{a} \\ \mathbf{x} \in B}} f(\mathbf{x}),$$

(ou um só conjunto A , nas mesmas condições, tal que não exista o limite de f no ponto \mathbf{a} relativo a A), então não existe $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x})$.

Exemplo 1.107 Seja $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}.$$

Mostremos que não existe o limite $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$.

Pondo

$$A = \{(x, 0) : x \in \mathbb{R} \setminus \{0\}\} \quad e \quad B = \{(0, y) : y \in \mathbb{R} \setminus \{0\}\}$$

temos

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \in A}} f(x, y) = \lim_{\substack{(x,y) \rightarrow (0,0) \\ y=0}} \frac{x^2 - y^2}{x^2 + y^2} = \lim_{x \rightarrow 0} \frac{x^2 - 0}{x^2 + 0} = 1$$

e

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \in B}} f(x, y) = \lim_{\substack{(x,y) \rightarrow (0,0) \\ x=0}} \frac{x^2 - y^2}{x^2 + y^2} = \lim_{y \rightarrow 0} \frac{0 - y^2}{0 + y^2} = -1.$$

Como existem dois limites relativos de f no ponto $(0, 0)$ distintos, podemos concluir que não existe $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$.

Observação 1.108 Em \mathbb{R}^2 , os limites relativos a rectas que passam pelo ponto (a, b) , i.e., os limites relativos a rectas da forma

$$x = a \quad e \quad y = b + \alpha(x - a),$$

chamam-se limites direccionais. Se o limite depender do declive da recta, α , então não existe limite. Por outro lado, se ao longo das rectas obtivermos sempre o mesmo valor para o limite, não significa que exista limite, pois poderá existir um outro caminho que passe em (a, b) ao longo do qual o limite seja diferente ou não exista.

Portanto, se os limites direccionais forem diferentes podemos concluir que não existe limite da função no ponto, mas se os limites direccionais forem todos iguais nada podemos concluir.

Exemplos 1.109 (1) Mostremos que a função $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

não tem limite no ponto $(0, 0)$.

Determinemos os limites direccionais ao longo da recta $y = mx$

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ y=mx}} f(x, y) = \lim_{\substack{(x,y) \rightarrow (0,0) \\ y=mx}} \frac{xy}{x^2 + y^2} = \lim_{x \rightarrow 0} \frac{mx^2}{x^2 + (mx)^2} = \lim_{x \rightarrow 0} \frac{mx^2}{x^2(1 + m^2)} = \frac{m}{1 + m^2}.$$

Como o limite direccional depende de m , podemos concluir que não existe o limite

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y).$$

(2) Seja $f : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$ a função definida por

$$f(x, y) = \frac{x^2 y}{x^4 + y^2}.$$

Verifiquemos se existe o limite $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$.

Determinemos os limites direccionais ao longo da recta $y = mx$

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ y=mx}} f(x, y) = \lim_{\substack{(x,y) \rightarrow (0,0) \\ y=mx}} \frac{x^2 y}{x^4 + y^2} = \lim_{x \rightarrow 0} \frac{mx^3}{x^4 + (mx)^2} = \lim_{x \rightarrow 0} \frac{mx^3}{x^4 + (mx)^2} = 0$$

(justifique!), e ao longo da recta $x = 0$

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ x=0}} f(x, y) = \lim_{\substack{(x,y) \rightarrow (0,0) \\ x=0}} \frac{x^2 y}{x^4 + y^2} = \lim_{y \rightarrow 0} \frac{0}{0 + y^2} = 0.$$

Como todos os limites direccionais são iguais a zero nada podemos concluir. Determinemos agora o limite ao longo da parábola $y = x^2$. Temos

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

como este limite é diferente dos limites direccionais, podemos concluir que não existe limite da função f no ponto $(0,0)$.

Exercício 1.110 Seja $f : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$ a função definida por

$$f(x, y) = \frac{x^2 y}{x^2 + y^2}.$$

Mostre que $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$.

Sugestão: basta ter em conta que $x^2 \leq x^2 + y^2$ e que $y \leq \sqrt{x^2 + y^2}$.

Observações 1.111 Em \mathbb{R}^2 , os limites

$$\lim_{x \rightarrow a} \left(\lim_{y \rightarrow b} f(x, y) \right) \quad \text{e} \quad \lim_{y \rightarrow b} \left(\lim_{x \rightarrow a} f(x, y) \right)$$

chamam-se limites iterados (ou limites repetidos) de f em (a, b) . Se existirem, em \mathbb{R} , e forem diferentes, podemos concluir que não existe o limite $\lim_{(x,y) \rightarrow (a,b)} f(x, y)$. No entanto, podem existir e serem iguais sem que exista $\lim_{(x,y) \rightarrow (a,b)} f(x, y)$, como se pode ver no Exemplo 1.112.1 abaixo.

Exemplos 1.112 (1) Seja $f : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$ uma função definida por

$$f(x, y) = \frac{xy}{x^2 + y^2}.$$

No Exemplo 1.109.1 mostrámos que f não tem limite no ponto $(0,0)$ mas, no entanto, os limites iterados existem e são iguais a zero (justifique).

(2) Mostremos, utilizando limites iterados, que a função $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$$

não tem limite no ponto $(0, 0)$.

Temos

$$\lim_{x \rightarrow 0} \left(\lim_{y \rightarrow 0} f(x, y) \right) = \lim_{x \rightarrow 0} \left(\lim_{y \rightarrow 0} \frac{x^2 - y^2}{x^2 + y^2} \right) = \lim_{x \rightarrow 0} \left(\frac{x^2}{x^2} \right) = 1$$

e

$$\lim_{y \rightarrow 0} \left(\lim_{x \rightarrow 0} f(x, y) \right) = \lim_{y \rightarrow 0} \left(\lim_{x \rightarrow 0} \frac{x^2 - y^2}{x^2 + y^2} \right) = \lim_{y \rightarrow 0} \left(\frac{-y^2}{y^2} \right) = -1.$$

Como os limites iterados existem em \mathbb{R} mas são diferentes, podemos concluir que não existe o limite da função dada no ponto $(0, 0)$, tal como já tínhamos visto no Exemplo 1.107.

Observação 1.113 Note-se que pode existir o limite $\lim_{(x,y) \rightarrow (a,b)} f(x, y)$ sem que existam os limites iterados (ver Exercício 5 da Ficha 3).

Mas se $\lim_{(x,y) \rightarrow (a,b)} f(x, y) = L$, e se existirem os limites unidimensionais

$$\lim_{x \rightarrow a} f(x, y) \quad \text{e} \quad \lim_{y \rightarrow b} f(x, y),$$

então existem os limites iterados e são iguais:

$$\lim_{x \rightarrow a} \left(\lim_{y \rightarrow b} f(x, y) \right) = \lim_{y \rightarrow b} \left(\lim_{x \rightarrow a} f(x, y) \right) = L.$$

(Demonstre!)

Continuidade

Definição 1.114 Uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 1$, diz-se contínua em $\mathbf{a} \in D \cap D'$ se $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Por outras palavras, f é contínua em \mathbf{a} se $f(\mathbf{x}^k) \rightarrow f(\mathbf{a})$ para cada sucessão $\{\mathbf{x}^k\} \rightarrow \mathbf{a}$, $k \rightarrow \infty$; ou $\forall \varepsilon > 0 \exists \delta > 0$ tal que

$$\text{se } \mathbf{x} \in D \text{ e } \|\mathbf{x} - \mathbf{a}\|_n \leq \delta \text{ então } \|f(\mathbf{x}) - f(\mathbf{a})\|_m \leq \varepsilon. \quad (1.5)$$

Se f não for contínua em \mathbf{a} diz-se descontínua em \mathbf{a} .

Definição 1.115 Diz-se que $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 1$, é contínua em D se f é contínua em cada $\mathbf{a} \in D$. (Se \mathbf{a} é um ponto isolado a função f considera-se contínua em \mathbf{a} por convenção.)

Observação 1.116 Intuitivamente, f é contínua em \mathbf{a} quando os pontos \mathbf{x} que estão próximos de \mathbf{a} têm imagens $f(\mathbf{x})$ próximas de $f(\mathbf{a})$.

Tendo em conta a Proposição 1.98 temos

Proposição 1.117 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, com coordenadas f_1, f_2, \dots, f_m , e $\mathbf{a} \in D$. Então f é contínua em \mathbf{a} sse f_j é contínua em \mathbf{a} , para todo o $j = 1, 2, \dots, m$.*

Tendo em conta a Proposição 1.103 obtemos o seguinte resultado

Proposição 1.118 *Sejam $f, g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ contínuas em $\mathbf{a} \in D$ e $\lambda \in \mathbb{R}$. Então as funções vectoriais $f + g$, λf e $f \times g$ no caso $m = 3$, e as funções escalares $\langle f, g \rangle$ e $\|f\|$ são contínuas em \mathbf{a} . Além disso, se $m = 1$ e $g(\mathbf{a}) \neq 0$ a função $\frac{f}{g}$ é contínua em \mathbf{a} .*

Proposição 1.119 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $g : E \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ com $f(D) \subset E$. Se f é contínua em \mathbf{a} e g é contínua em $f(\mathbf{a})$, então a função composta $g \circ f$ é contínua em \mathbf{a} .*

Demonstração:

Fixo $\varepsilon > 0$ arbitrário. Pela continuidade de g em $\mathbf{b} = f(\mathbf{a})$ existe $\delta_1 > 0$ tal que

$$\mathbf{y} \in E \text{ e } \|\mathbf{y} - \mathbf{b}\|_m < \delta_1 \text{ então } \|g(\mathbf{y}) - g(\mathbf{b})\|_p < \varepsilon.$$

Fixo tal δ_1 . Pela continuidade de f em \mathbf{a} existe $\delta_2 > 0$ tal que

$$\mathbf{x} \in D \text{ e } \|\mathbf{x} - \mathbf{a}\|_n < \delta_2 \text{ então } \|f(\mathbf{x}) - f(\mathbf{a})\|_m < \delta_1.$$

Assim, tomando $\delta = \delta_2$ tem-se

$$\mathbf{x} \in D \text{ e } \|\mathbf{x} - \mathbf{a}\|_n < \delta \text{ então } \|g(f(\mathbf{x})) - g(f(\mathbf{a}))\| < \varepsilon,$$

o que prova a continuidade de $(g \circ f)$ em \mathbf{a} . ■

Exemplo 1.120 *As funções*

$$f(x, y) = \sin(x^2 y), \quad g(x, y) = \ln(x^2 + y^2), \quad h(x, y) = \frac{e^{x+y}}{x+y}, \quad m(x, y) = \ln|\cos(x^2 + y^2)|.$$

são contínuas em todos os pontos em que estão definidas:

f é contínua em todos os pontos do plano, g em $\mathbb{R}^2 \setminus \{(0, 0)\}$, h é contínua em todos os pontos (x, y) tais que $x + y \neq 0$, m é contínua em todos os pontos para os quais $x^2 + y^2$ não seja múltiplo ímpar de $\pi/2$ (o conjunto de $(x, y) \in \mathbb{R}^2$ tais que $x^2 + y^2 = \frac{n\pi}{2}$, $n = 1, 3, 5, \dots$, é uma família de circunferências centradas na origem).

O exemplo anterior mostra que os pontos de descontinuidade podem ser pontos isolados, ou formar curvas ou famílias de curvas.

Uma função de duas variáveis pode ser contínua em relação a cada uma das variáveis separadamente e ser descontínua quando considerada como função das duas variáveis em conjunto. Este facto é ilustrado pelo exemplo seguinte.

Exemplo 1.121 Consideremos a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

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

Para pontos (x, y) do eixo dos x temos $y = 0$ e $f(x, y) = f(x, 0) = 0$, pelo que a função tem valor constante 0 ao longo de todo o eixo dos x . Deste modo, se fizermos $y = 0$ e considerarmos f unicamente como uma função de x , f é contínua em $x = 0$.

Analogamente, f toma o valor 0 em todos os pontos do eixo dos y , pelo que se fizermos $x = 0$ e considerarmos f unicamente como uma função de y , f é contínua em $y = 0$. Porém, f considerada como função de duas variáveis não é contínua na origem. Com efeito, em cada ponto da recta $y = x$ (excepto na origem) a função toma o valor constante $\frac{1}{2}$, pois

$$f(x, x) = \frac{x^2}{2x^2} = \frac{1}{2};$$

como existem sobre esta recta pontos tão próximos da origem quanto se queira e como $f(0, 0) \neq \frac{1}{2}$, a função não é contínua em $(0, 0)$.

Destacamos as seguintes importantes propriedades das funções contínuas definidas em conjuntos.

Proposição 1.122 Se $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é contínua e $X \subset D$ é um conjunto compacto em \mathbb{R}^n , então a imagem $f(X) := \{f(\mathbf{x}) : \mathbf{x} \in X\}$ é um conjunto compacto.

Demonstração:

Seja $\{f(\mathbf{x}^m)\} \subset f(X)$ uma sucessão de pontos arbitrária. Então $\{\mathbf{x}^m\} \subset X$ e como X é compacto existe uma subsucessão $\{\mathbf{x}^{k_m}\} \rightarrow \mathbf{x}^0$, para algum $\mathbf{x}^0 \in X$. Mas $\{f(\mathbf{x}^{k_m})\}$ é subsucessão de $\{f(\mathbf{x}^m)\}$ e converge para $f(\mathbf{x}^0)$ (justifique!). Donde se conclui que $f(X)$ é compacto. ■

Da proposição anterior sai imediatamente a generalização do Teorema de Weierstrass bem conhecida na Análise Matemática I.

Teorema 1.123 (Weierstrass) Se $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é contínua e $X \subset D$ é um conjunto compacto em \mathbb{R}^n , então f é limitada em X e atinge (em X) o seu valor máximo e mínimo; isto é, existem $\mathbf{a}, \mathbf{b} \in X$, tais que

$$f(\mathbf{a}) = \max_{\mathbf{x} \in X} f(\mathbf{x}) \quad \text{e} \quad f(\mathbf{b}) = \min_{\mathbf{x} \in X} f(\mathbf{x}).$$

Demonstração:

Da proposição anterior resulta, em particular, que $f(X)$ é um conjunto limitado e por conseguinte f é limitada em X (justifique!). Portanto o supremo e o ínfimo de f em X existem e são finitos. Seja, por exemplo, $\sup_{\mathbf{x} \in X} f(\mathbf{x}) = a$. Por definição de supremo existe uma sucessão $\{\mathbf{x}^m\} \subset X$ tal que $\lim_{m \rightarrow \infty} f(\mathbf{x}^m) = a$. Como X é compacto existe uma subsucessão $\{\mathbf{x}^{k_m}\} \rightarrow \mathbf{x}^0$, para algum $\mathbf{x}^0 \in X$, o que implica que $\{f(\mathbf{x}^{k_m})\} \rightarrow f(\mathbf{x}^0) = a$ (justifique!). Analogamente se mostra que o mínimo é atingido. ■

As aplicações contínuas também mantêm a conexidade dos conjuntos

Proposição 1.124 *Se $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é contínua num conjunto conexo (topologicamente ou por arcos) $X \subset D$, então o conjunto $f(X)$ é conexo (topologicamente ou por arcos, respectivamente).*

Demonstração:

Mostremos apenas que $f(X)$ é conexo por arcos, deixando a outra parte como exercício.

Sejam $c, d \in f(X)$ dois pontos diferentes, então existem $\mathbf{a}, \mathbf{b} \in X$, com $\mathbf{a} \neq \mathbf{b}$ tais que $f(\mathbf{a}) = c$ e $f(\mathbf{b}) = d$. Como X é conexo por arcos existe um caminho $r : [0, 1] \rightarrow X$ tal que $r(0) = \mathbf{a}$ e $r(1) = \mathbf{b}$, o que implica $f(r(0)) = f(\mathbf{a}) = c$ e $f(r(1)) = f(\mathbf{b}) = d$. Portanto o caminho $f \circ r$ liga os pontos c e d em $f(X)$. ■

Prolongamento por continuidade

Definição 1.125 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $\mathbf{a} \in D' \setminus D$. Diz-se que a função f é prolongável por continuidade ao ponto \mathbf{a} se existe o limite $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x})$. Neste caso, a função $\tilde{f} : D \cup \{\mathbf{a}\} \rightarrow \mathbb{R}^m$ definida por*

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{se } \mathbf{x} \in D \\ \lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) & \text{se } \mathbf{x} = \mathbf{a}, \end{cases}$$

chama-se prolongamento de f ao ponto \mathbf{a} .

Exemplos 1.126 1. *Seja $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ uma função definida por*

$$f(x, y) = \frac{xy^2}{x^2 + y^2}.$$

Vimos no Exemplo 1.100 que $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$. Logo a função dada é prolongável por continuidade ao ponto $(0, 0)$, e o prolongamento de f ao ponto $(0, 0)$ é a função $\tilde{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

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

2. *Seja $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ uma função definida por*

$$f(x, y) = \frac{x^2 y}{x^4 + y^2}.$$

Vimos no Exemplo 1.109.2 que não existe o limite $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$, logo a função dada não é prolongável por continuidade ao ponto $(0, 0)$.

Capítulo 2

Derivadas de funções escalares

Neste capítulo vamos estudar, em \mathbb{R}^n , os principais conceitos do cálculo diferencial aplicados a funções escalares, nomeadamente, derivada segundo um vector, derivada parcial, diferenciabilidade e gradiente de uma função escalar.

2.1 Derivadas direccionais e derivadas parciais

Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $\mathbf{a} \in \text{int } D$. Pretendemos avaliar a "taxa de variação" de f quando se atribuem "pequenos acréscimos" ao ponto \mathbf{x} , partindo da posição \mathbf{a} . Convém observar já que, enquanto no caso das funções reais de variável real os "acrécimos" 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 $\mathbf{a} \in \text{int } D$). 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 $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $\mathbf{a} \in \text{int } D$. Diz-se que f tem derivada em \mathbf{a} segundo o vector $\mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq \mathbf{0}$, se existir e for finito o limite*

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

O valor finito deste limite chama-se derivada (ou derivada direccionais) de f em \mathbf{a} segundo o vector \mathbf{v} e denota-se por $f'(\mathbf{a}; \mathbf{v})$ ou $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$.

Quando $\mathbf{v} = \mathbf{e}_i$ é um dos vectores coordenados ($i = 1, 2, \dots, n$) e

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

este valor comum chama-se a derivada parcial de f no ponto \mathbf{a} em ordem a x_i (onde x_i representa a i -ésima coordenada da variável \mathbf{x}), e denota-se também por $f'_{x_i}(\mathbf{a})$ ou $\frac{\partial f}{\partial x_i}(\mathbf{a})$.

Observações 2.2 1) *Na maioria dos casos consideramos vectores \mathbf{v} com norma 1. Se $\|\mathbf{v}\| \neq 1$ considera-se o vector $\bar{\mathbf{v}} := \frac{\mathbf{v}}{\|\mathbf{v}\|}$ e usa-se a propriedade*

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

que válida para qualquer $\lambda > 0$. Às vezes é conveniente considerar a derivada $f'(\mathbf{a}; \mathbf{v})$ quando $\mathbf{v} = \mathbf{0}$, logo temos $f'(\mathbf{a}; \mathbf{0}) = 0$.

2) Considerando a função real g , definida numa vizinhança de $h = 0$, por

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

é fácil ver que $f'(\mathbf{a}; \mathbf{v}) = g'_+(0)$, onde $g'_+(0)$ representa 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 vector \mathbf{v} .

Por exemplo, para $f(x, y) = x^2 + y^2$, $\mathbf{a} = (1, 1)$ e $\mathbf{v} = (1, 2)$ temos

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

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

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

3) Na prática, o cálculo da derivada parcial $\frac{\partial f}{\partial x_i}$ faz-se considerando todas as variáveis como se fossem constantes, excepto a i -ésima, e aplicam-se as regras usuais de derivação relativamente a essa variável.

Mostremos a afirmação anterior no espaço \mathbb{R}^2 . Seja $\mathbf{a} = (a_1, a_2)$. Como $\mathbf{e}_1 = (1, 0)$ e $\mathbf{e}_2 = (0, 1)$ então, pela definição de derivada direcciona, temos

$$f'(\mathbf{a}; \mathbf{e}_1) = \lim_{h \rightarrow 0^+} \frac{f(\mathbf{a} + h\mathbf{e}_1) - f(\mathbf{a})}{h} = \lim_{h \rightarrow 0^+} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}.$$

De outro lado,

$$\begin{aligned} -f'(\mathbf{a}; -\mathbf{e}_1) &= -\lim_{h \rightarrow 0^+} \frac{f(\mathbf{a} - h\mathbf{e}_1) - f(\mathbf{a})}{h} = \lim_{h \rightarrow 0^+} \frac{f(\mathbf{a} - h\mathbf{e}_1) - f(\mathbf{a})}{-h} \\ &= \lim_{h \rightarrow 0^-} \frac{f(\mathbf{a} + h\mathbf{e}_1) - f(\mathbf{a})}{h} = \lim_{h \rightarrow 0^-} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}. \end{aligned}$$

Logo

$$\frac{\partial f}{\partial x}(a_1, a_2) = \lim_{h \rightarrow 0} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}.$$

Analogamente se mostra que

$$\frac{\partial f}{\partial y}(a_1, a_2) = \lim_{h \rightarrow 0} \frac{f(a_1, a_2 + h) - f(a_1, a_2)}{h}.$$

Claro que as derivadas direccionais e parciais obedecem as regras de derivação usuais conhecidas da Análise Matemática I.

Proposição 2.3 Sejam $f, g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{a} \in \text{int } D$ 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 $\frac{\partial(\frac{f}{g})}{\partial \mathbf{v}}$ (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) $\frac{\partial\left(\frac{f}{g}\right)}{\partial \mathbf{v}}(\mathbf{a}) = \frac{\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})g(\mathbf{a}) - f(\mathbf{a})\frac{\partial g}{\partial \mathbf{v}}(\mathbf{a})}{(g(\mathbf{a}))^2}.$

Demonstração: Exercício.

Exemplos 2.4 1) Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ 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

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

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

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 $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ 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

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

Exercício 2.5 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 $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(a, b) \in \text{int } D$. Recordemos que o gráfico de f é o conjunto

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

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

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

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 num ponto, mesmo quando elas existem em todas as direcções, não garante a continuidade da função nesse ponto (ver Exemplo 2.6 abaixo). Por isso as derivadas direccionais não constituem uma extensão satisfatória do conceito unidimensional de derivada. Existe uma generalização mais adequada, cuja existência implica a continuidade e, ao mesmo tempo, permite-nos generalizar os principais teoremas da derivação unidimensional ao caso de dimensões superiores. É o chamado *diferencial total* ou simplesmente *diferencial*, que veremos em seguida.

No caso unidimensional a primeira afirmação do parágrafo anterior demonstra-se facilmente pela escolha de um $h \neq 0$ e escrevendo

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

Quando $h \rightarrow 0$ o segundo membro tende para $f'(a) \times 0 = 0$ e por isso $f(a+h) \rightarrow f(a)$. Isto mostra que a existência de $f'(a)$ implica a continuidade de f em a .

Aplicamos agora o mesmo raciocínio a uma função escalar qualquer. 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 $h \rightarrow 0^+$ 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

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

Tal significa que $f(\mathbf{x}) \rightarrow f(\mathbf{a})$ quando $\mathbf{x} \rightarrow \mathbf{a}$ ao longo da direcção \mathbf{v} . Se $f'(\mathbf{a}; \mathbf{v})$ existe para todo o vector \mathbf{v} , então $f(\mathbf{x}) \rightarrow f(\mathbf{a})$ quando $\mathbf{x} \rightarrow \mathbf{a}$ ao longo de todas as direcções. Tal facto parece sugerir que f é contínua em \mathbf{a} , o que não é necessariamente verdade. O seguinte exemplo mostra uma função escalar que admite derivada direccional no ponto $(0, 0)$ segundo qualquer direcção mas que não é contínuo nesse ponto.

Exemplo 2.6 *Seja f uma função escalar definida em \mathbb{R}^2 do seguinte modo:*

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

Sejam $\mathbf{a} = (0, 0)$ e $\mathbf{v} = (u, v)$ um vector qualquer de \mathbb{R}^2 . Para determinar $f'(\mathbf{a}; \mathbf{v})$ consideramos dois casos:

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

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

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

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

Portanto $f'(\mathbf{a}; \mathbf{v})$ existe em todas as direcções \mathbf{v} .

Além disso, Além disso, como no Exemplo 1.121, $f(x, y) \rightarrow 0$ quando $(x, y) \rightarrow (0, 0)$ ao longo de qualquer recta passando pela origem. Contudo, 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 $f(0, 0) = 0$, então a função f não é contínua em $(0, 0)$.

Antes de definirmos diferencial recordemos que, se $f :]\alpha, \beta[\subset \mathbb{R} \rightarrow \mathbb{R}$ é uma função diferenciável num ponto $a \in]\alpha, \beta[$, então, pela definição da derivada $f'(a)$, resulta imediatamente que, para h suficientemente pequeno (tal que $a + h \in]\alpha, \beta[$),

$$\Delta f := f(a + h) - f(a) = f'(a)h + o(h), \quad (2.2)$$

onde a função $o(h)$ verifica $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$ (uma tal função $o(h)$ diz-se um *o-pequeno de Landau* quando $h \rightarrow 0$). 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 (2.2) diz-nos que, para h pequeno, $\Delta f - y(h)$ tende para 0 "mais depressa" do que h quando $h \rightarrow 0$. Logo o acréscimo de f em torno de a pode ser substituído 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).

Exemplo 2.7 Consideremos a função $f(x) = x^2 + x + 1$. Sendo $f(2) = 7$ e $f'(2) = 5$ (já que $f'(x) = 2x + 1$) podemos escrever

$$f(x) = 7 + 5(x - 2) + o(x - 2),$$

com $o(h) = h^2$. Tem-se, obviamente, $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$.

Passemos agora para funções de várias variáveis. Como vimos no Exemplo 2.6 a existência de todas as derivadas direccionais de uma função num ponto não garante a continuidade da função nesse ponto. Esta garantia pode ser dada com uma aproximação linear análoga à considerada acima; daqui chegamos à seguinte definição:

Definição 2.8 Diz-se que a função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável num ponto $\mathbf{a} \in \text{int } D$ se existe uma aplicação linear $\mathbf{L} : \mathbb{R}^n \rightarrow \mathbb{R}$, tal que

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \mathbf{L}(\mathbf{h}) + o(\mathbf{h}), \quad (2.3)$$

onde a função $o(\mathbf{h})$ definida por esta igualdade (para $\mathbf{h} \in \mathbb{R}^n$ tal que $\mathbf{a} + \mathbf{h} \in D$) satisfaz

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

A aplicação linear \mathbf{L} (única pelo Teorema 2.13) chama-se *diferencial* (ou *diferencial total*) de f no ponto \mathbf{a} e denota-se por $df(\mathbf{a})$.

Observação 2.9 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. No exemplo anterior, a derivada de f no ponto 2 é o número 5. Enquanto que o diferencial de f no ponto 2 é a aplicação linear $df(\mathbf{a})(h) = 5h$.

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

Exemplo 2.10 A função $f(x, y, z) = xy + yz + zx - 4$ é diferenciável no ponto $(1, 2, 0)$ (e em todos os pontos de \mathbb{R}^3). De facto, fazendo as transformações algébricas é fácil verificar que

$$f(x, y, z) = -2 + 2(x - 1) + (y - 2) + 3z + (x - 1)(y - 2) + z(y - 2) + z(x - 1).$$

Esta expressão é formada por um termo constante:

$$-2 = f(1, 2, 0),$$

uma forma linear aplicada á diferença $(x, y, z) - (1, 2, 0) = (x - 1, y - 2, z)$:

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

e pelo resto:

$$(x - 1)(y - 2) + z(y - 2) + z(x - 1).$$

Exercício 2.11 Verifique que

$$\lim_{(x, y, z) \rightarrow (1, 2, 0)} \frac{(x - 1)(y - 2) + z(y - 2) + z(x - 1)}{\|(x, y, z) - (1, 2, 0)\|} = 0.$$

Sugestão: aplique a definição de limite.

O leitor atento terá notado que no exemplo anterior não mostrámos como obtivemos $df(1, 2, 0)$. Simplesmente sabemos que a forma linear

$$df(1, 2, 0)(h_1, h_2, h_3) = 2h_1 + h_2 + 3h_3$$

é o diferencial de f porque o resto dividido por $\|(x, y, z) - (1, 2, 0)\|$ tende para 0 (ver Exercício 2.11).

Vejamos algumas propriedades das funções diferenciáveis:

Proposição 2.12 *Sejam $f, g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ funções diferenciáveis em $\mathbf{a} \in \text{int } D$, 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

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

Demonstração:

Mostraremos apenas que, nas condições indicadas, $f + g$ é diferenciável em \mathbf{a} . Como f e g são diferenciáveis em \mathbf{a} , tem-se

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + df(\mathbf{a})(\mathbf{h}) + o_1(\mathbf{h}) \quad \text{e} \quad g(\mathbf{a} + \mathbf{h}) = g(\mathbf{a}) + dg(\mathbf{a})(\mathbf{h}) + o_2(\mathbf{h}),$$

onde

$$\lim_{\|\mathbf{h}\| \rightarrow 0} \frac{o_1(\mathbf{h})}{\|\mathbf{h}\|} = 0 \quad \text{e} \quad \lim_{\|\mathbf{h}\| \rightarrow 0} \frac{o_2(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Então

$$f(\mathbf{a} + \mathbf{h}) + g(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + g(\mathbf{a}) + df(\mathbf{a})(\mathbf{h}) + dg(\mathbf{a})(\mathbf{h}) + o_1(\mathbf{h}) + o_2(\mathbf{h})$$

e

$$\lim_{\|\mathbf{h}\| \rightarrow 0} \frac{o_1(\mathbf{h}) + o_2(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Como a aplicação $df(\mathbf{a}) + dg(\mathbf{a})$ é linear, tem-se o resultado para $f + g$. ■

O teorema que se segue mostra que se existe uma aplicação linear verificando a condição (2.3), então ela é única. Além disso, diz como calcular $df(\mathbf{a})$.

Teorema 2.13 Se $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função diferenciável em $\mathbf{a} \in \text{int } D$ com diferencial $df(\mathbf{a})$, então

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

$$f'(\mathbf{a}; \mathbf{v}) = df(\mathbf{a})(\mathbf{v}); \quad (2.4)$$

(ii) as derivadas parciais $\frac{\partial f}{\partial x_i}(\mathbf{a})$, $i = 1, 2, \dots, n$, existem e para $\mathbf{v} = (v_1, \dots, v_n)$ tem-se

$$f'(\mathbf{a}; \mathbf{v}) = \frac{\partial f}{\partial x_1}(\mathbf{a})v_1 + \frac{\partial f}{\partial x_2}(\mathbf{a})v_2 + \dots + \frac{\partial f}{\partial x_n}(\mathbf{a})v_n; \quad (2.5)$$

(iii) f é contínua em \mathbf{a} .

Demonstração:

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

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + df(\mathbf{a})(\mathbf{h}) + o(\mathbf{h}), \quad (2.6)$$

sendo \mathbf{h} tal que $\mathbf{a} + \mathbf{h} \in D$, onde $\frac{o(\mathbf{h})}{\|\mathbf{h}\|} \rightarrow 0$ quando $\|\mathbf{h}\| \rightarrow 0$. Pomos $\mathbf{h} = \lambda \mathbf{v}$, onde $\lambda > 0$ é suficientemente pequeno. Por linearidade temos

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

e por (2.6)

$$\frac{f(\mathbf{a} + \lambda \mathbf{v}) - f(\mathbf{a})}{\lambda} = df(\mathbf{a})(\mathbf{v}) + \frac{o(\mathbf{h})}{\lambda}. \quad (2.7)$$

Como

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

o segundo membro de (2.7) tende para $df(\mathbf{a})(\mathbf{v})$ quando $\lambda \rightarrow 0^+$. Portanto o primeiro membro tende para o mesmo limite, o que prova (2.4).

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

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

e usando (2.4) temos

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

Logo existem as derivadas parciais $\frac{\partial f}{\partial x_i}(\mathbf{a})$, $i = 1, 2, \dots, n$. Representando $\mathbf{v} = (v_1, \dots, v_n)$ na forma $\mathbf{v} = \sum_{k=1}^n v_k \mathbf{e}_k$ e novamente usando linearidade de $df(\mathbf{a})$ obtemos

$$\begin{aligned} f'(\mathbf{a}; \mathbf{v}) &= df(\mathbf{a})(\mathbf{v}) = df(\mathbf{a})\left(\sum_{k=1}^n v_k \mathbf{e}_k\right) = \sum_{k=1}^n v_k df(\mathbf{a})(\mathbf{e}_k) \\ &= \sum_{k=1}^n v_k f'(\mathbf{a}; \mathbf{e}_k) = \sum_{k=1}^n v_k \frac{\partial f}{\partial x_k}(\mathbf{a}). \end{aligned}$$

(iii) A fórmula (2.3) pode agora escrever-se na forma

$$f(\mathbf{x}) = f(\mathbf{a}) + \langle df(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle + o(\mathbf{x} - \mathbf{a}), \quad (2.8)$$

onde $o(\mathbf{x} - \mathbf{a}) \rightarrow 0$ quando $\mathbf{x} \rightarrow \mathbf{a}$.

De (2.8) tem-se

$$|f(\mathbf{x}) - f(\mathbf{a})| = |\langle df(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle + o(\mathbf{x} - \mathbf{a})|.$$

Aplicando a desigualdade triangular e a desigualdade de Cauchy-Schwarz obtemos

$$0 \leq |f(\mathbf{x}) - f(\mathbf{a})| \leq \|df(\mathbf{a})\| \|\mathbf{x} - \mathbf{a}\| + |o(\mathbf{x} - \mathbf{a})|.$$

Isto mostra que $f(\mathbf{x}) \rightarrow f(\mathbf{a})$ quando $\mathbf{x} \rightarrow \mathbf{a}$, pelo que f é contínua em \mathbf{a} . ■

Observação 2.14 O teorema anterior afirma, em particular, que se f é uma função diferenciável num ponto \mathbf{a} do interior do seu domínio, então existem todas as derivadas parciais de f em \mathbf{a} . No entanto, é importante notar que uma função pode possuir todas as derivadas parciais num certo ponto \mathbf{a} e não ser diferenciável nesse ponto. Com efeito, a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

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

tem todas as derivadas parciais no ponto $(0, 0)$ (justifique), mas não é diferenciável neste ponto porque não é contínua em $(0, 0)$ (ver Exemplo 1.109).

Exemplo 2.15 Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

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

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

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

$$\lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h^2 \sin \frac{1}{|h|}}{h} = 0$$

então $\frac{\partial f}{\partial x}(0, 0) = 0$, e também $\frac{\partial f}{\partial y}(0, 0) = 0$ porque f não se altera com a troca de x com y . Assim, por (2.5), $df(\mathbf{a})(\mathbf{h}) = 0$, qualquer que seja o vector $\mathbf{h} \in \mathbb{R}^2$. Teremos agora que mostrar que f satisfaz a definição de diferenciabilidade (Definição 2.8). Pondo em (2.3) $\mathbf{a} = (0, 0)$ e $\mathbf{h} = (x, y)$, basta verificar que se tem $f(x, y) = o(x, y)$ com $\lim_{(x,y) \rightarrow (0,0)} \frac{o(x,y)}{\|(x,y)\|} = 0$: com efeito

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

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

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

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

Interpretação geométrica

Da mesma forma que a diferenciabilidade de uma função real de variável real f , num ponto x_0 interior do seu domínio, está associada à existência de uma recta tangente ao gráfico de f no ponto $(x_0, f(x_0))$, também a diferenciabilidade de uma função escalar $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, num ponto $(x_0, y_0) \in \text{int } D$ está associada à existência de um plano tangente ao gráfico dessa função escalar no ponto $(x_0, y_0, f(x_0, y_0))$. Assim, em \mathbb{R}^2 , a aplicação linear $df(x_0, y_0)$ tem por gráfico um plano não vertical, passando pela origem, que translacionado de $(x_0, y_0, f(x_0, y_0))$ é tangente ao gráfico de f nesse ponto. Nomeadamente, pela definição do diferencial podemos escrever a equação do plano tangente (no ponto (x_0, y_0, z_0)):

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0). \quad (2.9)$$

Uma condição suficiente para a diferenciabilidade

Pelo Teorema 2.13 se f é diferenciável em \mathbf{a} então todas as derivadas parciais $\frac{\partial f}{\partial x_1}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a})$ existem. Contudo, como vimos na Observação 2.14, a existência de todas estas derivadas parciais não implica necessariamente que f seja diferenciável em \mathbf{a} . O teorema seguinte diz-nos sob que condições podemos concluir que f é diferenciável.

Teorema 2.16 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $\mathbf{a} \in \text{int } D$. 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:

Vamos mostrar o teorema no caso $n = 2$. Seja $\mathbf{a} = (a, b)$. Dado $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$ tal que $(a, b + h_2), (a + h_1, b + h_2) \in D$, representamos o acréscimo

$$\begin{aligned} f(a + h_1, b + h_2) - f(a, b) &= [f(a + h_1, b + h_2) - f(a, b + h_2)] \\ &\quad + [f(a, b + h_2) - f(a, b)]. \end{aligned} \quad (2.10)$$

Fixando $x = a$ a função $f(a, y)$, de variável real y , tem derivada próximo de b , que é a derivada parcial $f'_y(a, y)$. Aplicando o Teorema de Lagrange dos acréscimos finitos encontramos um ponto η entre b e $b + h_2$ (seja $\eta = b + \theta h_2$ para $0 < \theta < 1$) tal que

$$f(a, b + h_2) - f(a, b) = f'_y(a, b + \theta h_2) h_2.$$

Fixamos agora $y = b + h_2$ e consideramos a função $f(x, y + h_2)$ de variável x . Como em torno de a existe a derivada desta função, e é igual a $f'_x(x, b + h_2)$, novamente pelo Teorema de Lagrange existe $\xi = a + \tau h_1$ com $0 < \tau < 1$ tal que

$$f(a + h_1, b + h_2) - f(a, b + h_2) = f'_x(a + \tau h_1, b + h_2) h_1.$$

Observamos que θ e τ dependem de h_1 e h_2 (logo de \mathbf{h}) e $(a, b + \theta h_2) \rightarrow (a, b)$, $(a + \tau h_1, b + h_2) \rightarrow (a, b)$ quando $\|\mathbf{h}\| \rightarrow 0$. (Justifique!) Pela continuidade das derivadas parciais temos

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} f'_x(a + \tau h_1, b + h_2) = f'_x(a, b)$$

e

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

As últimas igualdades significam que

$$f'_x(a + \tau h_1, b + h_2) = f'_x(a, b) + \alpha(\mathbf{h}) \quad (2.11)$$

e

$$f'_y(a, b + \theta h_2) = f'_y(a, b) + \beta(\mathbf{h}) \quad (2.12)$$

onde $\alpha(\mathbf{h}) \rightarrow 0$ e $\beta(\mathbf{h}) \rightarrow 0$ quando $\mathbf{h} \rightarrow 0$. Substituindo (2.11) e (2.12) em (2.10) temos

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

Uma vez que

$$\frac{|h_1|}{\|\mathbf{h}\|} = \frac{|h_1|}{\sqrt{h_1^2 + h_2^2}} \leq 1$$

e analogamente

$$\frac{|h_2|}{\|\mathbf{h}\|} = \frac{|h_2|}{\sqrt{h_1^2 + h_2^2}} \leq 1$$

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

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

Daqui sai a diferenciabilidade de f no ponto (a, b) . ■

Observação 2.17 As condições do teorema anterior podem ser um pouco enfraquecidas. Nomeadamente, para diferenciabilidade basta supôr continuidade no ponto \mathbf{a} de $n - 1$ derivadas parciais enquanto a derivada restante pode só existir em \mathbf{a} .

Definição 2.18 Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{a} \in \text{int } D$. 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 contínuas 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 2.16 sai que qualquer função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 num conjunto aberto $U \subset D$ é diferenciável em todos os pontos de U .

Exercício 2.19 Mostre que a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = ye^x + x^3 + 1$ é diferenciável em todo $(x, y) \in \mathbb{R}^2$.

2.3 Gradiente

Definição 2.20 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $\mathbf{a} \in \text{int } D$. Suponhamos que f tem todas as derivadas parciais em \mathbf{a} . Chama-se gradiente de f em \mathbf{a} ao vector cujas coordenadas são as n derivadas parciais de f calculadas em \mathbf{a} , e representa-se por $\nabla f(\mathbf{a})$ (ou $\text{grad } f(\mathbf{a})$), ou seja*

$$\nabla f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \frac{\partial f}{\partial x_2}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right).$$

Usa-se ∇ (lê-se "nabla") para denotar o operador simbólico $\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right)$ que encontraremos várias vezes ao longo destes apontamentos.

Observação 2.21 *Se a função f é diferenciável em \mathbf{a} , pelo Teorema 2.13, temos*

$$f'(\mathbf{a}; \mathbf{v}) = df(\mathbf{a})(\mathbf{v}) = \langle \nabla f(\mathbf{a}), \mathbf{v} \rangle. \quad (2.13)$$

Exemplo 2.22 *O gradiente da função*

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

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

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

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

$$\nabla f(1, 2, 0) = (2, 1, 3).$$

Portanto, o diferencial de f em $(1, 2, 0)$ é a aplicação linear

$$df(\mathbf{a})(v_1, v_2, v_3) = \langle (2, 1, 3), (v_1, v_2, v_3) \rangle = 2v_1 + v_2 + 3v_3.$$

Teorema 2.23 *Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável em $\mathbf{a} \in \text{int } D$ e suponhamos que $\nabla f(\mathbf{a}) \neq 0$. Então, perto de \mathbf{a} e tomando como ponto de partida o ponto \mathbf{a} ,*

- (i) f é crescente na direcção de qualquer vector que faça um ângulo agudo com $\nabla f(\mathbf{a})$;*
- (ii) na direcção do vector $\nabla f(\mathbf{a})$ a função f cresce mais rapidamente do que em qualquer outra direcção.*

Demonstração:

Sem perda de generalidade, supomos que $\|\mathbf{v}\| = 1$. Conforme o sentido físico da derivada de uma função de varável real, f é crescente ou decrescente em torno do ponto \mathbf{a} na direcção do vector \mathbf{v} dependendo se $f'(\mathbf{a}; \mathbf{v})$ é positiva ou negativa, respectivamente. Então, pela fórmula 2.13

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

sendo θ 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. 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 $\mathbf{v} = \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$. ■

Observação 2.24 Abaixo veremos mais uma propriedade geométrica do gradiente, a qual envolve os conjuntos de nível de uma função de várias variáveis.

Exemplo 2.25 Suponhamos que a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = ye^x + x^3 + 1$$

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

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 que a temperatura sobe mais rapidamente. Nas direcções que fazem um ângulo agudo com o vector $(3, 1)$ a temperatura também sobe, mas não tão rapidamente. Nas direcções que fazem um ângulo obtuso com o vector $(3, 1)$ a temperatura desce.

Exercício 2.26 Suponha que a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

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

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

- Determine a taxa de variação de f no ponto $P = (2, 0)$ na direcção do vector $\overrightarrow{PQ} = Q - P$ onde $Q = (\frac{1}{2}, 2)$.
- Partindo do ponto P qual é a direcção em que a temperatura aumenta mais rapidamente?
- Qual é a máxima taxa de variação de f no ponto P ?

2.4 Derivada da função composta

Pela regra de derivação para a função composta de duas funções reais de variável real f e g sabemos que a derivada de $h = f \circ g$ 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 temos

Teorema 2.27 Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função escalar diferenciável em $\mathbf{a} = (a_1, \dots, a_n) \in \text{int } D$ e $g_i : [\alpha, \beta] \subset \mathbb{R} \rightarrow \mathbb{R}$, $\alpha < \beta$, $i = 1, \dots, n$, funções diferenciáveis num ponto $t_0 \in]\alpha, \beta[$, $a_i = g_i(t_0)$, $i = 1, \dots, n$. Então a função composta $\Phi(t) = f(g_1(t), \dots, g_n(t))$ é diferenciável em t_0 e tem-se

$$\Phi'(t_0) = \frac{\partial f}{\partial x_1}(x_0)g'_1(t_0) + \dots + \frac{\partial f}{\partial x_n}(x_0)g'_n(t_0). \quad (2.14)$$

Demonstração.

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

$$f(a_1 + h_1, \dots, a_n + h_n) - f(a_1, \dots, a_n) = \frac{\partial f}{\partial x_1}(\mathbf{a})h_1 + \dots + \frac{\partial f}{\partial x_n}(\mathbf{a})h_n + o(\mathbf{h}). \quad (2.15)$$

Atribuímos a h_i , $i = 1, \dots, n$, 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 $g_i(t_0 + \tau) - g_i(t_0) \rightarrow 0$ quando $\tau \rightarrow 0$. 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 (2.15) na forma:

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

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

$$\frac{\Phi(t_0 + \tau) - \Phi(t_0)}{\tau} = \frac{\partial f}{\partial x_1}(\mathbf{a}) \frac{g_1(t_0 + \tau) - g_1(t_0)}{\tau} + \dots + \frac{\partial f}{\partial x_n}(\mathbf{a}) \frac{g_n(t_0 + \tau) - g_n(t_0)}{\tau} + \frac{o(\mathbf{h})}{\tau}. \quad (2.16)$$

Como

$$\lim_{\tau \rightarrow 0} \frac{g_i(t_0 + \tau) - g_i(t_0)}{\tau} = g'_i(t_0)$$

e

$$\begin{aligned}\lim_{\tau \rightarrow 0} \frac{o(\mathbf{h})}{\tau} &= \lim_{\tau \rightarrow 0} \frac{o(\mathbf{h})}{\|\mathbf{h}\|} \frac{\|\mathbf{h}\|}{\tau} \\ &= \lim_{\tau \rightarrow 0} \frac{o(\mathbf{h})}{\|\mathbf{h}\|} \lim_{\tau \rightarrow 0} \frac{\sqrt{(g_1(t_0 + \tau) - g_1(t_0))^2 + \dots + (g_n(t_0 + \tau) - g_n(t_0))^2}}{\tau} \\ &= \lim_{\tau \rightarrow 0} \frac{o(\mathbf{h})}{\|\mathbf{h}\|} \lim_{\tau \rightarrow 0} \sqrt{\left(\frac{g_1(t_0 + \tau) - g_1(t_0)}{\tau}\right)^2 + \dots + \left(\frac{g_n(t_0 + \tau) - g_n(t_0)}{\tau}\right)^2} = 0,\end{aligned}$$

concluimos que o limite

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

existe e passando em (2.16) ao limite mostramos a fórmula (2.14). ■

Observação 2.28 Recordando a definição de gradiente podemos representar a fórmula (2.14) na forma vectorial

$$\Phi'(t_0) = \langle \nabla f(\mathbf{x}_0), g'(t_0) \rangle, \quad (2.17)$$

onde

$$g'(t_0) = (g'_1(t_0), \dots, g'_n(t_0)).$$

Na fórmula (2.17) temos que substituir \mathbf{x}_0 pelo vector $(g_1(t_0), \dots, g_n(t_0))$.

Exemplo 2.29 Sejam $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ e $g_i: \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, 2, 3$, definidas, respectivamente, por

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

Temos

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

e

$$g'(t) = (-\sin t, \cos t, 1),$$

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

$$\begin{aligned}(f \circ g)'(t) &= \langle \nabla f(g(t)), g'(t) \rangle = \langle (-2 \cos t, -2 \sin t, 2t), (-\sin t, \cos t, 1) \rangle \\ &= 2 \cos t \sin t - 2 \sin t \cos t + 2t = 2t.\end{aligned}$$

Observação 2.30 Se as variáveis x_1, \dots, x_n , no teorema anterior, dependem de várias variáveis, sejam u_1, \dots, u_k , então a derivada parcial da função composta em ordem a u_j , $j = 1, \dots, k$, (se existir) pode ser encontrada pela fórmula

$$\frac{\partial f}{\partial u_j} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial u_j} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial u_j} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial u_j}. \quad (2.18)$$

Aqui denotamos a função composta pela mesma letra f . Esta fórmula é conhecida por regra da cadeia.

Usando a observação anterior 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)

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i \quad (2.19)$$

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

Demonstração.

Se $x_i = x_i(u_1, \dots, u_k)$ são funções (de várias variáveis u_1, \dots, 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 (2.18)) 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 transformações 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 (2.19). ■

2.5 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

$$F(x, y) = 0. \quad (2.20)$$

Nomeadamente, se dado $D \subset \mathbb{R}$ e a cada $x \in D$ está associado um (único) y satisfazendo a equação (2.20) diz-se que: em D está definida *implicitamente* uma função $y = f(x)$ tal que

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

Por exemplo, para $F(x, y) = x^2 + y^2 - 25$ existem várias funções definidas no segmento $[-1, 1]$ que satisfazem a equação (2.20):

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

(de facto, existe uma infinidade de funções, sendo apenas duas delas contínuas: f_1 e f_2). Para definir uma função à custa da equação (2.20) 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 define-se 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 $x^2 + y^2 = 25$). Então existe uma vizinhança rectangular em torno 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 (2.20) tivermos a restrição de $F(x, y) = x^2 + y^2 - 25$ 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 única (por exemplo, à direita do ponto $(1, 0)$ a equação (2.20) 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 correctamente uma função implícita na vizinhança de um ponto dado. Observamos, no entanto, que apenas em situações bastantamente raras (como no 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 (2.20) não pode ser resolvida relativamente a y (nem a x) na vizinhança de nenhum ponto.

Teorema 2.32 *Seja $F : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^2$, uma função contínua numa vizinhança de um ponto $(x_0, y_0) \in \text{int } D$ tal que a derivada parcial $F'_y(x_0, y_0)$ existe e é contínua em (x_0, y_0) . Se $F'_y(x_0, y_0) \neq 0$ então existem $\delta > 0$ e $\varepsilon > 0$ tais que para cada $x \in]x_0 - \delta, x_0 + \delta[$ a equação (2.20) tem uma única solução (seja $y = f(x) \in]y_0 - \varepsilon, y_0 + \varepsilon[$), e a função $f :]x_0 - \delta, x_0 + \delta[\rightarrow$*

$]y_0 - \varepsilon, y_0 + \varepsilon[$, $f(x_0) = y_0$, 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:

$$f'(x_0) = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}. \quad (2.21)$$

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 em torno de x_0 , e verifica-se a fórmula (2.21) em todos os pontos dessa vizinhança (não só em x_0).

Demonstração.

Apresentamos apenas um esboço da 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)

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

Representando por $F'_y(x_0, y_0)$ a derivada usual da função $y \mapsto F(x_0, y)$, a última condição diz-nos que esta função é (monótona) 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 em $(x_0, y_0 \pm \varepsilon) \in U$ conclui-se que a função F mantém os mesmos sinais nalgumas vizinhanças desses pontos. Por outras palavras, existe $0 < \delta < \varepsilon$, tal que

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

e

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

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) :

$$F(x_0 + \tau, y_0 + h) - F(x_0, y_0) = F'_x(x_0, y_0)\tau + F'_y(x_0, y_0)h + o(\tau, h). \quad (2.22)$$

Pomos $h = f(x_0 + \tau) - f(x_0) \rightarrow 0$ quando $h \rightarrow 0$ e tomamos em conta que

$$F(x_0, y_0) = 0 \quad \text{e} \quad F(x_0 + \tau, y_0 + h) = F(x_0 + \tau, f(x_0 + \tau)) = 0.$$

De (2.22) sai que

$$F'_x(x_0, y_0)\tau + F'_y(x_0, y_0)(f(x_0 + \tau) - f(x_0)) + o(\tau, h) = 0.$$

Depois de dividir por τ obtemos

$$F'_x(x_0, y_0) + F'_y(x_0, y_0) \frac{f(x_0 + \tau) - f(x_0)}{\tau} + \frac{o(\tau, h)}{\tau} = 0. \quad (2.23)$$

Como

$$\frac{o(\tau, h)}{\tau} = \frac{o(\tau, h)}{\sqrt{\tau^2 + h^2}} \frac{\sqrt{\tau^2 + h^2}}{\tau} = \frac{o(\tau, h)}{\sqrt{\tau^2 + h^2}} \sqrt{1 + \left(\frac{f(x_0 + \tau) - f(x_0)}{\tau} \right)^2} \rightarrow 0,$$

então, passando ao limite em (2.23), obtemos

$$F'_x(x_0, y_0) + F'_y(x_0, y_0) f'(x_0) = 0,$$

donde sai (2.21) (porque $F'_y(x_0, y_0) \neq 0$). ■

Notemos que a fórmula (2.21) 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 explícita.

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 $F'_y(0, 0) = 1 \neq 0$. 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, $f'(0) = 0$ (isto é, a recta $y = 0$ é tangente ao gráfico da função f na origem).

Observação 2.34 *A fórmula (2.21) permite escrever a equação da recta tangente ao gráfico da função definida implicitamente. Pois, da Análise Matemática I, sabemos que a equação da recta tangente ao gráfico da função $y = f(x)$ num ponto (x_0, y_0) , $y_0 = f(x_0)$, é dada por*

$$y - y_0 = f'(x_0)(x - x_0). \quad (2.24)$$

Se $y = f(x)$ é uma função definida implicitamente em torno de (x_0, y_0) pela equação $F(x, y) = 0$, supondo que $F'_y(x_0, y_0) \neq 0$, então substituindo $f'(x_0)$ em (2.24) pela fórmula (2.21) obtemos a equação da recta tangente na forma simétrica:

$$\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y - y_0) = 0. \quad (2.25)$$

Observação 2.35 Suponhamos que F é continuamente diferenciável numa vizinhança de (x_0, y_0) com $F'_y(x_0, y_0) = 0$ mas $F'_x(x_0, y_0) \neq 0$. Então, aplicando o teorema anterior, existe uma única função $x = g(y)$, contínua e diferenciável, definida implicitamente em torno de y_0 , cuja derivada pode ser calculada pela fórmula (análoga a (2.21)):

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

Sendo $x - x_0 = g'(y_0)(y - y_0)$ a equação da recta tangente ao gráfico da função $x = g(y)$ no ponto (x_0, y_0) facilmente passamos à forma simétrica dessa equação (2.25) escrita em função das derivadas parciais de F .

Por outro lado, se ambas as derivadas $F'_x(x_0, y_0)$ e $F'_y(x_0, y_0)$ são iguais a zero (isto é $\nabla F(x_0, y_0) = 0$) então, no ponto (x_0, y_0) , não existe uma recta tangente bem definida (tal situação acontece por exemplo no caso de autointersecções da curva $F(x, y)$).

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

Teorema 2.36 Seja $F : D \times \mathbb{R} \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$, 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

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

Supomos também que a derivada parcial $F'_z(\mathbf{x}, z)$ existe e é 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) , f é diferenciável em \mathbf{a} , e as derivadas parciais em ordem de x_i , $i = 1, \dots, n$, são dadas por

$$f'_{x_i}(\mathbf{a}) = -\frac{F'_{x_i}(\mathbf{a}, z_0)}{F'_z(\mathbf{a}, z_0)}. \quad (2.26)$$

Tal como na Observação 2.34, de (2.26), e considerando $n = 2$ e $\mathbf{a} = (x_0, y_0)$, podemos obter a equação do plano tangente ao gráfico da função $z = f(x, y)$ num ponto (x_0, y_0, z_0) , $z_0 = f(x_0, y_0)$, em termos das derivadas parciais de F , sendo $F(x, y, z) = 0$ a equação implícita que define a função $z = f(x, y)$. Pois, recordando a definição do plano tangente na forma explícita (2.9) e substituindo os valores das derivadas, de (2.26) chegamos a:

$$\frac{\partial F}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0)(z - z_0) = 0. \quad (2.27)$$

Observação 2.37 Tal como na Observação 2.35, 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 , respectivamente, e a equação do plano tangente (2.27) 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 $\nabla F(x_0, y_0, z_0) = 0$.

2.6 Interpretação geométrica do gradiente

As equações (2.25) e (2.25), obtidas acima, permitem justificar o sentido geométrico do vector gradiente de uma função de duas ou três variáveis.

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

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

onde c é um dos valores que f toma. No caso $n = 2$ ou $n = 3$ os conjuntos de nível chamam-se, como sabemos, respectivamente, linhas de nível ou superfícies de nível.

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

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

Os conjuntos de nível $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = c\}$, para $c \geq 0$, são esferas centradas na origem com raio \sqrt{c} (o conjunto de nível 0 é constituído unicamente pela origem).

Se $f(x, y, z)$ representar a temperatura do ponto (x, y, z) de \mathbb{R}^3 , as superfícies de nível de f , para $c > 0$, são constituídas pelos pontos de \mathbb{R}^3 que estão à temperatura c e formam um conjunto de esferas concêntricas.

Teorema 2.39 *Seja $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^3$, uma função contínua numa vizinhança de $\mathbf{a} = (x_0, y_0, z_0) \in \text{int } D$, diferenciável nesse ponto e tal que $\nabla f(\mathbf{a}) \neq \mathbf{0}$. Então $\nabla f(\mathbf{a})$ é um vector ortogonal (normal) ao plano tangente à superfície de nível $L(c)$, $c = f(\mathbf{a})$, no ponto \mathbf{a} . Analogamente, se $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^2$, é diferenciável num ponto $\mathbf{a} = (x_0, y_0) \in \text{int } D$ e tal que $\nabla f(\mathbf{a}) \neq \mathbf{0}$, logo $\nabla f(\mathbf{a})$ é ortogonal (normal) à recta tangente à respectiva linha de nível $\{(x, y) \in \mathbb{R}^2 : f(x, y) = f(x_0, y_0)\}$ no ponto \mathbf{a} .*

Demonstração:

Para uma função f , definida em \mathbb{R}^3 , a superfície de nível tem a forma

$$L(c) = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = c\}.$$

Seja $F(x, y, z) = f(x, y, z) - c$ (uma função contínua em torno de \mathbf{a} , diferenciável em \mathbf{a} e tal que pelo menos uma das derivadas parciais $F'_x(x_0, y_0, z_0)$, $F'_y(x_0, y_0, z_0)$ ou $F'_z(x_0, y_0, z_0)$ é diferente de zero (justifique!). De acordo com a Observação 2.37, a equação da superfície de nível $L(c)$ pode ser considerada como o gráfico de uma função implícita de duas variáveis (de (x, y) ou (x, z) ou (y, z)). Em qualquer dos casos (2.27) é a equação do plano tangente ao gráfico da função no ponto $\mathbf{a} = (x_0, y_0, z_0)$. Na forma vectorial 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 a todos vectores $\mathbf{x} - \mathbf{a}$ que pertencem ao plano tangente.

No caso de uma função f definida num conjunto de \mathbb{R}^2 os raciocínios são os mesmos, só que usamos a Observação 2.35 e a equação da recta tangente na forma (2.25). ■

Exemplo 2.40 Para a função f do Exemplo 2.38 determinemos um vector ortogonal à superfície de nível de f que passa pelo ponto $(1, 2, 3)$, uma equação desta superfície e uma equação do plano tangente a essa superfície no referido ponto.

Como $f(1, 2, 3) = 14$, a tal superfície tem equação $x^2 + y^2 + z^2 = 14$ e é portanto uma esfera centrada na origem e raio $\sqrt{14}$. Assim, o vector perpendicular a esta esfera no ponto $(1, 2, 3)$ é o vector $\nabla f(1, 2, 3) = (2, 4, 6)$ e a equação do plano tangente é $\langle (x, y, z) - (1, 2, 3), (2, 4, 6) \rangle = 0$, isto é, $x + 2y + 3z = 14$.

Bibliografia

- [1] T. Apostol, *Cálculo*, vol. 2, Editorial Reverté, 1996.
- [2] J. Campos Ferreira, *Introdução à Análise em \mathbb{R}^n* , DMIST, 2003.
- [3] R. Courant, *Differential and Integral Calculus*, Vol. 2, Chap. 1-5, Wiley-Interscience, 1988.
- [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 Calouste 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.