

Noções Fundamentais do Cálculo Diferencial e Integral

ÍNDICE GERAL

1. Limite e continuidade de funções vectoriais	1
1.1. Pontos, vectores e abertos	1
1.2. Exemplos de funções vectoriais	2
1.3. Limite e continuidade de funções vectoriais	3
2. Diferenciabilidade de funções vectoriais	5
2.1. Derivada parcial e gradiente. Interpretação geométrica	5
2.2. Diferenciabilidade	6
2.3. Polinómio de Taylor	10
3. Aplicações da diferenciabilidade	11
3.1. Teoremas da função implícita e da função inversa	11
3.2. Dependência funcional	13
3.3. Curvas e Superfícies diferenciáveis	14
4. Extremos relativos	16
4.1. Condições de primeira ordem	16
4.2. Condições de segunda ordem	18
4.3. Extremos condicionados	20
ÍNDICE REMISSIVO	22

1. LIMITE E CONTINUIDADE DE FUNÇÕES VECTORIAIS

1.1. **Pontos, vectores e abertos.** Desde o ponto de vista da *Álgebra Linear* podemos considerar \mathbb{R}^n como um *espaço vectorial* ou como um *espaço afim*, mas que ninguém se assuste com este começo, pois isto não é mais de que uma forma aristocrática de dizer, que os seus elementos podem ser considerados como vectores ou como pontos. Ficará claro esta ambiguidade, denotando os pontos e os vectores com letras mais escuras.

Quando se fixa o sistema de referência canónico, cada $\mathbf{x} \in \mathbb{R}^n$, está representado pelas suas coordenadas. No caso dos pontos escreveremos $\mathbf{x} = (x_1, \dots, x_n)$, enquanto que para vectores as convenções da álgebra linear exigem que estes sejam descritos por matrizes $n \times 1$ em vez de $1 \times n$, i.e. sejam vectores coluna. Por razões tipográficas seremos forçados a escrever $\mathbf{x} = [x_1 \cdots x_n]^t$, onde o superíndice t indica a operação de transposição. O símbolo $\mathbf{0}$ representará o ponto ou o vector com todas as suas coordenadas nulas. Além disso, no espaço vectorial \mathbb{R}^n podemos medir comprimentos e ângulos, através da noção de *produto interno* definido por $\mathbf{x}|\mathbf{y} = \mathbf{x}^t\mathbf{y}$, pelo que \mathbb{R}^n é um *espaço euclideano*. A *norma* de um vector \mathbf{x} é dada por $\|\mathbf{x}\| = \sqrt{\mathbf{x}|\mathbf{x}}$, e o *ângulo* entre dois vectores \mathbf{x} e \mathbf{y} é o número real $\alpha \in [0, \pi]$ tal que $\mathbf{x}|\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\alpha$. Com estes conceitos podemos estabelecer uma geometria, que designaremos por \mathbb{R}^n -*metria*, onde a *distância* entre dois pontos, está determinada pelo comprimento do vector que determinam, i.e. se $\mathbf{a} = (a_1, \dots, a_n)$ e $\mathbf{b} = (b_1, \dots, b_n)$, então a distância entre \mathbf{a} e \mathbf{b} vem dada por

$$d(\mathbf{a}, \mathbf{b}) = \|\mathbf{b} - \mathbf{a}\| = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}.$$

Desta forma expressamos a nossa fé no *teorema de pitágoras* para além do caso bidimensional.

Enunciados todas estes entes matemáticos conhecidos, vamos pronunciar uma palavra que costuma causar pavor entre estudantes da Licenciatura em Matemática: *topologia*. Este conceito marca a diferença abismal entre o cálculo de uma e várias variáveis, e impede ou dificulta enormemente a escrita de um livro de este último tipo de cálculo, que satisfaça igualmente a estudantes com ambições utilitárias e a matemáticos agudos.

A dificuldade topológica em \mathbb{R}^n reside no facto de para $n \geq 2$ existir muito espaço, que permite aos objectos geométricos que lá vivem o poder-se retorcer e desfocar de muitas formas, criando com tais contorcionismos uma variedade complexa.

Também as propriedades locais de \mathbb{R}^n , que são básicas para estudar a continuidade e diferenciabilidade, se vêm afectadas pelas mudanças qualitativas topológicas no caso multidimensional. Por exemplo, os únicos subconjuntos de \mathbb{R} de uma peça só, são os intervalos, pelo que ao aproximar-nos de um ponto, somente podemos fazê-lo pela sua direita ou pela sua esquerda, se está na fronteira, e pelos dois lados se está no interior. No entanto, por exemplo, em \mathbb{R}^2 , não podemos aproximar-nos de $(0, 0)$ por elementos do conjunto $A = \{(x, y) \in \mathbb{R}^2 : x^4 \leq y \leq x^2\}$, pelo norte, pelo sul, pelo este, pelo oeste, ou por outra

qualquer direcção linear (cf. figura 1). Desta forma, existem não somente infinitas direcções, mas também caminhos curvos que não nos podemos esquecer. Neste curso não vamos estu-

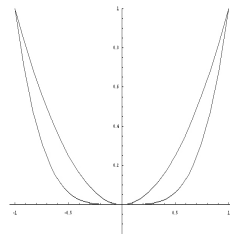


FIGURA 1. Conjunto A

dar a estrutura de espaço topológico genérico, mas tão somente os abertos, que são as peças básicas, no caso particular de \mathbb{R}^n . Estes conjuntos são os que contêm de alguma forma o que podemos designar pela vizinhança de cada um dos seus pontos. Geometricamente são os que não contêm pontos da sua fronteira.

Definição 1.1. Definimos *bola aberta em \mathbb{R}^n* de centro \mathbf{a} e raio r como o conjunto $B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) < r\}$. A partir delas criamos a definição geral, i.e. dizemos que $U \subset \mathbb{R}^n$ é *um aberto* se para todo $\mathbf{a} \in U$, existe $r > 0$ tal que $B_r(\mathbf{a}) \subset U$.

1.2. Exemplos de funções vectoriais. Vamos apresentar alguns exemplos de funções reais de várias variáveis reais que correspondem a noções geométricas, físicas, algébricas ou puramente matemáticas, que os estudantes já conhecem.

Exemplo 1.1. A *área de um rectângulo* expressa-se em termos do comprimento, x , e da largura, y , por $s(x, y) = xy$. Da mesma forma, o *volume de um paralelepipedo rectângulo* vem dado por $v(x, y, z) = xyz$, onde x, y, z correspondem ao comprimento, largura e altura do paralelepipedo.

Exemplo 1.2. A fórmula $pv = RT$ expressa a relação entre o volume v de uma quantidade de gás, temperatura T e pressão p .

Exemplo 1.3 (Definição por composição). A função de expressão analítica

$$f(x, y, z) = \ln(1 - x^2 - y^2 - z^2)$$

está definida em $D = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 1\}$, que è a esfera centrada em $(0, 0, 0)$ de raio 1.

Exemplo 1.4 (Soma Algébrica de funções). A função de expressão analítica

$$f(x, y, z) = \phi(x) + \phi(y) + \phi(z),$$

onde ϕ é uma função real de variável real definida em $[0, 1]$, está definida no cubo

$$D = \{(x, y, z) \in \mathbb{R}^3 : x \in [0, 1], y \in [0, 1], z \in [0, 1]\}.$$

Definição 1.2. Definimos *função vectorial*, \mathbf{F} , em \mathbb{R}^m com domínio $D \subset \mathbb{R}^n$, $n, m \in \mathbb{N}$ a uma aplicação de D em \mathbb{R}^m com expressão analítica

$$\mathbf{F}(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)),$$

onde as funções f_j com $j = 1, \dots, m$ são reais e com domínio D .

Para representarmos o *gráfico* de funções reais de duas variáveis, f , num referencial ortonormado $OXYZ$, i.e.

$$\mathcal{G} = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y), (x, y) \in D\},$$

devemos começar por representar as suas *curvas de nível*, i.e a intersecção de \mathcal{G} com planos paralelos a OXY . Bem como a intersecção com os planos OXZ ou OYZ .

Exemplo 1.5. Identifique geometricamente o gráfico da função real de duas variáveis reais, definida pela expressão analítica $f(x, y) = x^2 + y^2$.

As curvas de nível do gráfico de f são circunferências centradas na origem e raio $\sqrt{z_j}$ com $z_j \in \mathbb{R}^+$. Além disso, a intersecção do gráfico de f com o plano OXZ é a parábola $z = x^2$, e com o plano OYZ é a parábola $z = y^2$. Sendo assim, o gráfico de f é um *paraboloide de revolução*.

Vemos assim que é fundamental a introdução de conceitos geométricos como o de curvas e superfícies, bem como o de vectores, rectas e planos para o estudo de funções reais ou vectoriais de várias variáveis reais.

1.3. Limite e continuidade de funções vectoriais.

Definição 1.3. Sejam f uma função real definida num aberto D de \mathbb{R}^n e $\mathbf{a} \in D$. Dizemos que b é o *limite de f quando \mathbf{x} tende para \mathbf{a}* , e denotamo-lo por $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = b$, se para toda a sucessão não constante de pontos de D , (\mathbf{x}^n) que converge para \mathbf{a} se tem que a sucessão $(f(\mathbf{x}^n))$ converge para b .

Observação (Limite de funções vectoriais). Na definição de limite de funções reais definidas em abertos $D \subset \mathbb{R}^n$, dizemos que $\mathbf{x} = (x_1, \dots, x_n)$ tende para $\mathbf{a} = (a_1, \dots, a_n)$, quando e só quando $x_j \rightarrow a_j$ com $j = 1, \dots, n$. Da mesma forma dizemos que existe $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$ onde $\mathbf{b} = (b_1, \dots, b_m)$ e $\mathbf{f} = (f_1, \dots, f_m)$ é uma função vectorial definida no aberto $D \subset \mathbb{R}^n$ e $\mathbf{a} \in D$, se e somente se existirem os limites $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f_j(\mathbf{x}) = b_j$ para $j = 1, \dots, m$.

Teorema (Cauchy). *Sejam \mathbf{f} uma função vectorial definida num aberto D de \mathbb{R}^n e $\mathbf{a} \in D$. Uma condição necessária e suficiente para a existência do limite de \mathbf{f} quando \mathbf{x} tende para \mathbf{a} é dada por*

$$\forall \epsilon > 0 \exists \delta > 0 : |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| < \epsilon \Leftrightarrow \mathbf{x}, \mathbf{y} \in D, \|\mathbf{x} - \mathbf{a}\| < \delta \text{ e } \|\mathbf{y} - \mathbf{a}\| < \delta.$$

Teorema (Sandwich). *Sejam f, f_+, f_- três funções reais definidas numa vizinhança do ponto \mathbf{a} , tais que $f_- \leq f \leq f_+$. Se existem os limites $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f_-(\mathbf{x})$ e $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f_+(\mathbf{x})$ e forem iguais, então existe o limite $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x})$ e coincide com os limites anteriores.*

Observação . Este teorema tem uma versão vectorial, i.e. supondo que a função $\mathbf{f}: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ com D aberto, tem expressão analítica $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$, onde as funções f_j são reais definidas em D para $j = 1, \dots, m$. Se existirem funções f_j^+ e f_j^- tais que $f_j^- \leq f_j \leq f_j^+$ numa vizinhança de um ponto $\mathbf{a} \in D$, e os limites $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f_j^-(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} f_j^+(\mathbf{x}) = \ell_j$ para $j = 1, \dots, m$, então existe o limite $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = (\ell_1, \dots, \ell_m)$.

Definição 1.4. Sejam \mathbf{f} uma função vectorial definida num aberto D de \mathbb{R}^n e $\mathbf{a} \in D$. Dizemos que \mathbf{f} é contínua em \mathbf{a} se para toda a sucessão (\mathbf{x}^n) que converge para \mathbf{a} se tem que a sucessão $(\mathbf{f}(\mathbf{x}^n))$ converge para $\mathbf{f}(\mathbf{a})$.

Teorema (Composição de funções). Sejam \mathbf{f} e \mathbf{g} duas funções vectoriais definidas por

$$\mathbf{f}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{g}: V \subset \mathbb{R}^m \rightarrow \mathbb{R}^p.$$

com U e V abertos em \mathbb{R}^n e \mathbb{R}^m , respectivamente. Seja $\mathbf{a} \in U$ tal que $\mathbf{f}(\mathbf{a}) \in V$. Se \mathbf{f} é contínua em \mathbf{a} e \mathbf{g} contínua em $\mathbf{f}(\mathbf{a})$, então a função \mathbf{h} definida de U para \mathbb{R}^p por $\mathbf{h}(\mathbf{x}) = \mathbf{g} \circ \mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$, é contínua em \mathbf{a} .

Teorema (Limite por curvas). Sejam \mathbf{f} uma função vectorial definida numa vizinhança do ponto $\mathbf{a} \in \mathbb{R}^n$, e \mathbf{h} uma função vectorial definida num intervalo aberto $I \subset \mathbb{R}$, e contínua em $t_0 \in I$ com $\mathbf{h}(t_0) = \mathbf{a}$, que designaremos por curva contínua. Existe o limite $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x})$ e é igual a ℓ , quando e só quando existir e for igual a ℓ o limite $\lim_{t \rightarrow t_0} \mathbf{f} \circ \mathbf{h}(t)$, sobre todas as curvas contínuas \mathbf{h} que passam por \mathbf{a} .

Exemplo 1.6 (Limite direccional). Sejam \mathbf{f} uma função vectorial definida num aberto D de \mathbb{R}^n , $\mathbf{a} \in D$ e \mathbf{w} um vector de norma um. Suponhamos que uma vizinhança de $\mathbf{a} \in D$, $U \subset D$, contém o segmento de recta de equação $\mathbf{x} = \mathbf{a} + t\mathbf{w}$ com $t \in]-t_0, t_0[$. Considere a função vectorial φ , definida em $] -t_0, t_0[$ por $\varphi(t) = \mathbf{f} \circ \mathbf{h}(t)$, onde \mathbf{h} é a função vectorial definida em $] -t_0, t_0[$ por $\mathbf{h}(t) = \mathbf{a} + t\mathbf{w}$. O limite $\lim_{t \rightarrow 0} \varphi(t)$ é um caso particular da definição anterior e será designado por *limite de \mathbf{f} em \mathbf{a} segundo o vector \mathbf{w}* .

Observação . Ainda que não exista um algoritmo que nos permita decidir sobre a existência de limite, ou sobre a continuidade, de uma função num ponto, podemos começar por calcular os limites direccionais nesse ponto. Se o resultado alcançado depender da direcção ou do sentido do vector dado, então não existe limite. Caso contrário, teremos de procurar curvas mais complexas que os segmentos de recta dos limites direccionais ou tentar aplicar o teorema da sandwich.

Definição 1.5. Seja \mathbf{f} uma função vectorial definida num aberto D de \mathbb{R}^n . Dizemos que \mathbf{f} é *uniformemente contínua* em $A \subset D$ se

$$\forall \epsilon > 0 \exists \delta > 0 : \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| < \epsilon \iff \mathbf{x}, \mathbf{y} \in A, \|\mathbf{x} - \mathbf{y}\| < \delta.$$

Exemplo 1.7. Seja f a função real de variável real definida em $]0, 1[$, com expressão analítica $f(x) = \sin(1/x)$. Sabe-se que f assim definida é contínua. Considerando a sucessão de

pontos $x_n = 1/(n\pi + \pi/2)$, $n = 1, 2, \dots$ temos

$$f(x_{2n}) = 1, \quad f(x_{2n+1}) = -1 \quad \text{e} \quad x_{2n+1} - x_{2n} \rightarrow 0,$$

no entanto $|f(x_{2n+1}) - f(x_{2n})| = 2$, $n = 1, 2, \dots$, pelo que f não é uniformemente contínua.

Definição 1.6. Dizemos que $A \subset \mathbb{R}^n$ é um *compacto*, se toda a sucessão $(\mathbf{x}_n) \subset A$ tem uma subsucessão convergente com limite em A .

Teorema (Continuidade uniforme). *Seja f uma função real, definida e contínua sobre um conjunto compacto A . Então, são as seguintes condições são equivalentes:*

- (a) *O conjunto $\{y \in \mathbb{R} : f(\mathbf{x}) = y, \mathbf{x} \in A\}$ é limitado;*
- (b) $\exists \mathbf{x}_1, \mathbf{x}_2 \in A : \sup_{\mathbf{x} \in A} f(\mathbf{x}) = f(\mathbf{x}_1)$ e $\inf_{\mathbf{x} \in A} f(\mathbf{x}) = f(\mathbf{x}_2)$;
- (c) *f é uniformemente contínua.*

Observação . A condição (c) do teorema anterior é conhecido na literatura como teorema de Weierstrass.

2. DIFERENCIABILIDADE DE FUNÇÕES VECTORIAIS

2.1. Derivada parcial e gradiente. Interpretação geométrica. Considere-se a função de expressão analítica $z = f(x, y)$, definida num aberto $D \in \mathbb{R}^2$. Para y fixo, temos uma função real de variável real $z = f_y(x)$. À derivada desta função designamos por *derivada parcial de f relativamente a x* , e denotamo-la por $\frac{\partial f}{\partial x} = f'_x$. Note-se que a derivada parcial é ainda uma função real de duas variáveis reais.

Para cada y fixo, o gráfico de $z = f_y(x)$ representa uma curva que é a intersecção do gráfico de $z = f(x, y)$ com o plano paralelo a OXZ e distando y unidades dele. Desta forma, a *derivada parcial de f relativamente a x* , representa geometricamente o declive que a tangente à curva de equação $z = f_y(x)$ faz com a direcção positiva de OX .

Definição 2.1. Seja f uma função real de n variáveis reais (x_1, \dots, x_n) definida num aberto $D \subset \mathbb{R}^n$. Definimos *derivada parcial de f relativamente à variável x_j* , e denotamo-la por $\frac{\partial f}{\partial x_j}$, a derivada da função real de variável real de expressão analítica

$$z = f_{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n}(x_j).$$

Vimos já que a derivada parcial de f relativamente a x_j nos dá a variação de f na direcção do eixo OX_j , para cada $j = 1, 2, \dots, n$. Faz então sentido definir derivada de f segundo uma qualquer direcção.

Definição 2.2 (Derivada direccionada). Sejam f uma função real definida num aberto D de \mathbb{R}^n , $\mathbf{a} \in D$ e \mathbf{w} um vector de norma um. Suponhamos que uma vizinhança de $\mathbf{a} \in D$, $U \subset D$, contém o segmento de recta de equação $\mathbf{x} = \mathbf{a} + t\mathbf{w}$ com $t \in] -t_0, t_0[$. Considere a função real φ , definida em $] -t_0, t_0[$ por $\varphi(t) = f \circ \mathbf{h}(t)$, onde \mathbf{h} é a função vectorial definida em

$] - t_0, t_0[$ por $\mathbf{h}(t) = \mathbf{a} + t\mathbf{w}$. À derivada $\varphi'(0)$ designamos por *derivada de f em \mathbf{a} segundo o vector \mathbf{w}* , e denotamo-la por $\varphi'(0) = \frac{\partial f}{\partial \mathbf{w}}(\mathbf{a})$.

2.2. Diferenciabilidade. Para estudar, como primeira aproximação, uma função real na vizinhança de algum dos seus pontos procuramos, caso exista, a função linear que melhor aproxima o seu incremento. Definido este objectivo analisemos um resultado que nos guiará no estudo que agora iniciamos.

Teorema (Incrementos finitos). *Seja f uma função real contínua definida num conjunto aberto D de \mathbb{R}^n . Se existem as derivadas parciais de primeira ordem de f em qualquer ponto de D , então dados $\mathbf{a} \in D$ e $\mathbf{b} \in B_r(\mathbf{a}) \subset D$, existem pontos $\mathbf{t}_1, \dots, \mathbf{t}_n \in D$ e constantes reais $\alpha_1, \dots, \alpha_n$ tais que*

$$f(\mathbf{b}) - f(\mathbf{a}) = \alpha_1 \frac{\partial f}{\partial x_1}(\mathbf{t}_1) + \dots + \alpha_n \frac{\partial f}{\partial x_n}(\mathbf{t}_n).$$

Demonstração. Começemos por relembrar a definição de *conjunto convexo*, i.e. conjunto que contém os segmentos de recta entre cada par de pontos do conjunto. Como exemplo de conjunto convexo temos $B_r(\mathbf{a}) \subset D$.

Considere os $n + 1$ pontos de $B_r(\mathbf{a})$, $\mathbf{s}_j = (b_1, \dots, b_j, a_{j+1}, \dots, a_n)$ com $j = 1, \dots, n - 1$, $\mathbf{a} = (a_1, \dots, a_n)$ e $\mathbf{b} = (b_1, \dots, b_n)$. Note que por definição $\mathbf{s}_0 = \mathbf{a}$ e $\mathbf{s}_n = \mathbf{b}$. Assim,

$$f(\mathbf{b}) - f(\mathbf{a}) = (f(\mathbf{s}_n) - f(\mathbf{s}_{n-1})) + (f(\mathbf{s}_{n-1}) - f(\mathbf{s}_{n-2})) + \dots + (f(\mathbf{s}_1) - f(\mathbf{s}_0)).$$

Definindo as n funções reais de variável real contínuas, φ_j com $j = 1, \dots, n$ por

$$\begin{aligned} \varphi_j: [\min\{a_j, b_j\}, \max\{a_j, b_j\}] &\rightarrow \mathbb{R} \\ t &\mapsto \varphi_j(t) = f(b_1, \dots, b_{j-1}, t, a_{j+1}, \dots, a_n), \end{aligned}$$

temos que a expressão anterior toma a forma

$$f(\mathbf{b}) - f(\mathbf{a}) = \sum_{j=1}^n (\varphi_j(b_j) - \varphi_j(a_j)).$$

Tendo em atenção que $\varphi_j'(t) = \frac{\partial f}{\partial x_j}(\mathbf{t}^j)$ com $\mathbf{t}^j = (b_1, \dots, b_{j-1}, t, a_{j+1}, \dots, a_n)$, $j = 1, \dots, n$, vemos que as funções φ_j são deriváveis no seu intervalo aberto de definição.

Aplicando o teorema do valor médio de Lagrange a cada uma das funções φ_j no intervalo $[\min\{a_j, b_j\}, \max\{a_j, b_j\}]$, temos que existe $l_j \in]\min\{a_j, b_j\}, \max\{a_j, b_j\}[$, $j = 1, \dots, n$ tais que

$$\varphi_j(b_j) - \varphi_j(a_j) = (b_j - a_j)\varphi_j'(l_j), \quad j = 1, \dots, n.$$

Assim,

$$f(\mathbf{b}) - f(\mathbf{a}) = \sum_{j=1}^n (b_j - a_j) \frac{\partial f}{\partial x_j}(\mathbf{l}_j^j),$$

onde no enunciado do teorema temos $\alpha_j = b_j - a_j$, e $\mathbf{t}_j = \mathbf{l}_j^j$, $j = 1, \dots, n$. *c. q. d.*

Estamos em condição de definir o conceito fundamental do Cálculo Diferencial.

Definição 2.3. Dizemos que a função real f definida no aberto $D \subset \mathbb{R}^n$ é *diferenciável em* $\mathbf{a} \in D$ se existem duas funções reais, uma linear que denotamos por $df(\mathbf{a})$ definida em \mathbb{R}^n , e a outra ϵ definida numa vizinhança da origem, $V \subset \mathbb{R}^n$, tais que

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\epsilon(\mathbf{h}), \text{ e } \lim_{\mathbf{h} \rightarrow \mathbf{0}} \epsilon(\mathbf{h}) = 0.$$

Nesta condições, dizemos que $df(\mathbf{a})$ é o *diferencial de f em \mathbf{a}* .

Observação .

- (1) Por definição de função linear, $df(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}$, tem expressão analítica $df(\mathbf{a})(\mathbf{h}) = [\alpha_1 \dots \alpha_n]|\mathbf{h}$, pelo que passaremos a denotar a expressão analítica de $df(\mathbf{a})$ por $df(\mathbf{a})\mathbf{h}$.
- (2) Se f é diferenciável em \mathbf{a} existe uma única diferencial de f em \mathbf{a} .
- (3) Tomando limite quando \mathbf{h} tende para $\mathbf{0}$ em ambos os membros da igualdade anterior vemos que se f é diferenciável em \mathbf{a} é também contínua nesse ponto. O recíproco é falso, em geral. Dizemos assim, que *uma condição necessária de diferenciabilidade de uma função real f em \mathbf{a} é que f seja contínua em \mathbf{a}* .
- (4) Tomando $\mathbf{h} = [h_1 \ 0 \dots 0]$ na definição anterior

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \alpha_1 h_1 + |h_1|\epsilon(h_1, 0, \dots, 0), \text{ com } \lim_{h_1 \rightarrow 0} \epsilon(h_1, 0, \dots, 0) = 0.$$

Pelo que, dividindo ambos os membros por h_1 e tomando o limite quando h_1 tende para zero, obtemos que $\alpha_1 = \frac{\partial f}{\partial x_1}(\mathbf{a})$.

Tomando $\mathbf{h} = [0 \dots 0 \ h_j \ 0 \dots 0]$ e procedendo como no caso anterior obtemos $\alpha_j = \frac{\partial f}{\partial x_j}(\mathbf{a})$, $j = 1, \dots, n$.

- (5) Se a função real de n variáveis reais for diferenciável em \mathbf{a} então

$$\begin{aligned} df(\mathbf{a}) : \mathbb{R}^n &\rightarrow \mathbb{R} \\ \mathbf{h} &\mapsto df(\mathbf{a})\mathbf{h} = \text{grad } f(\mathbf{a})|\mathbf{h}. \end{aligned}$$

- (6) Se f é diferenciável em \mathbf{a} , então a derivada direccional de f em \mathbf{a} segundo um qualquer vector, \mathbf{u} , de norma um, vem dada por $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{a}) = df(\mathbf{a})\mathbf{u}$, ou em notação matricial $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{a}) = \text{grad } f(\mathbf{a})|\mathbf{u}$.

- (7) **Interpretação geométrica do diferencial:** Dizemos que a superfície, S , de equação cartesiana $z = f(x, y)$ tem o plano π como *plano tangente* no ponto $\mathbf{P}_0 = (x_0, y_0, z_0)$ de S , se o ângulo que π faz com uma secante $\mathbf{P}_0\mathbf{P}$ onde $\mathbf{P} = (x, y, z) \in S$, tende para zero quando (x, y) tende para (x_0, y_0) . Assim, f , como função real de duas variáveis reais, é diferenciável em (x_0, y_0) quando e só quando a superfície, S de equação $z = f(x, y)$ tem plano tangente em $(x_0, y_0, z_0) \in S$. Neste caso, o plano π tem equação cartesiana

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

Problema 2.1. Seja $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por

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

Determine as constantes a, b, c para que o valor máximo do módulo da derivada direccional no ponto $(1, 1, 1)$ seja 13, na direcção do vector $(1, 5, 0)$.

Resolução. Como a função f é diferenciável em \mathbb{R}^3 a derivada direccional de f segundo o vector $\mathbf{u} = (1, 5, 0)/\sqrt{26}$, vem dada por $\frac{\partial f}{\partial \mathbf{u}}(1, 1, 1) = (2a + c, a + 2b, b + 2c)|\mathbf{u}$, e é máxima quando o ângulo entre estes dois vectores for zero. Assim,

$$2a + c = \alpha, \quad a + 2b = 5\alpha, \quad b + 2c = 0, \quad \text{i.e.}$$

$a = 2b = -c = \alpha$, e como a $\|\frac{\partial f}{\partial \mathbf{u}}(1, 1, 1)\| = 13$, i.e. $\|\text{grad } f(1, 1, 1)\| = \|(\alpha, 3\alpha, 0)\| = \sqrt{\alpha^2 + 9\alpha^2} = 13$ e portanto, $\alpha = 1/\sqrt{26}$. c.q.d.

Definição 2.4. Dizemos que a função vectorial $\mathbf{f}: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é diferenciável em $\mathbf{a} \in D$ se existem duas funções em \mathbb{R}^m , uma linear definida em \mathbb{R}^n que denotamos por $d\mathbf{f}(\mathbf{a})$, e a outra ϵ definida numa vizinhança da origem, $V \subset \mathbb{R}^n$, tais que

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\mathbf{h} + \|\mathbf{h}\|\epsilon(\mathbf{h}), \quad \text{e} \quad \lim_{\mathbf{h} \rightarrow \mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}.$$

Nestas condições, dizemos que a função $d\mathbf{f}(\mathbf{a})$ é o diferencial de \mathbf{f} em \mathbf{a} .

Observação . Tendo em atenção que \mathbf{f} tem expressão analítica $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ onde as funções f_j são funções reais definidas em $D \subset \mathbb{R}^n$, a condição de diferenciabilidade reescreve-se como

$$f_j(\mathbf{a} + \mathbf{h}) - f_j(\mathbf{a}) = df_j(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\epsilon_j(\mathbf{h}), \quad j = 1, \dots, m,$$

e portanto \mathbf{f} é diferenciável em \mathbf{a} se e somente se as funções reais f_j o são, para todo o $j = 1, \dots, m$. Além disso, a matriz que representa o diferencial de \mathbf{f} em \mathbf{a} vem dada por

$$J_{\mathbf{f}}(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

e designa-se por *jacobiano* de \mathbf{f} em \mathbf{a} .

Definição 2.5. Seja \mathbf{f} uma função vectorial em \mathbb{R}^m , definida num aberto $D \subset \mathbb{R}^n$, por $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$. Dizemos que \mathbf{f} é de classe C^1 num ponto $\mathbf{a} \in D$ se as funções derivadas de primeira ordem de f_j com $j = 1, \dots, m$ forem contínuas em \mathbf{a} . Dizemos que \mathbf{f} é de classe C^1 em D se esta propriedade se tiver em todos os pontos de D .

Teorema (Condição suficiente de diferenciabilidade). Se \mathbf{f} é de classe C^1 em $\mathbf{a} \in D$ onde D é um aberto de \mathbb{R}^n , então \mathbf{f} é diferenciável em \mathbf{a} .

Demonstração. Seguindo o processo descrito no teorema dos incrementos finitos e pela continuidade das funções derivadas parciais temos o resultado desejado. c.q.d.

Problema 2.2. Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \frac{3x^2y - 2x^3}{x^2 + y^4}, \quad (x, y) \neq (0, 0) \quad \text{e} \quad f(0, 0) = 0.$$

- (a) Verifique que f admite em $(0,0)$, derivada direccional segundo qualquer vector unitário.
- (b) Mostre que f não é diferenciável em $(0,0)$. Comente o resultado obtido.

Resolução. (a) Por definição de derivada direccional de f segundo um vector \mathbf{u} unitário, i.e $\mathbf{u} = (\cos \alpha, \sin \alpha)$ temos

$$\frac{\partial f(0,0)}{\partial \mathbf{u}} = \lim_{t \rightarrow 0} \frac{f(t \cos \alpha, t \sin \alpha) - f(0)}{t} = \begin{cases} 0 & , \cos \alpha = 0 \\ -2 \cos \alpha + 3 \sin \alpha & , \cos \alpha \neq 0 \end{cases}.$$

(b) As derivadas parciais de f em $(0,0)$ vêm dadas por $\frac{\partial f}{\partial x}(0,0) = -2$ e $\frac{\partial f}{\partial y}(0,0) = 0$. Assim, se f é diferenciável em $(0,0)$ então a derivada direccional f segundo um vector unitário \mathbf{u} vem dado por $\frac{\partial f(0,0)}{\partial \mathbf{u}} = (-2, 0) | (\cos \alpha, \sin \alpha) = -2 \cos \alpha$. Como este resultado não coincide com o calculado em (a), pela alínea (b) da observação que aparece a seguir à definição 2.3, temos que f não é diferenciável. *c.q.d.*

Teorema (Composição de funções). *Sejam \mathbf{f} e \mathbf{g} duas funções vectoriais definidas por*

$$\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{g} : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^p.$$

com U e V abertos em \mathbb{R}^n e \mathbb{R}^m , respectivamente. Seja $\mathbf{a} \in U$ tal que $\mathbf{f}(\mathbf{a}) \in V$. Se \mathbf{f} é diferenciável em \mathbf{a} e \mathbf{g} diferenciável em $\mathbf{f}(\mathbf{a})$, então a função \mathbf{h} definida de U para \mathbb{R}^p por $\mathbf{h}(\mathbf{x}) = \mathbf{g} \circ \mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$, é diferenciável em \mathbf{a} e

$$J_{\mathbf{h}}(\mathbf{a}) = J_{\mathbf{g}}(\mathbf{f}(\mathbf{a})) \cdot J_{\mathbf{f}}(\mathbf{a}).$$

Problema 2.3. *Seja f a função real definida em \mathbb{R}^2 por*

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

- (a) *Determine o seu domínio de continuidade.*
- (b) *Defina as funções $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ indicando os respectivos domínios de continuidade.*
- (c) *Estude f quanto à diferenciabilidade.*

Resolução. (a) A função f é a composição das funções contínuas, g real de variável real definida em \mathbb{R} com expressão analítica $g(x) = x^2 \sin(1/x)$, $x \neq 0$ e $g(0) = 0$, e a função real, h , definida em \mathbb{R}^2 , com expressão analítica $h(x, y) = \sqrt{x^2 + y^2}$. Assim, $f = g \circ h$ em \mathbb{R}^2 , e portanto é uma função contínua.

(b) As funções derivadas parciais estão definidas em \mathbb{R}^2 por

$$\begin{aligned} \frac{\partial f}{\partial x} &= \begin{cases} 2x \sin(1/\sqrt{x^2 + y^2}) - x/\sqrt{x^2 + y^2} \cos(1/\sqrt{x^2 + y^2}) & , (x, y) \neq (0, 0) \\ 0 & , (x, y) = (0, 0) \end{cases} \\ \frac{\partial f}{\partial y} &= \begin{cases} 2y \sin(1/\sqrt{x^2 + y^2}) - y/\sqrt{x^2 + y^2} \cos(1/\sqrt{x^2 + y^2}) & , (x, y) \neq (0, 0) \\ 0 & , (x, y) = (0, 0) \end{cases}. \end{aligned}$$

Note-se que o cálculo de $\frac{\partial f}{\partial x}(0,0)$ e de $\frac{\partial f}{\partial y}(0,0)$ tem de ser efectuado por definição, i.e.

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h \rightarrow 0} \frac{f(h,0) - f(0,0)}{h} = 0 = \lim_{k \rightarrow 0} \frac{f(0,k) - f(0,0)}{k} = \frac{\partial f}{\partial y}(0,0).$$

Pode ver-se que as funções que acabámos de definir são contínuas em $\mathbb{R}^2 \setminus \{(0,0)\}$.

(c) Como a função h é diferenciável em $\mathbb{R}^2 \setminus \{(0,0)\}$ e g diferenciável em $\mathbb{R} \setminus \{0\}$, f é diferenciável em $\mathbb{R}^2 \setminus \{(0,0)\}$.

A análise da diferenciabilidade de f no ponto $(0,0)$ vai ser feita por definição:

$$f(h_1, h_2) - f(0,0) = (h_1^2 + h_2^2) \sin(1/\sqrt{h_1^2 + h_2^2}) = \sqrt{h_1^2 + h_2^2} \epsilon(h_1, h_2),$$

onde $\epsilon(h_1, h_2) = \sqrt{h_1^2 + h_2^2} \sin(1/\sqrt{h_1^2 + h_2^2})$ é tal que $\lim_{(h_1, h_2) \rightarrow (0,0)} \epsilon(h_1, h_2) = 0$, logo f é diferenciável em $(0,0)$. c.q.d.

Observação . Acabámos de apresentar um exemplo de uma função diferenciável num ponto que não é de classe C^1 nesse ponto.

Problema 2.4. *Averigúe se a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por*

$$f(x, y) = \begin{cases} (x^2 + x^2 \sin(1/x), y) & , \quad x \neq 0 \\ (0, y) & , \quad x = 0 \end{cases}$$

é diferenciável em \mathbb{R}^2 .

Resolução. Aplique as ideias apresentadas na resolução do problema anterior. c.q.d.

2.3. Polinómio de Taylor. É sabido que para funções reais de variável real, f , uma *função analítica em $I \subset \mathbb{R}$* é aquela que admite uma representação em *série de Taylor*, uniformemente convergente numa vizinhança de cada ponto $x_0 \in I$, i.e.

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n, \quad \{x \in \mathbb{R} : |x - x_0| < r\} \subset I.$$

Vejamos como estender este tipo de representações para funções reais de n variáveis reais. Para tal considere-se uma função real f analítica definida num aberto $D \subset \mathbb{R}^n$. Defina-se para cada $\mathbf{x}^0 \in D$ a função auxiliar ϕ como a composição de f com a função vectorial \mathbf{X} definida em $[0, 1]$ por $\mathbf{X}(t) = \mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0)$ com $\mathbf{X}([0, 1]) \subset D$, i.e.

$$\begin{aligned} \phi: [0, 1] &\rightarrow \mathbb{R} \\ t &\mapsto \phi(t) = f(\mathbf{X}(t)). \end{aligned}$$

Assim, como ϕ é uma função real de variável real analítica, pois é a composição de funções analíticas f e \mathbf{X} , temos que

$$f(\mathbf{X}(t)) = \sum_{n=0}^{\infty} \frac{\phi^{(n)}(0)}{n!} t^n, \quad \{t \in \mathbb{R} : |t| < r\},$$

onde $\phi(0) = f(\mathbf{x}^0)$, $\phi'(0) = \text{grad } f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)^t = d f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)$,

$$\phi''(0) = (\mathbf{x} - \mathbf{x}^0) \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}_{\mathbf{x}^0} (\mathbf{x} - \mathbf{x}^0)^t = d^2 f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0),$$

e mais geralmente,

$$\phi^{(n)}(0) = d^n f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)^n, \quad n = 0, 1, \dots$$

Tomando $t = 1$ e denotando $\mathbf{h} = \mathbf{x} - \mathbf{x}^0$ temos que a *série de Taylor de f em \mathbf{x}^0* vem dada por

$$f(\mathbf{x}) = f(\mathbf{x}^0) + d f(\mathbf{x}^0)\mathbf{h} + \frac{d^2 f(\mathbf{x}^0)\mathbf{h}}{2!} + \cdots + \frac{d^n f(\mathbf{x}^0)\mathbf{h}}{n!} + \cdots,$$

e o *polinómio de Taylor de ordem n de f em \mathbf{x}^0* é dado por

$$T_n(f; \mathbf{x}^0) = f(\mathbf{x}^0) + d f(\mathbf{x}^0)\mathbf{h} + \frac{d^2 f(\mathbf{x}^0)\mathbf{h}}{2!} + \cdots + \frac{d^n f(\mathbf{x}^0)\mathbf{h}}{n!}.$$

Definição 2.6. À matriz da forma bilinear $d^2 f(\mathbf{x}^0)$, i.e.

$$\mathcal{H}_f(\mathbf{x}^0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}_{\mathbf{x}^0}$$

designamos por *hessiana de f em \mathbf{x}^0* .

3. APLICAÇÕES DA DIFERENCIABILIDADE

3.1. Teoremas da função implícita e da função inversa. A partir de funções reais de duas variáveis reais podemos definir funções reais de variável real. De facto, dada uma função real diferenciável de duas variáveis reais, f , considere-se a equação $f(x, y) = 0$. Em geral esta equação define em \mathbb{R}^2 um conjunto que designaremos por *curva*. Estes conjuntos representam frequentemente gráficos de funções reais de variável real de expressão analítica $y = \phi(x)$ ou $x = \psi(y)$. Neste caso, dizemos que estas *funções estão definidas implicitamente pela equação $F(x, y) = 0$* .

Exemplo 3.1. A equação $x^2 + y^2 - r^2 = 0$ define, por exemplo, as funções de expressão analítica $f_1(x) = \sqrt{r^2 - x^2}$, $x \in]-r, r[$ e $f_2(y) = -\sqrt{r^2 - y^2}$, $y \in]-r, r[$.

Note-se que uma equação pode não definir nenhuma função real de variável real, como é o caso de $x^2 + y^2 + 1 = 0$.

Definição 3.1. Sejam $U \subset \mathbb{R}^n \times \mathbb{R}$ uma vizinhança de um ponto $(a_1, \dots, a_n; b)$ e F uma função real definida em U . Dizemos que a *equação $F(\mathbf{x}; y) = 0$ define implicitamente y como função de \mathbf{x} , em U* , se existir uma vizinhança, $U_{\mathbf{a}}$ e uma única função real contínua, f , tal que

$$b = f(\mathbf{a}), \quad F(\mathbf{x}; f(\mathbf{x})) = 0, \quad \mathbf{x} \in U_{\mathbf{a}}.$$

Teorema (Existência da função implícita). *Considere a função real, F , definida e contínua numa vizinhança do ponto $(\mathbf{a}; b)$. Se $F(\mathbf{a}; b) = 0$ e F é estritamente monótona considerada como função da variável y , então a equação $F(\mathbf{x}; y) = 0$ define y como função implícita de \mathbf{x} numa vizinhança de $(\mathbf{a}; b)$.*

Observação . Se a condição de monotonia da função F considerada como função de y , for substituída por, a função derivada parcial de F em ordem a y é contínua numa vizinhança de $(\mathbf{a}; b)$, $\frac{\partial F}{\partial y}(\mathbf{a}; b) \neq 0$, e existirem as derivadas parciais de F relativamente a x_j em $(\mathbf{a}; b)$, então existem as derivadas parciais de f , numa vizinhança de \mathbf{a} , e são dadas por

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) = - \left[\frac{\partial F}{\partial y}(\mathbf{x}; f(\mathbf{x})) \right]^{-1} \frac{\partial F}{\partial x_j}(\mathbf{x}; f(\mathbf{x})), \quad \mathbf{x} \in U_{\mathbf{a}}, \quad j = 1, \dots, n.$$

Teorema (Diferenciabilidade da função implícita). *Seja a função real, F , definida e contínua numa vizinhança do ponto $(\mathbf{a}; b)$. Se $F(\mathbf{a}; b) = 0$, $\frac{\partial F}{\partial y}(\mathbf{a}; b) \neq 0$ e F é diferenciável em $(\mathbf{a}; b)$, então a equação $F(\mathbf{x}; y) = 0$ define y como função implícita de \mathbf{x} , diferenciável, numa vizinhança de $(\mathbf{a}; b)$. Além disso, as derivadas parciais da função assim definida, $y = f(x_1, \dots, x_n)$, são tais que*

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) = - \left[\frac{\partial F}{\partial y}(\mathbf{x}; f(\mathbf{x})) \right]^{-1} \frac{\partial F}{\partial x_j}(\mathbf{x}; f(\mathbf{x})), \quad j = 1, \dots, n.$$

Observação . Desta forma se vê que as condições de regularidade de F passam directamente para a função definida implicitamente pela equação $F(\mathbf{x}; y) = 0$, i.e. se F é de classe C^k numa vizinhança de $(\mathbf{a}; b)$ então existe uma vizinhança de \mathbf{a} onde f é de classe C^k .

Estes teoremas vão ser agora estendidos ao caso de $m \in \mathbb{N}$ equações.

Teorema (Funções implícitas). *Sejam $U \subset \mathbb{R}^n \times \mathbb{R}^m$ uma vizinhança de um ponto $(\mathbf{a}; \mathbf{b}) = (a_1, \dots, a_n; b_1, \dots, b_m)$ e \mathbf{F} uma função vectorial definida de U para \mathbb{R}^m , por $\mathbf{F}(\mathbf{x}; \mathbf{y}) = (F_1(\mathbf{x}; \mathbf{y}), \dots, F_m(\mathbf{x}; \mathbf{y}))$, onde as funções F_j são reais, definidas em U . Se \mathbf{F} é de classe C^1 em U , $\mathbf{F}(\mathbf{a}; \mathbf{b}) = \mathbf{0}$ e o determinante do jacobiano de \mathbf{F} considerado como função de \mathbf{y} é não nulo, i.e.*

$$\left| \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{a}; \mathbf{b}) \right| = \left| \frac{\partial (F_1, \dots, F_m)}{\partial (y_1, \dots, y_m)}(\mathbf{a}; \mathbf{b}) \right| = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{vmatrix}_{(\mathbf{a}; \mathbf{b})} \neq 0$$

então a equação $\mathbf{F}(\mathbf{x}; \mathbf{b}) = \mathbf{0}$ define implicitamente \mathbf{y} como função de \mathbf{x} , de classe C^1 numa vizinhança de $(\mathbf{a}; \mathbf{b})$, i.e existem uma vizinhança de \mathbf{a} , $U_{\mathbf{a}}$, e uma única função de classe C^1 , \mathbf{f} , em \mathbb{R}^m , definida em $U_{\mathbf{a}}$ por $\mathbf{y} = \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ tal que $\mathbf{F}(\mathbf{x}; \mathbf{f}(\mathbf{x})) = \mathbf{0}$, $\mathbf{x} \in U_{\mathbf{a}}$. Além disso, a matriz que representa o diferencial de \mathbf{f} em $U_{\mathbf{a}}$, $J_{\mathbf{f}} = \frac{\partial (f_1, \dots, f_m)}{\partial (x_1, \dots, x_n)}$ vem dada por

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{bmatrix}_{(\mathbf{x}; \mathbf{f}(\mathbf{x}))}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_n} \end{bmatrix},$$

ou em notação funcional $J_{\mathbf{f}} = - \left[\frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} \right]^{-1} \frac{\partial(F_1, \dots, F_m)}{\partial(x_1, \dots, x_n)}$.

Definição 3.2. Dizemos que a função \mathbf{f} definida do conjunto C para D , é *bijectiva*, se cada elemento de D é a imagem de um só elemento de C . Neste caso definimos função inversa de \mathbf{f} , e designamo-la por \mathbf{f}^{-1} à função de D para C , que a cada elemento $\mathbf{d} \in D$ corresponde um único $\mathbf{c} \in C$ tal que $\mathbf{f}(\mathbf{c}) = \mathbf{d}$, e portanto $\mathbf{f}^{-1}(\mathbf{d}) = \mathbf{c}$ é equivalente a $\mathbf{f}(\mathbf{c}) = \mathbf{d}$.

Vamos analisar a existência de inversa local de uma função vectorial, i.e. a existência de inversa na vizinhança de um ponto determinado de C , considerado aberto de \mathbb{R}^n .

Teorema (Função inversa). *Sejam \mathbf{f} uma função de $V \subset \mathbb{R}^n$ definida e de classe C^1 num aberto $U \subset \mathbb{R}^n$, e $\mathbf{a} \in U$. Se o jacobiano de \mathbf{f} em \mathbf{a} , $J_{\mathbf{f}}(\mathbf{a})$, for invertível, i.e. $|J_{\mathbf{f}}(\mathbf{a})| \neq 0$, então f admite inversa local em \mathbf{a} , i.e. existem vizinhanças, $U_{\mathbf{a}} \subset U$ de \mathbf{a} e $V_{\mathbf{b}} \subset V$ de $\mathbf{b} = \mathbf{f}(\mathbf{a})$ tais que \mathbf{f} considerada como uma função de $U_{\mathbf{a}}$ em $V_{\mathbf{b}}$ é bijectiva. Além disso, a função \mathbf{f}^{-1} inversa de f é de classe C^1 em $U_{\mathbf{a}}$ e tem-se*

$$J_{\mathbf{f}^{-1}}(\mathbf{f}(\mathbf{x})) = [J_{\mathbf{f}}(\mathbf{x})]^{-1}, \quad J_{\mathbf{f}}(\mathbf{x}) = [J_{\mathbf{f}^{-1}}(\mathbf{f}(\mathbf{x}))]^{-1}.$$

Problema 3.1. *Considere o sistema de equações*

$$e^{x-u} + yv + 3 = 0, \quad e^{y+v} - xu = 0.$$

- Mostre que o sistema define, nas condições do teorema das funções implícitas, (u, v) como função de (x, y) numa vizinhança do ponto $(x_0, y_0, u_0, v_0) = (1, 2, 1, -2)$.*
- Calcule a matriz Jacobiana (ou o diferencial) da função, \mathbf{g} , definida na alínea anterior no ponto $(1, 2)$.*
- Justifique que \mathbf{g} é localmente invertível e indique sem calcular $|J_{\mathbf{g}^{-1}}(1, -2)|$.*

Resolução. (a) Considere as funções reais f_1, f_2 definidas em \mathbb{R}^4 por $f_1(x, y, u, v) = e^{x-u} + yv + 3$, $f_2(x, y, u, v) = e^{y+v} - xu$, que são funções diferenciáveis em \mathbb{R}^4 . Além disso,

$$f_1(1, 2, 1, -2) = 0 = f_2(1, 2, 1, -2), \quad \left| \frac{\partial(f_1, f_2)}{\partial(x, y)}(1, 2, 1, -2) \right| = -1 \neq 0,$$

como queríamos demonstrar.

(b) Aplicando o teorema da função implícita temos

$$J_{\mathbf{g}}(x, y) = - \begin{bmatrix} e^{x-u} & v \\ -u & e^{y+v} \end{bmatrix}_{(x, y, \mathbf{g}(x, y))}^{-1} \begin{bmatrix} -e^{x-u} & y \\ -x & e^{y+v} \end{bmatrix}_{(x, y, \mathbf{g}(x, y))},$$

$$\text{e portanto } J_{\mathbf{g}}(1, 2) = \begin{bmatrix} -3 & 4 \\ -2 & 3 \end{bmatrix}.$$

$$(c) |J_{\mathbf{g}^{-1}}(1, -2)| = 1/|J_{\mathbf{g}}(1, 2)| = -1.$$

c. q. d.

3.2. Dependência funcional.

Definição 3.3. Sejam $D \in \mathbb{R}^n$ um aberto e $f_j, j = 1, \dots, m$ funções reais definidas em D de classe C^1 em D . Dizemos que f_1, \dots, f_m são *funcionalmente dependentes em D* , se existir uma função real $F \not\equiv 0$ definida em \mathbb{R}^m tal que

$$F(f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) = 0, \quad \mathbf{x} \in D.$$

Dizemos que f_1, \dots, f_m são *funcionalmente independentes em D* , se não forem funcionalmente dependentes.

Observação (Linearmente independente). Note-se que se um conjunto de funções é funcionalmente independente, então também é linearmente independente. No entanto o recíproco não se tem. De facto, as funções $\sin : \mathbb{R} \rightarrow \mathbb{R}$ e $\cos : \mathbb{R} \rightarrow \mathbb{R}$ são linearmente independentes, mas verificam $\sin^2 x + \cos^2 x - 1 = 0, x \in \mathbb{R}$, pelo que existe uma função $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ não constantemente nula, tal que $F(\sin x, \cos x) = 0, x \in \mathbb{R}$. Basta tomar F com expressão analítica $F(x, y) = x^2 + y^2 - 1$. Logo \sin e \cos são funcionalmente dependentes.

Teorema (Condição necessária). Sejam $D \in \mathbb{R}^n$ um aberto e $f_j, j = 1, \dots, m$ funções reais definidas em D de classe C^1 em D , com $m \leq n$. Se as funções f_j com $j = 1, \dots, m$ são funcionalmente dependentes em D , então o determinante do jacobiano das funções f_1, \dots, f_m relativamente a m das n variáveis x_1, \dots, x_n é nulo em D , i.e.

$$\left| \frac{\partial(f_1, \dots, f_m)}{\partial(x_{j_1}, \dots, x_{j_m})}(\mathbf{x}) \right| = 0, \quad \mathbf{x} \in D,$$

onde os $j_k \in \{1, \dots, n\}$ e são todos diferentes.

Definição 3.4. Seja A uma matriz de dimensão $m \times n$. Dizemos que A tem *característica* $r \leq \min\{n, m\}$ se todos os determinantes de ordem $r + 1$ que se podem formar com linhas e colunas de A , i.e. os *menores de ordem $r + 1$* , são nulos e existe um menor de ordem r não nulo.

Teorema (Condição suficiente). Sejam $D \in \mathbb{R}^n$ um aberto e $f_j, j = 1, \dots, m$ funções reais definidas em D de classe C^1 em D , com $m \leq n$. Se a característica do jacobiano da função vectorial, \mathbf{G} , definida em D por $\mathbf{G}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ é $r \leq m$, i.e. $\text{car } J_{\mathbf{G}} = r$, então para cada $\mathbf{a} \in D$ existe uma vizinhança $U_{\mathbf{a}} \subset D$ tal que r das funções f_1, \dots, f_m são funcionalmente independentes e as $m - r$ restantes dependem funcionalmente das primeiras.

3.3. Curvas e Superfícies diferenciáveis. Definimos já o significado de curva contínua em \mathbb{R}^n (cf. teorema limite por curvas). Vamos somente acrescentar que as *curvas* que considerarmos são de classe C^1 e *simples*, i.e. sendo \mathbf{h} uma curva em \mathbb{R}^n definida em $I \subset \mathbb{R}$ por $\mathbf{h}(t) = (h_1(t), \dots, h_n(t))$ onde as funções h_j são todas contínuas em I , $\|\mathbf{h}(t_1) - \mathbf{h}(t_2)\| \neq 0, t_1 \neq t_2 \in I$ ou admitem um número finito de pontos que não verifiquem esta condição que são designados por *pontos múltiplos*.

O vector $\mathbf{t} = \mathbf{h}'(t_0), t_0 \in I$ de coordenadas $(h'_1(t_0), \dots, h'_n(t_0))$ é designado por *vector tangente à curva \mathbf{h} em t_0* .

Definição 3.5. Sejam f_1, \dots, f_n funções reais de classe C^1 definidas num aberto D de \mathbb{R}^2 . Definimos *superfície diferenciável* $S \subset \mathbb{R}^n$ ao conjunto dos pontos \mathbf{x} de coordenadas

$$x_1 = f_1(s, t), \quad \dots, \quad x_n = f_n(s, t), \quad (s, t) \in D,$$

para os quais a característica do jacobiano das funções f_j é igual a dois, i.e.

$$\text{car} \frac{\partial(f_1, \dots, f_n)}{\partial(s, t)} = 2.$$

Observação. Nas condições anteriores sabemos pelo teorema da condição suficiente da secção 3.2, que existe numa vizinhança, $U_{\mathbf{a}} \subset S$, de cada ponto $\mathbf{a} \in S$, uma função F real de classe $C^1(U_{\mathbf{a}})$ tal que a *equação cartesiana* da restrição de S a $U_{\mathbf{a}}$ é dada $F(x_1, \dots, x_n) = 0$ com $\|\text{grad } F\| \neq 0$.

Esta observação é fundamental, pois é um exercício simples demonstrar que um *vector normal a uma superfície* de equação cartesiana $F(x_1, \dots, x_n) = 0$ num ponto \mathbf{x}^0 é $\mathbf{n} = \text{grad } F(\mathbf{x}^0)$. Na verdade basta tomar uma curva de classe C^1 , $C \subset S$ que contenha \mathbf{x}^0 , de equação $x_j = g_j(t)$, $t \in I$, $j = 1, \dots, n$, i.e. $\mathbf{x}^0 = (g_1(t_0), \dots, g_n(t_0))$, para algum $t_0 \in I$. Então, $F(g_1(t), \dots, g_n(t)) = 0$, $t \in I$, e diferenciando obtemos $\text{grad } F(g_1(t_0), \dots, g_n(t_0)) | \mathbf{t} = 0$, onde \mathbf{t} é o vector tangente à curva C em \mathbf{x}^0 , logo um vector normal à superfície S em \mathbf{x}^0 é o vector $\text{grad } F(\mathbf{x}^0)$.

De forma geral temos que para superfícies em \mathbb{R}^3 o vector \mathbf{n} normal a S em $\mathbf{x}^0 \in S$ é dado por

$$\mathbf{n} = \left(\left| \frac{\partial(f_2, f_3)}{\partial(s, t)} \right|, - \left| \frac{\partial(f_1, f_3)}{\partial(s, t)} \right|, \left| \frac{\partial(f_1, f_2)}{\partial(s, t)} \right| \right).$$

Problema 3.2. Um vector normal à superfície de equação $x = u \cos v$, $y = u \sin v$, $z = v$, $(u, v) \in \mathbb{R}^2$ no ponto de coordenadas $(\sqrt{2}/2, \sqrt{2}/2, \pi/4)$ é $(\sqrt{2}/2, -\sqrt{2}/2, 1)$.

Resolução. Vimos já que os vectores normais a uma superfície são múltiplos escalares do vector \mathbf{n} definido anteriormente, considerando que

$$f_1(u, v) = u \cos v, \quad f_2(u, v) = u \sin v, \quad f_3(u, v) = v.$$

Assim, em cada ponto da superfície $\mathbf{n} = (\sin v, -\cos v, u)$. O ponto dado é atingido quando $(u, v) = (1, \pi/4)$, assim $\mathbf{n} = (\sqrt{2}/2, -\sqrt{2}/2, 1)$. c.q.d.

Problema 3.3. Determine uma equação do plano tangente e da recta normal à superfície de equações paramétricas

$$x = s + t, \quad y = s - t, \quad z = s^2 - t^2,$$

em $(s, t) = (1, 0)$ e $(s, t) = (1, 1)$.

Resolução. As equações do plano tangente e da normal à superfície no ponto que corresponde aos parâmetros $(s, t) = (1, 0)$, i.e. o ponto de coordenadas $\mathbf{P}_0 = (1, 1, 1)$ são dadas em termos do vector \mathbf{n} normal à superfície em \mathbf{P}_0 por

$$(x - 1, y - 1, z - 1) | \mathbf{n} = 0, \quad (x - 1, y - 1, z - 1) = t \mathbf{n}, \quad t \in \mathbb{R},$$

respectivamente.

Das considerações efectuadas aquando da resolução do problema anterior temos que o vector \mathbf{n} é um múltiplo escalar de $(1, 1, -1)$. *c.q.d.*

Problema 3.4. *Determine as equações da recta tangente e do plano normal à curva C de equação cartesiana*

$$2x^2 + 3y^2 + z^2 = 47, \quad x^2 + 2y^2 = z,$$

no ponto $\mathbf{P}_0 = (-2, 1, 6)$.

Resolução. Temos de calcular um vector tangente à curva C no ponto \mathbf{P}_0 , que denotaremos por \mathbf{t} . Começemos por ver que o ponto \mathbf{P}_0 pertence à curva. De facto,

$$2(-2)^2 + 3(1)^2 + (6)^2 = 47, \quad (-2)^2 + 2(1)^2 = 6.$$

A equações pedidas viriam dadas por:

$$\text{Recta tangente: } (x + 2, y - 1, z - 6) = \alpha \mathbf{t}, \quad \alpha \in \mathbb{R}.$$

$$\text{Plano normal: } (x + 2, y - 1, z - 6) \cdot \mathbf{t} = 0.$$

Determinação do vector tangente, \mathbf{t} .

Primeiro método: Suponha que a curva C está parametrizada por $x = X(s)$, $y = Y(s)$, $z = Z(s)$, $s \in I \subset \mathbb{R}$, e que $\mathbf{P}_0 = (X(s_0), Y(s_0), Z(s_0))$. Então o vector tangente a C vem dado por $(X'(s_0), Y'(s_0), Z'(s_0))$. O cálculo destas derivadas faz-se a partir do sistema

$$4XX' + 6YY' + 2ZZ' = 0, \quad 2XX' + 4YY' - Z' = 0,$$

resultante das equações cartesianas da curva C, por derivação. Resolvendo este sistema obtemos

$$X'(s_0) = -\frac{4Z(s_0) + 3}{2X(s_0)}Z'(s_0), \quad Y'(s_0) = \frac{Z(s_0) + 1}{Y(s_0)}Z'(s_0),$$

pelo que um vector tangente a C em \mathbf{P}_0 viria dado por $\mathbf{t} = (27, 28, 4)$.

Segundo método: Determine dois vectores, \mathbf{n}_1 , perpendicular à superfície de equação $2x^2 + 3y^2 + z^2 = 47$ no ponto \mathbf{P}_0 , i.e. $\mathbf{n}_1 = (4x, 6y, 2z)_{\mathbf{P}_0}$, e outro \mathbf{n}_2 , perpendicular à superfície de equação $x^2 + 2y^2 = z$ no ponto \mathbf{P}_0 , i.e. $\mathbf{n}_2 = (2x, 4y, -1)_{\mathbf{P}_0}$. O vector tangente a C em \mathbf{P}_0 vem dado pelo produto vectorial destes dois vectores, i.e.

$$\mathbf{t} = \mathbf{n}_1 \times \mathbf{n}_2 \begin{vmatrix} i & j & k \\ 4x & 6y & 2z \\ 2x & 4y & -1 \end{vmatrix}_{\mathbf{P}_0}.$$

c.q.d.

4. EXTREMOS RELATIVOS

4.1. Condições de primeira ordem. Se uma função real de duas ou mais variáveis reais, f , definida no aberto D , tem um *máximo* (respectivamente, um *mínimo*) no ponto $(x_0, y_0) \in D$, i.e. $f(x_0, y_0) \geq f(x, y)$ (respectivamente $f(x_0, y_0) \leq f(x, y)$), $(x, y) \in \mathcal{V}_{(x_0, y_0)}$; então é

também máximo (respectivamente, mínimo) de f sobre qualquer curva de intersecção de $z = f(x, y)$ com um plano paralelo a OXZ ou OYZ .

Teorema (Condição necessária de extremo). *Uma condição necessária para que o ponto de coordenadas $(x_1^0, \dots, x_n^0) \in D \subset \mathbb{R}^n$ seja ponto extremante de f , i.e. maximizante ou minimizante de f , é que*

$$\frac{\partial f}{\partial x_1}(x_1^0, \dots, x_n^0) = 0, \dots, \frac{\partial f}{\partial x_n}(x_1^0, \dots, x_n^0) = 0.$$

Um ponto $\mathbf{x}^0 \in D$ que verifica o sistema anterior diz-se *ponto crítico* de f em D .

Observação . Neste ponto assume particular importância os chamados raciocínios indirectos, i.e. seja f uma função diferenciável que atinge um mínimo em D . Se somente existir um ponto que verifica a condições necessárias de extremo então o mínimo da função deve alcançar-se nesse ponto.

Problema 4.1. *Construir uma caixa de latão de volume fixo V usando a menor quantidade de material.*

Resolução. Denotando por x, y os lados da base da caixa, a altura, h , virá dada por $V/(xy)$, pelo que a superfície da caixa vem dada por

$$S(x, y) = xy + \frac{V}{xy}(2x + 2y) = xy + 2V\left(\frac{1}{x} + \frac{1}{y}\right).$$

Note-se que $x, y > 0$ e o problema reduz-se a minimizar a função S em $D = \mathbb{R}^+ \times \mathbb{R}^+$.

Pela condição necessária de extremo temos

$$\frac{\partial S}{\partial x} = y - \frac{2V}{x^2} = 0, \quad \frac{\partial S}{\partial y} = x - \frac{2V}{y^2} = 0.$$

Assim $(x_0, y_0, h_0) = (\sqrt[3]{2V}, \sqrt[3]{2V}, \sqrt[3]{V/4})$ é o candidato a solução do nosso problema. Note-se que, como D não é um compacto, não podemos garantir que S atinga o seu valor máximo em D .

Vejam os como demonstrar que este ponto é a solução que procuramos. Para tal seja (x_1, y_1) um ponto arbitrário de D e $N = S(x_1, y_1)$. Tome-se $R \in \mathbb{R}^+$ tal que $R > N$, $2VR > N$ e construa-se um quadrado de lado R , D_R , cujos lados são paralelos aos eixos coordenados e estão a uma distância igual a $1/R$ dos eixos, e o vértice inferior esquerdo está na posição $(1/R, 1/R)$.

Procuremos um infimo para S em $D \setminus D_R$. Assim em pontos de D tais que $x < 1/R$, temos que

$$S(x, y) > 2V/x > 2VR > N.$$

Da mesma forma se mostra que para pontos de D tais que $y < 1/R$ então $S(x, y) > N$. Além disso, nos pontos de D tais que $x > R/2$ e $y < 1/R$ temos

$$S(x, y) > xy > \frac{1}{R}R^2 = R > N,$$

e também nos pontos tais que $x < 1/R$ e $y > R + 1/R$ se tem $S(x, y) > N$.

Acabámos de provar que para todos os pontos de D situados fora do quadrado D_R se verifica $S(x, y) > N$, pelo que o mínimo de S terá de ser alcançado em pontos de D_R .

Agora D_R é um compacto, então S atinge aí o seu valor máximo e mínimo. *c.q.d.*

4.2. Condições de segunda ordem.

Definição 4.1. Seja A uma matriz $n \times n$. Dizemos que a matriz A é *definida positiva* (respectivamente, *definida negativa*) se para todo o $\mathbf{x} \in \mathbb{R}^n$ se tiver

$$\mathbf{x}^t A \mathbf{x} > 0, \quad \text{respectivamente,} \quad \mathbf{x}^t A \mathbf{x} < 0.$$

A matriz A é *semi-definida positiva* (respectivamente, *negativa*) se as desigualdades estritas forem substituídas por desigualdades em sentido lato.

Definindo *forma quadrática*, f , associada a uma matriz $n \times n$, A , como uma função real definida em \mathbb{R}^n por $f(\mathbf{x}) = \mathbf{x}^t A \mathbf{x}$. Nestas condições dizemos que A é a *matriz da forma quadrática* f . Diremos que a *forma quadrática* é (*semi*)-*definida positiva* (respectivamente, *negativa*) se a matriz da forma quadrática for (*semi*)-definida positiva (respectivamente, negativa).

Exemplo 4.1. A matriz hessiana de f em \mathbf{x}^0 , $\mathcal{H}_f(\mathbf{x}^0)$, é a matriz da forma quadrática $d^2 f(\mathbf{x}^0)$.

Teorema (Condição suficiente de extremo). *Considere-se uma função real f definida num aberto $D \subset \mathbb{R}^n$, de classe C^2 . Seja $\mathbf{x}^0 \in D$ um ponto crítico de f , i.e. $\text{grad } f(\mathbf{x}^0) = \mathbf{0}$. Então:*

- (1) *Se $d^2 f(\mathbf{x}^0)$ é definida positiva então f atinge em \mathbf{x}^0 um mínimo relativo.*
- (2) *Se $d^2 f(\mathbf{x}^0)$ é definida negativa então f atinge em \mathbf{x}^0 um máximo relativo.*
- (3) *Se $d^2 f(\mathbf{x}^0)$ não é semi-definida então \mathbf{x}^0 não é extremo de f .*

Observação . Se $d^2 f(\mathbf{x}^0)$ é semi-definida e $d f(\mathbf{x}^0) = \text{grad } f(\mathbf{x}^0) = \mathbf{0}$ este resultado pode aplicar-se, quando e só quando $|\mathcal{H}_f(\mathbf{x}^0)| \neq 0$.

Ideia da Demonstração. Como f é uma função de classe C^2 , da fórmula de Taylor sabemos que

$$f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) = d f(\mathbf{x}^0) \mathbf{h} + \frac{1}{2!} d^2 f(\mathbf{x}^0) \mathbf{h} + \|\mathbf{h}\|^2 \epsilon(\mathbf{h}), \quad \lim_{\mathbf{h} \rightarrow \mathbf{0}} \epsilon(\mathbf{h}) = 0.$$

Mas $d f(\mathbf{x}^0) \mathbf{h} = \mathbf{0}$, pois \mathbf{x}^0 é ponto crítico de f , pelo que o sinal $[f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0)]$ vem determinado pelo de sinal $[d^2 f(\mathbf{x}^0)(\mathbf{h}/\|\mathbf{h}\|) + \epsilon(\mathbf{h})]$. Mais, uma vez por ser f uma função de classe C^2 , a função $d^2 f(\mathbf{x}^0)(\mathbf{h}/\|\mathbf{h}\|)$ é contínua sobre, $\bar{B}_1(\mathbf{0}) = \{\mathbf{x} \in \mathbb{R}^2 : d(\mathbf{x}, \mathbf{0}) \leq 1\}$, que é um compacto. Pelo teorema de Weierstrass, $d^2 f(\mathbf{x}^0)$, atinge em $\bar{B}_1(\mathbf{0})$ o seu valor máximo e mínimo. Se o sinal de $d^2 f(\mathbf{x}^0)(\mathbf{h}/\|\mathbf{h}\|)$ estiver determinado temos

$$\text{sinal}[f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0)] = \text{sinal}[d^2 f(\mathbf{x}^0) \mathbf{h}/\|\mathbf{h}\|],$$

que é que se tem nos casos enunciados no teorema.

c.q.d.

Veamos um resultado auxiliar de álgebra linear para a análise dos pontos críticos.

Teorema (Sylvester). *Seja $A = [a_{j,k}]$ uma matriz de dimensão $n \times n$ simétrica. Então A é definida positiva se e somente se os menores principais de A , i.e.*

$$\Delta_1 = a_{1,1}, \quad \Delta_2 = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}, \quad \dots, \quad \Delta_n = |A|,$$

são positivos.

Observação .

- (1) Uma condição necessária e suficiente para que a matriz A seja semi-definida positiva é que $\Delta_k \geq 0$, $k = 1, \dots, n$.
- (2) Dizemos que uma matriz, A , é (semi)-definida negativa, quando e só quando a matriz $-A$ é (semi)-definida positiva, i.e. $(-1)^n \Delta_k$ (não-negativos) positivos, $k = 1, \dots, n$.
- (3) Para aplicar o teorema de Sylvester à matriz hessiana de f , temos de substituir os elementos da matriz A por,

$$a_{j,k} = \frac{\partial^2 f}{\partial x_j \partial x_k}, \quad j, k = 1, \dots, n.$$

Problema 4.2. *Considere-se o sistema material $\{(\mathbf{P}_1, m_1), \dots, (\mathbf{P}_n, m_n)\}$, onde os pontos \mathbf{P}_j têm coordenadas (x_j, y_j, z_j) e m_j são as massas de \mathbf{P}_j , $j = 1, \dots, n$. Determinar o ponto $\mathbf{P} = (x, y, z) \in \mathbb{R}^3$ que minimiza o momento de inércia do sistema dado, i.e. que minimiza a função, M , de expressão analítica*

$$M(x, y, z) = \sum_{j=1}^n m_j ((x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2).$$

Resolução. Começemos por determinar os pontos críticos de M , i.e.

$$2 \sum_{j=1}^n (x - x_j) m_j = 0, \quad 2 \sum_{j=1}^n (y - y_j) m_j = 0, \quad 2 \sum_{j=1}^n (z - z_j) m_j = 0,$$

e portanto temos um único ponto crítico \mathbf{P}_0 com as coordenadas do *baricentro* ou *centro de massa* do sistema, i.e. no ponto de coordenadas

$$x = \frac{1}{m} \sum_{j=1}^n m_j x_j, \quad y = \frac{1}{m} \sum_{j=1}^n m_j y_j, \quad z = \frac{1}{m} \sum_{j=1}^n m_j z_j,$$

onde $m = \sum_{j=1}^n m_j$. Verifiquemos que neste ponto a função M atinge o seu valor mínimo.

Para tal calculemos

$$d^2 M(\mathbf{P}_0) = 2m \mathbf{I}_{n \times n},$$

que é uma matriz definida positiva.

c.q.d.

Problema 4.3. *Determinar os extremos relativos da função real, f , definida em \mathbb{R}^2 por $f(x, y) = x^4 + y^4 - 2x^2 + 4xy - 2y^2$.*

Resolução. Começemos por determinar os pontos estacionários de f , i.e.

$$\frac{\partial f}{\partial x} = 4x^3 - 4x + 4y = 0, \quad \frac{\partial f}{\partial y} = 4y^3 - 4y + 4x = 0,$$

ou ainda,

$$x^3 = -y^3, \quad 4x(x^2 - 2) = 0,$$

e portanto, os pontos críticos são $\mathbf{P}_1 = (0, 0)$, $\mathbf{P}_2 = \pm(\sqrt{2}, -\sqrt{2})$.

Analisemos para cada um destes pontos as condições de segunda ordem. Assim

$$d^2 f = \begin{bmatrix} 12x^2 - 4 & 4 \\ 4 & 12y^2 - 4 \end{bmatrix},$$

pelo que

$$d^2 f(\mathbf{P}_2) = \begin{bmatrix} 20 & 4 \\ 4 & 20 \end{bmatrix}, \quad d^2 f(\mathbf{P}_1) = \begin{bmatrix} -4 & 4 \\ 4 & -4 \end{bmatrix}.$$

Assim, nos pontos \mathbf{P}_2 a função f atinge um máximo relativo, pois os menores principais são positivos. Mas para $(0, 0)$ o determinante da matriz $d^2 f$ é nulo, pelo que o teorema não pode ser aplicado. No entanto, $f(x, y) = x^4 + y^4 - 2(x - y)^2$, pelo que em qualquer vizinhança de $(0, 0)$, existem pontos para os quais f é negativo (por exemplo, em $(h, -h)$ e $h < 1$) e f é positivo (por exemplo, em (h, h) e $h < 1$), donde se conclui que $(0, 0)$ não é ponto extremante de f . *c.q.d.*

4.3. Extremos condicionados.

Teorema (Multiplicadores de Lagrange). *Sejam f uma função real de n variáveis reais, de classe C^1 e \mathbf{g} uma função vectorial de classe C^1 definida em \mathbb{R}^n , com contradomínio em \mathbb{R}^m e $m < n$, com funções coordenadas g_1, \dots, g_m , funcionalmente independentes. Para que f tenha um extremo relativo em \mathbf{x}^0 condicionado pela condição $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ é necessário que existam parâmetros reais $\lambda_1, \dots, \lambda_m$ tais que a função de Lagrange ou lagrangiano*

$$\phi = f + \lambda_1 g_1 + \dots + \lambda_m g_m$$

tenha um ponto crítico em \mathbf{x}^0 , i.e. $d\phi(\mathbf{x}^0) = \mathbf{0}$, que verifique as condições

$$\mathbf{g}(\mathbf{x}^0) \equiv (g_1(\mathbf{x}^0), \dots, g_m(\mathbf{x}^0)) = \mathbf{0}.$$

Teorema (Condição suficiente de extremo condicionado). *Sejam f e \mathbf{g} como no teorema anterior mas de classe C^2 . Uma condição suficiente para que um ponto crítico \mathbf{x}^0 seja extremo de f condicionado por $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ é que $d^2 \phi(\mathbf{x}^0)\mathbf{h}$ tenha sinal determinado, sobre todas as direcções \mathbf{h} que verificam $d\mathbf{g}(\mathbf{x}^0)\mathbf{h} = \mathbf{0}$.*

Problema 4.4. *Determine os extremos relativos da função*

$$f(x, y) = x \ln x + y \ln y + z \ln z$$

definida para $x, y, z > 0$ e sujeita à seguinte condição $x + y + z = 3a$.

Resolução. Definimos o lagrangiano, que é uma função real definida em $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$ por $\phi(x, y, z) = f(x, y, z) - \lambda(x + y + z - 3a)$, e resolvemos o sistema

$$\ln x = \ln y = \ln z = \lambda - 1, \quad x + y + z = 3a.$$

Assim, $(x, y, z) = (a, a, a)$ é o único candidato a solução do problema proposto.

Como o conjunto $\{(x, y, z) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ : x + y + z = 3a\}$ não é um compacto temos de passar às condições de segunda ordem.

Começamos por determinar as direcções ortogonais a $\text{grad } g(a, a, a)$ onde $g(x, y, z) = x + y + z - 3a$, i.e. as direcções $\mathbf{u} = (u_1, u_2, u_3)$ tais que $(u_1, u_2, u_3) \cdot (1, 1, 1) = 0$. Agora

$$d^2 \phi(a, a, a)\mathbf{u} = \begin{bmatrix} u_1 & u_2 & -u_1 - u_2 \end{bmatrix} \begin{bmatrix} 1/x & 0 & 0 \\ 0 & 1/y & 0 \\ 0 & 0 & 1/z \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ -u_1 - u_2 \end{bmatrix}$$

pelo que $d^2 \phi(a, a, a) = u_1^2 + u_2^2 + (u_1 + u_2)^2$ é positivo, logo (a, a, a) é um minimizante do problema dado. *c.q.d.*

Exercício . Seja C a curva de equação cartesiana

$$z = x^3 + y, \quad \frac{3}{2}x^2 + y^2 = \frac{7}{4}.$$

- (a) Identifique o lagrangiano que lhe permite calcular os pontos de cota máxima de C .
- (b) Sabendo que os pontos candidatos a extremante são

$$\pm (0, \sqrt{7}/2, \sqrt{7}/2), \quad \pm (1, 1/2, 3/2), \quad \pm (\sqrt{6}/6, \sqrt{6}/2, 19\sqrt{6}/36),$$

resolva o problema enunciado na alínea anterior. Justifique convenientemente a sua resposta.

ÍNDICE REMISSIVO

- \mathbb{R}^n -metria, 1
- área
 - rectângulo, 2
- baricentro, 19
- bola aberta, 2

- centro de massa, 19
- conjunto
 - aberto, 2
 - compacto, 5
- curva
 - contínua, 4
 - de nível, 3
 - em \mathbb{R}^2 , 11
 - simples, 14
- derivada
 - parcial, 5
 - representação geométrica, 5
- diferenciabilidade
 - condição necessária, 7
 - condição suficiente, 8
 - função composta, 9
- diferencial
 - função vectorial, 8
 - função real, 7
- diferenciável
 - função real, 7
 - função vectorial, 8
- distância, 1

- espaço
 - afim, 1
 - euclideo, 1
 - vectorial, 1
- extremos
 - condição necessária, 17
 - condição suficiente, 18

- forma quadrática, 18
 - definida negativa, 18
 - definida positiva, 18
 - semi-definida negativa, 18
 - semi-definida positiva, 18
- funções
 - analíticas, 10
 - bijectivas, 13
 - contínuas, 4
 - de Lagrange, 20
 - definidas implicitamente, 11
 - funcionalmente dependentes, 13
 - funcionalmente independentes, 13
 - uniformemente contínuas, 4
 - vectoriais, 3

- lagrangiano, 20
- limite
 - direcciona, 4
 - função real, 3
 - função vectorial, 3

- matriz
 - característica, 14
 - definida negativa, 18, 19
 - definida positiva, 18, 19
 - forma quadrática, 18
 - menores, 14
 - menores principais, 19
 - semi-definida negativa, 18, 19
 - semi-definida positiva, 18, 19
- momento
 - de inércia, 19
- Multiplicadores de Lagrange
 - condição suficiente, 20
 - método, 20

- ponto
 - crítico, 17
 - de mínimo, 17
 - de máximo, 17
 - extremante, 17
 - múltiplo, 14

- superfície diferenciável, 15
 - equação cartesiana, 15
 - vector normal, 15

- Taylor
 - polinómio, 11
 - série, 11
- teorema
 - da função implícita
 - diferenciabilidade, 12
 - existência, 12
 - da sandwich, 3
 - de Cauchy, 3
 - de Sylvester, 19
 - de Weierstrass, 5
 - função implícita, 12
 - função inversa, 13
 - funcionalmente dependentes
 - condição necessária, 14
 - condição suficiente, 14
 - incrementos finitos, 6
 - limite por curvas, 4