João Carlos Beirão & Bhangy Cassy

CÁLCULO DIFERENCIAL EM $\mathbb{R}^n$

Maputo
2005

Ficha Técnica

Título: Análise Matemática: Cálculo Diferencial em $\mathbb{R}^n$

Autores: João Carlos Beirão e Bhangy Cassy

Editor: Imprensa Universitária, UEM

Composição: José António Nhavoto

Maquetização e Impressão: Imprensa Universitária, UEM

Prefácio

O objectivo fundamental deste livro é apresentar uma exposição sistemática e moderna do cálculo diferencial de funções de várias variáveis que possa servir como texto de apoio aos cursos que sobre essa matéria são geralmente ministrados no ensino superior aos estudantes de matemática, ciências e engenharia.

Os conteúdos abrangidos são os normalmente incluídos em obras deste tipo e foram programados de modo a poderem ser tratados num curso semestral com três horas semanais de aula. De acordo com a orientação modernamente utilizada neste domínio houve, entretanto, a preocupação de alicerçar a exposição da matéria no estudo prévio dos espaços métricos, por se considerar ser essa a melhor forma de introduzir o estudante na linguagem e nos conceitos fundamentais da matemática moderna e de o capacitar para uma ulterior ampliação dos seus conhecimentos.

Como pré-requisito para a utilização do livro, pressupõe-se que o leitor possua os conhecimentos geralmente obtidos nos cursos básicos de cálculo, álgebra e geometria analítica.

Tendo em vista a necessidade de garantir o grau de rigor que é de exigir no ensino superior, houve a preocupação de apresentar a demonstração de todos os resultados importantes com excepção de alguns poucos que, pela sua delicadeza, exigem um nível de conhecimentos mais avançados e que por isso se remetem para a bibliografia. Tais resultados foram de uma forma geral designada como Proposições, reservando-se a designação Teorema apenas para aqueles que se destacam pela sua importância ou que tradicionalmente são designados desse modo.

Com vista a desenvolver técnicas numéricas e analíticas para abordar os problemas que surgem nas aplicações práticas, incluem-se ao longo do texto inúmeros exemplos resolvidos e problemas com resposta.

Os Autores

Índice

Bibliografia **221**

Capítulo 1

Espaços métricos

O conceito de espaço métrico, formulado no inicio do século XX pelo matemático francês Maurice Fréchet, assenta num conjunto de axiomas que não são mais do que o refinamento matemático da noção intuitiva de distância utilizada na vida corrente.

À noção de distância devem-se associar dois conceitos de fundamental importância: **convergência** e **limite**. Com efeito,

(i) uma sucessão (a_n) de elementos de um dado conjunto A diz-se **convergente** para o elemento $a \in A$ quando, para n suficientemente grande, os termos a_n se tornam arbitrariamente próximos de a.

De forma análoga:

(ii) uma função $f : A \to B$ diz-se que converge para o limite ℓ ao tender x para a quando, para valores de x suficientemente próximos de a, os valores de $f(x)$ se tornam arbitrariamente próximos de ℓ.

É evidente pois que ambas definições assentam no pressuposto de que é possível averiguar se dois elementos de um conjunto estão ou não "próximos" um do outro. Ora isso implica que para as definições dadas terem sentido é indispensável que os conjuntos considerados possuam uma estrutura que permita estabelecer a "proximidade" dos seus elementos.

Portanto quando por exemplo se pretende verificar qual dentre os elementos x e y de um dado conjunto A está mais próximo de um outro elemento $a \in A$, não se pretende mais do que avaliar a distância de cada um deles a a. Todavia, isso exige que no conjunto A se tenha previamente definido a noção de distância. Quando assim acontece o conjunto A deixa de ser um conjunto qualquer e passa a constituir o que denomina um **espaço métrico.**

1.1 Definição do espaço métrico

Definição 1.1 (**Métrica**).

*Dado um conjunto não vazio E, chama-se **métrica** (ou função distância) nesse conjunto, a qualquer aplicação $d : E \times E \to \mathbb{R}$ que associa a cada par de pontos $(x, y) \in E \times E$ um número real $d(x, y)$, chamado distância do ponto x ao ponto y, e que goza das seguintes propriedades:*

(i) ***Positividade:***
$$d(x,y) > 0 \quad \text{se} \quad x \neq y; \quad \text{e} \quad d(x,x) = 0;$$

(ii) ***Simetria:***
$$d(x,y) = d(y,x);$$

(iii) ***Desigualdade triangular:***
$$d(x,z) \leq d(x,y) + d(y,z).$$

A propriedade (iii) chama-se desigualdade triangular (Figura 9.11) porque exprime que cada lado dum triângulo não excede a soma dos outros dois.

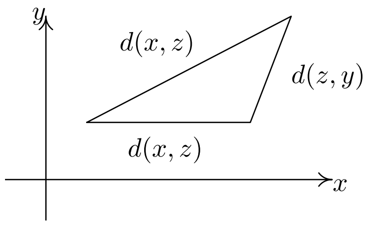

Figura 1.1: Desigualdade triangular no plano

Como decorre da própria definição, num mesmo conjunto E podem definir-se tantas métricas (ou funções distância), quantas as aplicações distintas de $E \times E$ em $\mathbb{R}$ que satisfaçam as propriedades indicadas.

Definição 1.2 (Espaço métrico).

> *Chama-se **espaço métrico** ao par (E,d) formado por um conjunto E e uma métrica d definida em E.*

Por comodidade, sempre que não haja perigo de confusão, podemos referir-nos ao espaço métrico (E,d), dizendo apenas "**espaço métrico** E", isto é, deixando subentendida a métrica d.

Exemplo 1.1 (Recta real $\mathbb{R}$). O exemplo mais importante de espaço métrico é o conjunto $\mathbb{R}$ dos números reais, munido da métrica $d(x,y) = |x-y|$. É fácil de verificar que a aplicação $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ assim definida satisfaz as condições (i), (ii), (iii). Com efeito, como se sabe do estudo das propriedades dos números reais:

(a) $d(x, y) = |x - y| > 0$ se $x \neq y$ e $d(x, x) = |0| = 0$;

(b) $d(x, y) = |x - y| = |y - x| = d(y, x)$;

(c) $d(x, z) = |x - z| = |x - y + y - z| \leq$

$$\leq |x - y| + |y - z| = d(x, y) + d(y, z).$$

O conjunto $\mathbb{R}$ munido desta métrica, isto é, o espaço métrico $(\mathbb{R}, d)$, com $d(x, y) = |x - y|$, chama-se **recta real**. Sempre que $\mathbb{R}$ se considera um espaço métrico, sem indicar expressamente qual a métrica de que está munido, subentende-se que é a função distância acima indicada. ◄

Exemplo 1.2. A função distância atrás considerada não é, porém, evidentemente, a única que se pode considerar no conjunto dos números reais $\mathbb{R}$. Assim, por exemplo, se f for uma função real estritamente monótona definida sobre $\mathbb{R}$, a função

$$d(x, y) = |f(x) - f(y)|$$

é também uma métrica em $\mathbb{R}$.

Com efeito,

(a) $d(x, y) = |f(x) - f(y)| > 0$ se $x \neq y$ e $d(x, x) = |0| = 0$;

(b) $d(x, y) = |f(x) - f(y)| = |f(y) - f(x)| = d(y, x)$;

(c) $d(x, y) = |f(x) - f(z)| = |f(x) - f(y) + f(y) - f(z)| \leq$

$$\leq |f(x) - f(y)| + |f(y) - f(z)| = d(x, y) + d(y, z). \quad ◄$$

Exemplo 1.3 (Plano Euclidiano $\mathbb{R}^2$). No plano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, pode-se definir uma métrica tomando o conjunto de todos os pares ordenados com a distância euclidiana d definida por,

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

Com efeito,

(a) $d(x, y) > 0 \quad \text{se } x \neq y \quad \text{e} \quad d(x, x) = 0;$

(b) $d(x, y) = d(y, x);$

(c) $$\begin{aligned} d(x, z) &= \sqrt{(x_1 - z_1)^2 + (x_2 - z_2)^2} = \\ &= \sqrt{(x_1 - y_1 + y_1 - z_1)^2 + (x_2 - y_2 + y_2 - z_2)^2} \leq \\ &\leq \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} + \sqrt{(y_1 - z_1)^2 + (y_2 - z_2)^2} = \\ &= d(x, y) + d(y, z), \end{aligned}$$

de acordo com a **desigualdade de Minkowski** .◀

Exemplo 1.4. Ainda em $\mathbb{R}^2$ pode também definir-se uma métrica, pondo, para $x = (x_1, x_2)$ e $y = (y_1, y_2)$:

$$d(x, y) = |x_1 - y_1| + |x_2 - y_2|.$$

Na verdade:

(a) $d(x, y) > 0 \quad \text{se } x \neq y \quad \text{e} \quad d(x, x) = 0;$

(b) $d(x, y) = d(y, x);$

(c) $$\begin{aligned} d(x, z) &= |x_1 - z_1| + |x_2 - z_2| = \\ &= |x_1 - y_1 + y_1 - z_1| + |x_2 - y_2 + y_2 - z_2| \leq \\ &\leq |x_1 - y_1| + |y_1 - z_1| + |x_2 - y_2| + |y_2 - z_2| = \\ &= d(x, y) + d(y, z). \end{aligned}$$ ◀

Exemplo 1.5. Seja E um conjunto qualquer. Pode-se definir em E uma métrica d pondo $d(x,x)=0$ e $d(x,y)=1$ se $x \neq y$.

Com efeito:

(a) $d(x,y)>0 \quad$ se $x \neq y \quad$ e $\quad d(x,x)=0$;

(b) $d(x,y)=d(y,x)$ quer seja $x=y$, quer seja $x \neq y$;

(c) Se $x \neq z$ e $x \neq y$ e $y \neq z$, então
$$d(x,z)=1<1+1=d(x,y)+d(y,z);$$

Se $x \neq z$, mas $x=y$ ou $y=z$, então
$$d(x,z)=1=1+0=d(x,y)+d(y,z)$$

e portanto, em conclusão, qualquer que seja y, é

$$d(x,z) \leq d(x,y)+d(y,z). \quad \blacktriangleleft$$

A métrica assim definida chama-se **métrica discreta**. É uma métrica trivial, normalmente só usada em contra-exemplos. Por ela se vê, porém, que qualquer conjunto se pode tornar um espaço métrico.

1.1.1 Subespaço métrico

Todo o subconjunto X dum espaço métrico E possui uma estrutura natural de espaço métrico. Basta, para isso, definir a distância entre dois pontos de X como a distância entre esses pontos quando considerados como pontos de E. A métrica assim definida em X chama-se a **métrica induzida em X pela métrica E** e o espaço métrico X resultante diz-se um **subespaço métrico** de E.

1.1.2 Isometrias

Dados dois espaços métricos (E, d_1) e (F, d_2), chama-se **isometria** a toda a aplicação bijectiva $f : E \to F$ que preserva as distâncias, isto é, que satisfaz, para quaisquer $x,\ y \in E$, a condição

$$d_2(f(x), f(y)) = d_1(x, y).$$

Dois espaços E e F dizem-se **isométricos** se existir uma isometria de E sobre F. Nessa altura toda a propriedade demonstrada em E relativa apenas a distâncias entre elementos de E tem em F uma propriedade correspondente.

Se, embora preservando as distâncias, a aplicação f de E em F for apenas injectiva, então denomina-se uma **imersão isométrica** de E em F.

1.2 Noções métricas

De um modo geral, dá-se o nome de **noções métricas** num conjunto E a todas as noções que se podem definir logicamente a partir da métrica (ou função distância) própria de E. Entre elas figura, evidentemente, a noção de distância entre dois pontos, que é a noção métrica primitiva. Vejamos algumas outras.

Definição 1.3 (**Bola aberta**).

> *Dados um espaço métrico (E, d), um ponto a de E e um número real $r > 0$, chama-se **bola aberta** com centro em a e raio r (Figura 1.2) ao conjunto dos pontos x de E cuja distância ao ponto a é inferior a r, isto é, ao conjunto $B(a, r)$ tal que*
>
> $$B(a, r) = \{x \in E : d(a, x) < r\}.$$

De forma análoga:

Definição 1.4 (**Bola fechada**).

> *Dados um espaço métrico (E, d), um ponto a de E e um número real $r > 0$, chama-se **bola fechada** de centro a e raio r (Figura 1.3) ao conjunto dos pontos x de E tais que $d(a, x) \leq r$, isto é, ao conjunto*
>
> $$B'(a, r) = \{x \in E : d(a, x) \leq r\}.$$

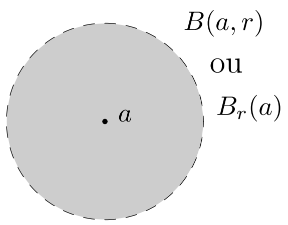

Figura 1.2: Bola aberta

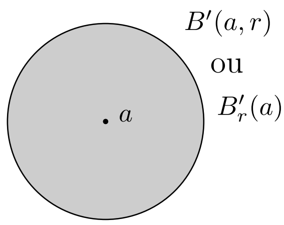

Figura 1.3: Bola fechada

Definição 1.5 (**Superfície esférica**). *Das definições 1.3 e 1.4, a bola fechada $B'(a, r)$ é formada pela bola aberta $B(a, r)$ e pelos pontos x de E cuja distância ao ponto a é igual a r. Estes últimos pontos formam um conjunto que se anota $S(a, r)$ e se chama a esfera de centro a e raio r ou **superfície esférica** (Figura 1.4) e denota-se:* $S(a, r) = \{x \in E : d(a, x) = r\}$.

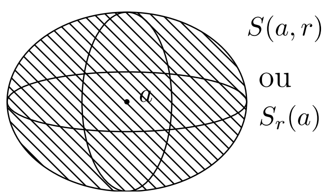

Figura 1.4: Superfície esférica

Exemplo 1.6. Na recta $\mathbb{R}$, a bola aberta $B(a,r)$ é o intervalo aberto $]a-r,\ a+r[$, a bola fechada $B'(a,r)$ é o intervalo fechado $[a-r,\ a+r]$ e a esfera $S(a,r)$ reduz-se ao par de pontos $\{a-r,\ a+r\}$. ◀

Exemplo 1.7 (Espaço métrico discreto). Num espaço com métrica discreta, uma bola fechada de raio $r<1$ reduz-se ao centro, uma bola fechada de raio $r=1$ é o espaço inteiro e todas as esferas de raio $r=2$ coincidem e são vazias. ◀

Exemplo 1.8. No espaço $\mathbb{R}^2$, a bola aberta $B(a,r)$ chama-se disco e é o interior da circunferência de centro a e raio r; a esfera $S(a,r)$ chama-se circunferência e é a circunferência de centro a e raio r. ◀

Definição 1.6.

> *Dados um espaço métrico (E,d), um ponto $a\in E$ e um subconjunto não vazio $A\subset E$, chama-se* ***distância de a a*** *A ao ínfimo das distâncias de a aos pontos x de A, ou seja:*
> $$d(a,A)=\inf_{x\in A} d(a,x).$$

Assim, a distância $d(\delta,A)$ do ponto a ao conjunto A é o único número real m tal que

(a) $m\le d(a,x),\ \forall x\in A$;

(b) Dado qualquer $\epsilon>0$, existe $x\in E$ tal que $d(a,x)-\epsilon<m$.

Se $a\in A$, então $d(a,A)=0$, mas a recíproca é falsa.

Na recta real, por exemplo, se A for o intervalo aberto $]1,2[$ tem-se $d(1,A)=0$ e $d(2,A)=0$ sem que, no entanto, $1\in A$ nem $2\in A$.

Em geral, $d(a, A) = 0$ significa que dado $\epsilon > 0$ arbitrário, existe $x \in A$ tal que $d(a, x) < \epsilon$, isto é, existem pontos de A arbitrariamente próximos de a.

Definição 1.7.

> *Dados dois subconjuntos não vazios A e B de um espaço métrico (E, d), chama-se **distância de** A **a** B ao ínfimo das distâncias dos pontos de A aos pontos de B, isto é:*
>
> $$d(A, B) = \inf_{x \in A, y \in B} d(x, y).$$

Como é evidente, se os conjuntos A e B não forem disjuntos, isto é, se $A \cap B \neq \emptyset$, a sua distância é nula: $d(A, B) = 0$. O recíproco, porém, não é necessariamente verdadeiro, isto é, dois conjuntos podem ser disjuntos e a sua distância ser nula. É o que acontece, por exemplo, em $\mathbb{R}^2$, quando $A = \{(x, 0) : x \in \mathbb{R}\}$ é o eixo real e $B = \{(x, \frac{1}{x}) : x \in \mathbb{R}\}$ é um ramo de hipérbole. Nesse caso, é $A \cap B = \emptyset$, mas $d(A, B) = 0$, pois se $x = (x, 0)$ e $y = (x, \frac{1}{x})$, $d(x, y) = \frac{1}{x}$ pode tornar-se tão pequena quanto se queira, desde que tome x suficientemente grande.

Definição 1.8.

> *Seja A um conjunto não vazio do espaço métrico (E, d). Chama-se **diâmetro do conjunto** A ao supremo das distâncias de dois pontos variáveis de A, isto é, ao número:* $\delta(A) = \sup_{x, y \in A} d(x, y)$.

De acordo com a definição 1.8 o diâmetro dum conjunto é sempre um número real positivo ou infinito.

O diâmetro é nulo se e só se o conjunto A for formado por um só ponto. Nessa altura, e só então, tem-se $\delta(A) = 0$.

Definição 1.9.

Um subconjunto A dum espaço métrico (E, d) diz-se limitado quando o seu diâmetro é finito; equivale isso a dizer que existe pelo menos uma bola de E que contém A.

1.3 Subconjuntos particulares dos espaços métricos

1.3.1 Partes abertas

Definição 1.10 (**Parte aberta ou conjunto aberto**).

*Diz-se que um subconjunto A de um espaço métrico (E, d) é uma **parte aberta** (ou conjunto aberto) de E quando todo o ponto $a \in A$ é centro de uma bola aberta inteiramente contida em A, isto é,*

$$\forall a \in A, \ \exists \epsilon > 0 : \ B(a, \epsilon) \subset A.$$

Ou seja, para cada $a \in A$ existe $\epsilon > 0$ tal que, se $x \in E$ e $d(a, x) < \epsilon$, então $x \in A$.

Proposição 1.1. *Toda a bola aberta $B(a,r)$ de um espaço métrico (E,d) é um subconjunto aberto de E.*

Demonstração. Com efeito, para cada ponto $x \in B(a,r)$ tem-se $d(a,x) < r$. E, escolhendo $\epsilon = r - d(a,x) > 0$ é fácil ver que a bola $B(x,\epsilon)$ está contida em $B(a,r)$. Donde, se $y \in B(x,\epsilon)$ então $d(x,y) < \epsilon = r - d(a,x)$ e portanto

$$d(a,y) \leq d(a,x) + d(x,y) < r - d(a,x) + d(a,x) = r.$$

Logo, $y \in B(a,r)$ e portanto $B(x,\epsilon) \subset B(a,r)$. ■

Proposição 1.2. *Os subconjuntos abertos de um espaço métrico (E,d) gozam das seguintes propriedades:*

P1. *O espaço inteiro E e o conjunto vazio são subconjuntos abertos de E.*

Demonstração. Com efeito, o espaço total E é aberto, visto que qualquer bola aberta centrada num dos seus pontos está contida em E. Como é lógico, os pontos que não pertencem a um dado espaço métrico não têm qualquer relevância nas questões relativas ao contexto desse espaço. Ora, um conjunto é aberto ou não apenas em relação a um específico espaço métrico que o contém e nunca em relação a si próprio. Assim, para mostrar que $\emptyset$ é aberto basta notar que um subconjunto $X \subset E$ só não é aberto quando existe um ponto $x \in X$ tal que nenhuma bola de centro x está contida em X. Ora, como não existe nenhum $x \in \emptyset$ o conjunto vazio não viola a condição de definição das partes abertas.

P2. *A união de qualquer família (finita ou infinita)de subconjuntos abertos de E é um subconjunto aberto de E.*

Demonstração. Com efeito, seja $(A_i), i \in I$, uma família qualquer de subconjuntos abertos de E e designemos por S a sua união (isto é, $S = \bigcup_{i \in I} A_i$). Se x for um elemento qualquer de S, então pertence, pelo menos, a um dos conjuntos da família, por exemplo $A_k (k \in I)$. Mas como A_k é, por hipótese, aberto, existe um $\delta > 0$ tal que $B(x, \delta) \subset A_k \subset S$ e, portanto, S é aberto.

P3. *A intersecção de um número finito de subconjuntos abertos de E é um subconjunto aberto de E.*

Demonstração. Com efeito, seja $A_1, A_2, \ldots, A_n$ uma colecção finita de subconjuntos abertos de E e designemos por S a sua intersecção $S = \bigcap_{i=1}^{n} A_i$. Se a intersecção for vazia, nada há a demonstrar, visto que $\emptyset$ é aberto. Suponhamos, porém, que assim não é e seja x um elemento arbitrário de S. Se $x \in S$, então $x \in A_i$ para cada $i = 1, 2, 3, \ldots, n$, isto é, x é elemento comum a todos os conjuntos. Mas, como todos eles são abertos, existem $\epsilon_1, \epsilon_2, \epsilon_3, \ldots, \epsilon_n$ tais que

$$\begin{array}{ccc} B(x, \epsilon_1) & \subset & A_1 \\ B(x, \epsilon_2) & \subset & A_2 \\ \ldots\ldots & \ldots & \ldots \\ B(x, \epsilon_n) & \subset & A_n. \end{array}$$

Designando por $\epsilon = \min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$ então $B(x, \epsilon)$ estará contida em todos os conjuntos (isto é, $B(x, \epsilon) \subset A_1 \cap A_2 \cap \ldots \cap A_n = S$), o que significa que S é aberto. ∎

Nota. Em geral, uma intersecção de infinitos conjuntos abertos não é aberta. Por exemplo a intersecção dos intervalos $]-1/n, 1/n[$ de $\mathbb{R}$ é o conjunto formado por um único ponto $\{0\}$ que não é aberto(vd. Proposição 1.6).

Observação. *A família dos conjuntos abertos de um espaço métrico denomina-se frequentemente a* ***topologia*** *desse espaço.*

Proposição 1.3 (**Proposição da separação de Hausdorff**). *Dados dois pontos distintos a e b dum espaço métrico (E, d) existem dois subconjuntos abertos e disjuntos que contém a e b respectivamente.*

Demonstração. Seja $d(a, b)$ a distância entre os dois pontos e consideremos as bolas abertas com centro em a e b e raio $d(a, b)/2$. Se estas bolas tivessem um ponto comum $x \in E$, seria

$$d(a, x) < \frac{d(a, b)}{2};$$

$$d(b, x) < \frac{d(a, b)}{2}$$

e portanto, pela desigualdade triangular , resultaria

$$d(a, b) \leq d(a, x) + d(x, b) < d(a, b)$$

o que é absurdo. Por consequência, as bolas abertas são distintas. ■

Um espaço que goza desta propriedade (Proposição 1.3) denomina-se **Espaço separável** ou de **Hausdorff**. Todo o espaço métrico é, portanto, espaço de Hausdorff.

1.3.2 Partes fechadas

Definição 1.11 (Subconjunto fechado).

> *Um subconjunto A de um espaço métrico (E, d) diz-se fechado quando o seu complementar $E - A$ é aberto.*

De acordo com a definição, o conjunto vazio e o espaço total E são ambos conjuntos fechados.

Proposição 1.4. *Uma bola fechada é um conjunto fechado. Uma esfera é um conjunto fechado.*

Demonstração. Com feito, se $x \notin B'(a, r)$, então

$$d(x, B'(a, r)) \geq d(a, x) - r > 0$$

e, portanto, a bola aberta de centro em x e raio $d(a, x) - r$ está no complementar de $B'(a, r)$, o que demonstra que esse complementar é aberto (pois é união de bola abertas).

O complementar da esfera $S(a, r)$ é a reunião da bola $B(a, r)$ e do complementar de $B'(a, r)$ e, portanto, é aberta. ■

Proposição 1.5. *Os subconjuntos fechados de um espaço métrico (E, d) gozam das seguintes propriedades:*

P1. *O conjunto vazio $\emptyset$ e o espaço inteiro E são fechados.*

> **Demonstração.** Na verdade, $\emptyset$ e E são complementares de E e $\emptyset$, respectivamente, logo são fechados;

P2. *A intersecção de uma família qualquer (finita ou infinita) de subconjuntos fechados de E é um subconjunto fechado de E;*

Demonstração. Seja $(A_i), i \in I$, uma família qualquer (finita ou infinita) de subconjuntos fechados de E e designemos por S a sua intersecção:

$$S = \bigcap_{i \in I} A_i.$$

De acordo com as leis de De Morgan, o complementar de S é

$$E - S = \bigcup_{i \in I} (E - A_i)$$

e, portanto, $E - S$ é aberto e S fechado.

P3. *A união de um número finito de subconjuntos fechados é um subconjunto fechado de E.*

Demonstração. Seja $A_1, A_2, \ldots, A_n$ uma colecção finita de subconjuntos fechados de E e designemos por S a sua união:

$$S = \bigcup_{i=1}^{n} A_i.$$

De acordo com as leis de De Morgan, o complementar de S é

$$E - S = \bigcap_{i \in I}^{n} (E - A_i).$$

Ora, como os A_i são fechados, os $E - A_i$ serão abertos e, portanto, $E - S$ é aberto por ser a intersecção dum número finito de conjuntos abertos. Mas, se $E - S$ é aberto, então S é fechado, como se pretendia provar. ■

Proposição 1.6. *Um conjunto formado por um só ponto é fechado. Geralizando tem-se que: todo o conjunto finito é fechado.*

Demonstração. Seja $a \in E$ e consideremos o conjunto $\{a\}$. Para provar que este conjunto é fechado, basta provar que o seu complementar $E - \{a\}$ é aberto. Com efeito, se

$$x \in E - \{a\} \quad \text{então} \quad d(a,x) > 0.$$

Mas então a bola aberta $B(x,\epsilon)$ com $\epsilon = d(a,x)$, não contém o ponto a e portanto $B(x,\epsilon) \subset E - \{a\}$. Mais geralmente, o complementar $E - F$ de qualquer subconjunto finito $F = \{a_1, a_2 \ldots, a_n\}$ de E é um subconjunto aberto de E, pois se $x \in (E - F)$ então o número

$$\epsilon = \min\{d(a_1,x), \ldots, d(a_n,x)\}$$

é maior que 0 e a bola aberta $B(x,\epsilon)$ não contém nenhum dos pontos $a_1, a_2, \ldots, a_n$, isto é, $B(x,\epsilon) \subset E - F$ e portanto F é fechado. ■

1.3.3 Vizinhanças

Definição 1.12 (Vizinhança).

> *Seja A um subconjunto não vazio de um espaço métrico (E,d). Chama-se vizinhança de A a qualquer subconjunto aberto de E que contenha A. Se A se reduz a um único ponto, $A = \{a\}$, as vizinhanças dizem-se então vizinhanças do ponto a.*

Assim, qualquer bola aberta com centro num ponto a é uma vizinhança desse ponto. Salvo indicação em contrário, serão deste tipo as vizinhanças que doravante se utilizarão:

(i) A bola aberta com centro no ponto a e raio $\epsilon > 0$ denomina-se vizinhança ϵ de a e representa-se por $V(a, \epsilon)$.

(ii) Se desta vizinhança se excluir o ponto a, então a vizinhança diz-se **vizinhança reduzida** de a e representa-se por $V'(a, \epsilon)$. E, como é evidente, $V'(a, \epsilon) = V(a, \epsilon) - \{a\}$.

1.3.4 Interior dum conjunto

Definição 1.13 (Ponto interior).

*Seja A um subconjunto dum espaço métrico (E, d). Um ponto x de E diz-se **ponto interior** a A se existir uma vizinhança de x inteiramente contida em A.*

O conjunto de todos os pontos interiores a A chama-se o **interior** de A e designa-se por $int(A)$.

O interior de um conjunto tem as seguintes propriedades básicas:

(i) O interior de um conjunto A é o maior conjunto aberto contido em A;

(ii) Um conjunto A é aberto se e só se coincide com o seu interior, isto é: $A = int(A)$

(iii) $int(A)$ é a união de todos os subconjuntos abertos de A.

1.3.5 Exterior de um conjunto

Definição 1.14.

> *Um ponto x diz-se **exterior** ao conjunto A quando for interior ao seu complementar $E - A$ (isto é, quando existir pelo menos uma vizinhança de x que está contida no complementar de A).*

Ao conjunto dos pontos exteriores a A chama-se o **exterior** de A e representa-se por $ext(A)$. Como resulta da própria definição, é $ext(A) = int(E - A)$ e portanto $ext(A)$ é aberto.

1.3.6 Fronteira dum conjunto

Definição 1.15 (Ponto fronteiro).

> *Seja A um subconjunto dum espaço métrico (E, d). Um ponto x de E diz-se que é **ponto fronteiro** a A, quando não é interior nem exterior a A. Equivale isto a dizer que em cada vizinhança de x existe pelo menos um ponto interior a A e um ponto exterior a A.*

O conjunto dos pontos fronteiros a A chama-se **fronteira de** A e representa-se por $front(A)$. Como é evidente, tem-se:

$$E = int(A) \cup ext(A) \cup front(A)$$

sendo os conjuntos disjuntos dois a dois.

Por outro lado, da definição resulta que:

$$E - front(A) = int(A) \cup ext(A)$$

e, portanto, $front(A)$ é fechado, visto o seu complementar ser aberto.

1.3.7 Aderência de um conjunto

Definição 1.16 (**Ponto aderente**).

> *Seja A um subconjunto dum espaço métrico (E, d). Diz-se que um ponto x de E é ponto aderente a A quando não é exterior a A, isto é, quando é interior ou fronteiro a A. Equivale isto a dizer que em cada vizinhança de x existe pelo menos um ponto de A.*

O conjunto de todos os pontos aderentes de A chama-se a **aderência** ou **fecho de** A e representa-se por $\overline{A}$. De acordo com a definição é

$$\overline{A} = A \cup front(A).$$

Da própria definição resulta também imediatamente que:

(a) A aderência de um conjunto A é o complementar do exterior de A;

(b) Um conjunto A é fechado se e só se coincide com seu fecho, isto é, $A = \overline{A}$.

1.3.8 Pontos de acumulação. Conjunto derivado

Definição 1.17 (**Ponto de acumulação**).

> *Seja A um subconjunto dum espaço métrico (E, d). Um ponto a de E diz-se* ***ponto de acumulação*** *do conjunto A se qualquer sua vizinhança contém uma infinidade de pontos de A.*

Um ponto de acumulação dum conjunto pode pertencer ou não a esse conjunto. Por exemplo, o conjunto dos números racionais do intervalo [1, 2] tem como pontos de acumulação não só todos os racionais desse intervalo (que pertencem ao conjunto), mas também todos os irracionais entre 1 e 2 (que não pertencem ao conjunto).

O conjunto de todos os pontos de acumulação de um conjunto A denomina-se **derivado de** A e representa-se por A'.

Definição 1.18 (**Ponto isolado**).

> *Um ponto x dum conjunto A diz-se* ***ponto isolado*** *de A se existe uma vizinhança de x que não contém nenhum ponto de A distinto do próprio x.*

Como é fácil de verificar:

Proposição 1.7. *Todo o ponto aderente a um conjunto A ou é ponto de acumulação ou ponto isolado de A.*

O fecho ou aderência dum conjunto A é, portanto, constituído em geral por três tipos de pontos:

- pontos isolados de A;
- pontos de acumulação de A pertencentes a A;
- pontos de acumulação de A não pertencentes a A.

Assim o fecho ou aderência de A é a união de A com o seu derivado, isto é: $\overline{A} = A \cup A'$. Portanto:

Proposição 1.8. *Um conjunto A é fechado se e só se contém todos os seus pontos de acumulação, isto é: $A' \subset A$.*

Um conjunto que é fechado e cujos pontos são todos de acumulação diz-se um **conjunto perfeito**.

1.3.9 Conjunto denso

Definição 1.19 (**Conjunto denso**).

> *Num espaço métrico (E, d) um conjunto A diz-se denso em relação a um conjunto B quando cada ponto de B é ponto aderente a A, isto é, quando para todo o $x \in B$ qualquer vizinhança de x contém pontos de A.*

Tem-se então $B \subset \overline{A}$ o que significa que A é denso em B se e só se todo o ponto de B é ponto de acumulação de A ou ponto de A (ou ambas as coisas).

Se um conjunto A é denso em relação a todo o espaço E, diz-se então que é **denso em todas as partes de** E ou simplesmente **denso em** E. Por exemplo, o conjunto $\mathbb{Q}$ dos números racionais é denso em $\mathbb{R}$.

1.4 Sucessões em espaços métricos

1.4.1 Convergência

Dados dois conjuntos quaisquer E e F chama-se **aplicação** (ou **função**) de E em F a toda correspondência que a cada elemento de E associa um e um só elemento de F.

Designado por uma letra, por exemplo f, a correspondência, escreve-se $f : E \to F$ que se lê "f é uma aplicação de E em F". O conjunto E é o domínio da aplicação f.

Se y for o elemento de F que corresponde ao elemento x de E através da aplicação f , escreve-se $y = f(x)$ e diz-se que y é a imagem de x através de f. Ao conjunto de todas as imagens através de f dos elementos de E chama-se o **contradomínio** da aplicação e representa-se por $f(E)$.

Em vez das expressões **aplicação** ou **função** de E em F usam-se também com o mesmo sentido as expressões **transformação** ou **operador**.

Um caso particular importante e que recebe designação especial é o das aplicações do conjunto $\mathbb{N}$ dos números naturais num outro conjunto qualquer E, isto é, o das aplicações $f : \mathbb{N} \to E$. Tais aplicações denominam-se **sucessões** e não são mais do que aplicações que a cada número natural n associam um elemento $f(n)$ de E que se representa geralmente por x_n (se for x a letra utilizada para designar elementos de E).

A sucessão $x_0, x_1, x_2, \dots, x_n, \dots$ representa-se abreviadamente por (x_n). Com n variável, x_n denomina-se o **termo geral** da sucessão.

Se o conjunto E estiver algebrizado podem nele definir-se operações com sucessões. Assim, por exemplo, se E for um corpo, dadas duas sucessões (x_n) e (y_n) de pontos de E, chama-se soma, produto e quociente dessas sucessões a sucessão (w_n) cujos termos têm, respectivamente, as expressões:

(a) $W_n = x_n + y_n$;

(b) $W_n = x_n \cdot y_n$;

(c) $Wn = \dfrac{x_n}{y_n}$ se $y_n \neq 0$ para todo n.

Definição 1.20 (Sucessão convergente).

> *Seja (x_n) uma sucessão de pontos do espaço métrico (E, d). Diz-se que a* ***sucessão é convergente*** *para o ponto $x \in E$ quando a todo o $\epsilon > 0$ corresponde um número natural n_0 tal que, para $n > n_0$, é $d(x_n, x) < \epsilon$.*

A definição equivale a afirmar que a todo $\epsilon > 0$ corresponde uma ordem n_0 a partir da qual os termos da sucessão estão contidos na vizinhança $V(x, \epsilon)$. Escreve-se então $\lim\limits_{n\to\infty} x_n = x$ ou $x_n \to x$ e diz-se que x é o **limite** da sucessão ao tender n para o infinito.

Quando não haja dúvida de que n tende para o infinito escreve-se simplesmente $\lim x_n = x$.

É claro que se (x_n) converge para x, então:

(a) é finito o número de termos da sucessão situados no exterior de qualquer vizinhança de x.

(b) se a sucessão tem uma infinidade de termos distintos, então o limite x é ponto de acumulação do conjunto dos seus termos.

A ordem n_0 a partir da qual se verifica a condição da definição $d(x_n, x) < \epsilon$ depende em geral do ϵ escolhido, facto que se costuma traduzir escrevendo $n_0 = n_0(\epsilon)$.

A convergência da sucessão para x pode exprimir-se simbolicamente da seguinte forma:

$$\forall \epsilon > 0,\ \exists n_0 = n_0(\epsilon) \in \mathbb{N} : \quad n > n_0 \Longrightarrow d(x_n, x) < \epsilon$$

ou

$$n > n_0 \Longrightarrow x_n \in V(x, \epsilon).$$

Na recta real $\mathbb{R}$, como $d(x, y) = |x - y|$, a definição toma a forma

$$\forall \epsilon > 0,\ \exists n_0(\epsilon) \in \mathbb{N} : \quad n > n_0 \Longrightarrow |x_n - x| < \epsilon.$$

Uma sucessão (x_n) diz-se **limitada** se o conjunto dos seus termos $\{x_0, x_1, x_2, \ldots\}$ é limitado.

Se (x_n) é uma sucessão qualquer e (n_k) é uma sucessão de inteiros positivos tais que $n_1 < n_2 < \ldots < n_k < \ldots$ então a sucessão (x_{n_k}) chama-se uma **subsucessão** da sucessão dada.

Proposição 1.9. *Num espaço métrico o limite de uma sucessão, se existe, é único.*

Demonstração. Suponha-se que assim não sucedia e que a sucessão (x_n) admitia dois limites x e x' tais que $d(x, x') = \epsilon$. Então, para esse $\epsilon > 0$, existiriam dois naturais $n_0(\epsilon)$ e $n_0'(\epsilon)$ tais que

$$\begin{aligned} n > n_0 &\implies d(x_n, x) > \frac{\epsilon}{2} \\ n > n_0' &\implies d(x_n, x') < \frac{\epsilon}{2}. \end{aligned}$$

Mas nessas condições, tomando $N = \max\{n_0, n_0'\}$, ter-se-ia, para $n > N$,

$$d(x, x') \leq d(x, x_n) + d(x_n, x') < \epsilon = d(x, x'),$$

o que é contraditório. Portanto deve ser $d(x, x') = 0$ o que significa que $x = x'$. ■

Proposição 1.10. *Toda a sucessão convergente de pontos dum espaço métrico* (E, d) *é limitada .*

Demonstração. Basta mostrar que o conjunto $\{x_0, x_1, x_2, \ldots\}$ dos termos da sucessão está contido numa bola de (E, d). Se $\lim x_n = x$, escolhido $\epsilon > 0$ existe n_0 tal que, para qualquer $n > n_0$, é $d(x, x_n) < \epsilon$. Fixado ϵ seja $r = \max\{d(x, x_0), d(x, x_1), \ldots, d(x, x_{n_0}), \epsilon\}$. Então o conjunto $\{x_0, x_1, x_2, \ldots\}$ está contido na bola fechada $B'(x, r)$ com centro em x e raio r, o que prova que é limitado. ■

Proposição 1.11. *Num espaço métrico, toda a subsucessão duma sucessão convergente converge para o mesmo limite.*

Demonstração. Se $\lim x_n = x$, dado $\epsilon > 0$ existe n_0 tal que $d(x, x_n) < \epsilon$ sempre que $n > n_0$. Como $n_{n_m} \geq m$ para todos os inteiros positivos m, tem-se: $d(x, x_{n_m}) < \epsilon$ para $m > n_0$ o que significa que $\lim x_{n_m} = x$ como se queria provar. ■

Proposição 1.12. *Se x é ponto de acumulação do subconjunto D do espaço métrico (E, d), é sempre possível construir uma sucessão de pontos distintos de D convergente para x.*

Demonstração. Se x é o ponto de acumulação de D, qualquer bola aberta com centro em x contém uma infinidade de pontos de D. Então, tomando o raio $r_0 = 1$ escolha-se na bola $B(x, r_0)$ um ponto de D, $x_0 \neq x$. Em seguida tome-se

$$r_1 = \frac{d(x, x_0)}{2} < \frac{1}{2}$$

e na bola $B(x, r_1)$ escolha-se um ponto $x_1 \neq x$. Continuando a proceder do mesmo modo, obtém-se na bola $B(x, r_n)$ com

$$r_n = \frac{d(x, x_{n-1})}{2} < \frac{1}{2^n}$$

um ponto $x_n \neq x$. Os pontos escolhidos pertencem todos a D e são distintos uns dos outros formando uma sucessão $x_0, x_1, \ldots, x_n, \ldots$. Como

$$r_n = d(x, x_{n-1}) < \frac{1}{2^n},$$

vem $d(x, x_0) < \dfrac{1}{2^n}$ o que mostra que a todo o $\epsilon > 0$ corresponde um n_0 de modo que $d(x, x_n) < \epsilon$ para $n > n_0$ (para o que, escolhido ϵ, basta fazer $\dfrac{1}{2^n} < \epsilon$, resultando daí n_0). Logo é $\lim x_n = x$, como se pretendia provar. ■

Proposição 1.13. *Um ponto x é ponto de acumulação dum subconjunto D de um espaço métrico (E, d) se e só se existe uma sucessão de pontos distintos de D convergente para x.*

Demonstração. De facto, se x é ponto de acumulação de D, nos termos da proposição anterior é sempre possível construir uma sucessão (x_n) de pontos distintos de D convergente para x. Reciprocamente, se x é limite duma sucessão de pontos distintos de D, a todo o $\epsilon > 0$ corresponde um $n_0(\epsilon)$ tal que é $d(x, x_n) < \epsilon$ para $n > n_0$, e, portanto, qualquer bola aberta com centro em x contém uma infinidade de pontos de D, o que significa que x é ponto de acumulação de D. ∎

Proposição 1.14. *O limite duma sucessão com uma infinidade de termos distintos é único ponto de acumulação do conjunto dos seus termos.*

Demonstração. De facto, como se verificou já, se a sucessão tem uma infinidade de termos distintos e é convergente, o limite é ponto de acumulação do conjunto dos termos da sucessão. E não pode haver outro, pois que se houvesse podia extrair-se da sucessão uma subsucessão convergente para ele, que teria, portanto, limite distinto do da sucessão, o que não é possível nos termos da Proposição 1.11. ∎

Observação. Da proposição anterior decorre que o limite de uma sucessão convergente ou é ponto de acumulação do conjunto dos seus termos ou termo indefinidamente repetido na sucessão.

1.4.2 Sucessões de Cauchy. Espaços completos

Definição 1.21 (Sucessão de Cauchy).

> *Uma sucessão (x_n) de pontos de um espaço métrico (E, d) diz-se uma sucessão de Cauchy se a todo o $\epsilon > 0$ corresponde um número natural n_0 tal que $d(x_p, x_q) < \epsilon$ sempre que $p, q > n_0$.*

Portanto, toda a **sucessão convergente** é **sucessão de Cauchy**. De facto, se (x_n) converge para x, então para todo o $\epsilon > 0$ existe um natural n_0 tal que para $n > n_0$ é

$$d(x_n, x) < \frac{\epsilon}{2}.$$

Para quaisquer $p, q > n_0$ tem-se portanto

$$d(x_p, x_q) \leq d(x_p, x) + d(x, x_q) < \epsilon$$

o que mostra que a sucessão é, de facto, uma sucessão de Cauchy.

O recíproco, porém, é falso, isto é, **nem todas as sucessões de Cauchy são convergentes**. Por exemplo, no subespaço $A =]0, 1]$ da recta real a sucessão de termo geral $x_n = \dfrac{1}{n}$ é sucessão de Cauchy mas não é convergente já que 0 não pertence a A.

Como é evidente:

Proposição 1.15. *Toda a subsucessão de uma sucessão de Cauchy é também uma sucessão de Cauchy.*

Proposição 1.16. *Toda a sucessão de Cauchy é limitada.*

Se (x_n) é sucessão de Cauchy, para cada $\epsilon > 0$ existe um n_0 tal que $d(x_p, x_q) < \epsilon$ para quaisquer $p, q > n_0$. Então, fixado um $q > n_0$, todos os termos da sucessão ficarão contidos na bola fechada com centro em x_q e raio $r = \max\{d(x_0, x_q), d(x_1, x_q), \ldots, d(x_n, x_q), \epsilon\}$.

Proposição 1.17. *É convergente toda a sucessão de Cauchy que tenha uma subsucessão convergente.*

Demonstração. Sejam (x_n) uma sucessão de Cauchy do espaço métrico (E, d) e (x_{n_k}) uma sua subsucessão para o ponto $x \in E$. Como a subsucessão converge para x, dado $\epsilon > 0$ existe n_0' tal que para $n_k > n_0'$ é

$$d(x_{n_k,x}) < \frac{\epsilon}{2}.$$

Mas como a sucessão dada é sucessão de Cauchy, ao mesmo ϵ escolhido corresponde um n_0'' tal que é

$$d(x_p, x_q) < \frac{\epsilon}{2}$$

sempre que $p, q > n_0''$. Então, se for $n_0 = \max\{n_0', n_0''\}$, para $n > n_0$ e $n_k > n_0$ tem-se

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

o que significa que a sucessão (x_n) converge também para x. ■

Definição 1.22 (Espaço completo).

> *Um espaço métrico (E, d) diz-se um **espaço métrico completo** quando toda a sucessão de Cauchy de pontos de E é convergente em E.*

Das proposições 1.12, 1.6 e 1.14 resulta que num espaço métrico E há identidade entre os subconjuntos fechados de E e os subespaços completos de E.

Teorema 1.1. *A recta real $\mathbb{R}$ é um espaço métrico completo.*

Demonstração. Seja x_n uma sucessão de Cauchy de números reais e A o conjunto definido pela forma:

$$A = \{a \in \mathbb{R} : a \leq x_n, \quad \text{para um número infinito de naturais } n\}.$$

Como toda a sucessão de Cauchy é limitada, o conjunto A é limitado superiormente e, portanto, admite supremo. Seja $x = \sup A$. Como (x_n) é sucessão de Cauchy, para qualquer $\epsilon > 0$ existe um natural n_0 tal que para $p, q > n_0$ é

$$d(x_p, x_q) < \frac{\epsilon}{2}.$$

Para o mesmo ϵ tem-se então, de acordo com a definição de supremo, que

$$x - \frac{\epsilon}{2} \in A \quad \text{mas} \quad x + \frac{\epsilon}{2} \notin A,$$

o que significa que há uma infinidade de naturais n para os quais $x - \frac{\epsilon}{2} \leq x_n$ e apenas um número finito para os quais $x + \frac{\epsilon}{2} \leq x_n$.

Pode portanto determinar-se um natural $q > n_0$ tal que seja

$$x - \frac{\epsilon}{2} \leq x_q \quad \text{e} \quad x + \frac{\epsilon}{2} > x_q$$

e por consequência

$$|x - x_q| \leq \frac{\epsilon}{2}.$$

Então para $p > n_0$ tem-se

$$|x_p - x| = |(x_p - x_q) + (x_q - x)| \leq |x_p - x_q| + |x_q - x| < \epsilon$$

o que prova que a sucessão (x_n) converge para x. Portanto em $\mathbb{R}$ toda a sucessão de Cauchy é convergente e, por consequência, $\mathbb{R}$ é completo. $\blacksquare$

A importância dos espaços completos assenta no facto de que neles, para se provar que uma sucessão é convergente, basta provar que ela é sucessão de Cauchy (ou, como se costuma dizer, que verifica o **Critério de Cauchy**) sem necessidade de se determinar o seu limite.

1.5 Exercícios

1. Deduza a desigualdade de Cauchy

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \sum_{k=1}^{n} a_k^2 \cdot \sum_{k=1}^{n} b_k^2.$$

2. Deduza a desigualdade de Minkowski

$$\sqrt{\sum_{k=1}^{n}(a_k + b_k)^2} \leq \sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}.$$

3. Seja $C[a,b]$ o conjunto de todas as funções reais contínuas no intervalo $[a,b]$. Prove que $d(f,g) = \sup_{x\in[a,b]} |f(x) - g(x)|$ é métrica nesse conjunto.

4. Seja $C[0,1]$ o conjunto das funções reais contínuas no intervalo $[0,1]$. Prove que $d(f,g) = \int_0^1 |f(x) - g(x)| dx$ é métrica nesse conjunto.

5. Dados os conjuntos $A = [0,1]$ e $B =]1,2]$ de $\mathbb{R}$, determine a distância entre esses conjuntos quando $\mathbb{R}$ está munido:
   (a) Da métrica usual $d(x,y) = |x - y|$;
   (b) Da métrica discreta.

6. Seja (E,d) um espaço discreto e A um subconjunto não vazio de E. Determine a distância de um ponto qualquer x de E ao conjunto A.

7. Determine os pontos interiores, os pontos exteriores e os pontos fronteiros do conjunto dos irracionais do intervalo $]1,5[$.

8. Prove que o interior de um subconjunto aberto A é o maior conjunto aberto contido em A.

9. Prove que se $A \subset B$ também $int(A) \subset int(B)$.

10. Determine o interior, o exterior e a fronteira dos seguintes conjuntos:

    (a) O intervalo $I = [a, b]$ de $\mathbb{R}$;

    (b) O conjunto $\mathbb{Q}$ dos racionais.

11. Prove que $\mathbb{R} - \overline{A} = int(\mathbb{R} - A)$.

12. Determine o derivado dos seguintes subconjuntos de $\mathbb{R}$:

    (a) $]0, 1]$;

    (b) Conjunto $\mathbb{Z}$ dos inteiros.

13. Determine o fecho ou aderência dos seguintes subconjuntos da recta real

    (a) $\mathbb{Q}$;

    (b) $]0, +\infty[$;

    (c) $]-1, 0[\cup]0, 1[$.

14. Sejam (x_n) e (y_n) sucessões do espaço métrico (E, d) convergentes para x e y respectivamente. Prove que

$$d(x_n, y_n) \to d(x, y).$$

15. Prove que qualquer aplicação dum espaço discreto noutro espaço métrico é contínua.

16. Prove que a união de um conjunto finito de subconjuntos compactos de um espaço métrico é compacta.

17. Prove que uma aplicação f de um espaço métrico (E, d_1) noutro espaço métrico (F, d_2) é contínua se e só se a imagem inversa por f de qualquer fechado de F é um fechado de E.

18. Prove que a aplicação identidade $i : x \rightarrow x$ de um espaço métrico (E, d) sobre si mesmo é contínua.

19. Prove que um subconjunto A da recta real $\mathbb{R}$ é conexo se e só se contém qualquer número real c compreendido entre dois quaisquer dos seus elementos, isto é: se $a \in A$, $b \in A$, $c \in \mathbb{R}$ e $a < c < b$ então $c \in A$.

20. Mostre que se todo o subconjunto infinito dum conjunto A tem pelo menos um ponto de acumulação em A, então A é fechado e limitado.

Capítulo 2

Espaços vectoriais normados

2.1 Espaços vectoriais

2.1.1 Definição de espaço vectorial

Definição 2.1.

Dados um conjunto E de elementos u, v, $z, \ldots$ quaisquer e um corpo comutativo K de elementos $\alpha, \beta, \ldots$ diz-se que o conjunto E é um espaço vectorial sobre o corpo K, quando se verificam as seguintes condições:

(i) Está definida em E uma adição que a cada par (u, v) de elementos de E faz corresponder um e um só elemento de E, que se chama soma de u com v e se representa por $u + v$, de tal modo que a operação assim definida confere a E estrutura de grupo comutativo;

(ii) A cada par (α, u), formado por um elemento α de

K e um elemento u de E, corresponde um determinado elemento de E, que se chama produto escalar de α por u e se representa por αu, de tal modo que para quaisquer $u, v \in E$ e $\alpha, \beta \in K$:

1. $(\alpha + \beta)u = \alpha u + \beta u$
   (distributividade para os elementos dos corpos K).
2. $\alpha(u + v) = \alpha u + \alpha v$
   (distributividade para os elementos de E).
3. $\alpha(\beta u) = (\alpha\beta)u$
   (Associatividade)
4. $1 \cdot u = u$.

Nestas condições os elementos de E chamam-se **vectores** e os de K **escalares**.

Nas aplicações à Análise só interessa, geralmente, o caso em que K é o corpo $\mathbb{R}$ ou o corpo complexo $\mathbb{C}$. No primeiro caso diz-se que E é um **espaço vectorial real** e, no segundo caso, um **espaço vectorial complexo** . Só destes dois casos se tratará em tudo o que se segue, salvo indicação expressa em contrário.

Exemplo 2.1. A recta real $\mathbb{R}$, com a adição e multiplicação usuais de números reais é um **espaço vectorial** sobre si mesmo. Na realidade é mais do que isso visto ser um corpo. ◀

Exemplo 2.2. Seja E o conjunto dos polinómios $P(x)$ com coeficientes reais definidos no intervalo $[0, 1]$, incluindo os polinómios de grau zero e o polinómio identicamente nulo. Com as operações usuais de adição de polinómios e de multiplicação de um polinómio por um numero real, E é um **espaço vectorial real**. ◀

Exemplo 2.3. Se em $\mathbb{R}^2$ se definir soma de dois elementos $x = (x_1, x_2)$ e $y = (y_1, y_2)$ como sendo o elemento de $\mathbb{R}^2$

$$x + y = (x_1 + y_1, x_2 + y_2)$$

e produto de um número real λ por x como sendo o elemento de $\mathbb{R}^2$

$$\lambda x = (\lambda x_1, \lambda x_2)$$

é fácil verificar que as operações assim definidas conferem a $\mathbb{R}^2$ estrutura de espaço vectorial real (se $\lambda \in \mathbb{C}$ o espaço será um **espaço vectorial complexo**). ◀

Exemplo 2.4. Designando por $\mathbb{R}^\infty$ o conjunto de todas as sucessões de números reais

$$(x_1, x_2, \ldots, x_n, \ldots)$$

e definindo neste conjunto, de forma análoga à do exemplo anterior, operações de adição e de multiplicação por escalar, o conjunto $\mathbb{R}^\infty$ converte-se num espaço vectorial real. ◀

Exemplo 2.5. Dado um intervalo I qualquer de $\mathbb{R}$, seja $\mathcal{F}(I, \mathbb{R})$ o conjunto de todas as funções reais definidas em I. Definindo soma $f + g$ de duas funções $f, g \in \mathcal{F}(I, \mathbb{R})$ e produto λf de um número $\lambda \in \mathbb{R}$ por uma função $f \in \mathcal{F}(I, \mathbb{R})$ pela forma usual, isto é

$$(f + g)(x) = f(x) + g(x), \quad \forall x \in I$$

$$(\lambda f)(x) = \lambda f(x), \quad \forall x \in I$$

facilmente se reconhece que $\mathcal{F}(I, \mathbb{R})$ é um espaço vectorial real. ◀

Exemplo 2.6. De forma análoga são espaços vectoriais reais os conjuntos $\Re(X, \mathbb{R})$ das funções reais limitadas definidas num conjunto X e o conjunto $C[a, b]$ das funções reais contínuas num intervalo $[a, b]$ de $\mathbb{R}$. ◄

Nesta ordem de ideias, pode-se chamar vector a cada função f e **coordenadas** ou **componentes** do vector f aos valores $f(x)$ que f toma nos diferentes pontos x do seu domínio.

2.1.2 Independência linear

Se $\{x_1, x_2, \ldots, x_n\}$ é um subconjunto finito de vectores do espaço vectorial E sobre o corpo K, chama-se **combinação linear** desse vector a qualquer vector da forma

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$$

em que $\lambda_1, \lambda_1, \ldots, \lambda_n$ são escalares pertencentes ao corpo K que se denominam **coeficientes da combinação**.

Um conjunto de vectores $\{x_1, x_2, \ldots, x_n\}$ diz-se **linearmente independente** se qualquer combinação linear de coeficientes não todos nulos é diferente de zero, isto é, se qualquer relação da forma

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = 0$$

implica

$$\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0.$$

Se assim não acontecer, diz-se que os vectores são **linearmente dependentes** sobre K (ou que formam um sistema ligado). Significa isso que há pelo menos uma combinação linear desses vectores, de coeficientes não todos nulos, que é nula.

2.1.3 Subespaços vectoriais

Chama-se subespaço dum espaço vectorial E a qualquer subconjunto não vazio E' de E que seja ele próprio um espaço vectorial relativamente ás operações definidas sobre E.

Como é evidente, um subconjunto não vazio E' de um espaço vectorial E é um subespaço vectorial de E se e só se for fechado em relação à adição e à multiplicação escalar, isto é, se contém qualquer combinação linear dos seus vectores: $\lambda_1 x + \lambda_2 y \in E'$ sempre que $x, y \in E'$ e $\lambda_1, \lambda_2 \in K$.

Se o subespaço E' for um subconjunto próprio do espaço E diz-se que E' é um **subespaço próprio** de E. O espaço inteiro E e o subconjunto $\{0\}$ constituído pelo vector nulo são sempre subespaços de E, denominados **subespaços impróprios**.

Como resulta da própria definição, se S for um subconjunto não vazio do espaço vectorial E, o conjunto de todas as combinações lineares dos vectores de S é um subconjunto de E, que contém S e se diz gerado por S.

Os vectores de S denominam-se então os geradores do subespaço e este representa-se por $\{S\}$.

2.1.4 Base e dimensão dum espaço vectorial

Chama-se **base** dum espaço vectorial E a qualquer conjunto B de vectores linearmente independentes de E que gera E, isto é, tal que $E = \{B\}$.

Assim, um subconjunto $B = \{x_1, x_2, \ldots, x_n\}$ dum espaço vectorial E é uma base de E se e só se qualquer vector $x \in E$ se pode exprimir de um só modo como combinação linear dos x_i e estes são linearmente independentes.

O número de vectores de uma base pode ser finito ou infinito.

Se um espaço vectorial E admite, pelo menos, um sistema finito de geradores linearmente independentes, tal sistema constitui uma **base finita** do espaço E.

Os vectores que constituem uma base representam-se, geralmente, pela letra e afectada de índices: e_i.

Como facilmente se pode comprovar, se um espaço E tem uma base finita $B = \{e_1, e_2, \ldots, e_n\}$ com n elementos, então qualquer outra base de E tem o mesmo número de n elementos.

Ao número de vectores de qualquer base do espaço vectorial E chama-se a **dimensão** desse espaço.

Se um espaço E admite uma base constituída por n vectores então diz-se que E é um **espaço de dimensão finita** n e escreve-se $dim(E) = n$. Caso contrário, o espaço diz-se de **dimensão infinita**.

Da definição resulta que num espaço de dimensão n qualquer sistemas de $n+1$ vectores é necessariamente dependente e só existem sistemas linearmente independentes de k vectores se for $k \leq n$.

Claro que qualquer subespaço W dum espaço E de dimensão n tem dimensão não superior a n, isto é, $dim(W) \leq n$.

Se, em particular, for $dim(W) = n$ então necessariamente $W = E$.

2.1.5 Isomorfismos

Dois espaços vectoriais E e F sobre o mesmo corpo K dizem-se isomorfos se existe uma bijecção f de E sobre F que respeite as operações desses espaços, isto é, tal que:

$$\begin{aligned} f(x+y) &= f(x)+f(y) \\ f(\lambda x) &= \lambda f(x). \end{aligned}$$

2.1.6 Conjuntos convexos. Funções convexas

Se E é um espaço vectorial real e x e y são dois pontos de E, chama-se segmento fechado com extremidades nesses pontos, e representa-se por $[x, y]$, ao conjunto dos pontos p de E tais que

$$p = \lambda x + (1-\lambda)y, \quad 0 \leq \lambda \leq 1.$$

De forma análoga se define segmento aberto $]x, y[$ de extremidades x e y, tomando $0 < \lambda < 1$.

Definição 2.2 (**Conjunto denso**).

> *Um subconjunto A dum espaço métrico vectorial E diz-se conjunto convexo se contém o segmento fechado que liga dois quaisquer dos seus pontos.*

Portanto o subconjunto A de E é convexo se e só se, sempre que $x, y \in A$, também o segmento $[x, y]$ está contido em A.
Como é evidente, se E for convexo todo o subespaço vectorial de E

é convexo e em $\mathbb{R}$ os únicos subconjuntos convexos não vazios são os intervalos.

Um conjunto convexo A diz-se um **cone convexo** se, para todo o x de A e todo $\lambda \geq 0$, também $\lambda x \in A$.

Se A_1 e A_2 são conjuntos convexos e x e y são pontos comuns a ambos, então o segmento que liga x e y está necessariamente contido tanto em A_1 como em A_2.

Portanto, a **intersecção de conjuntos convexos é um conjunto convexo** (a união em geral não é.)

Definição 2.3.

Seja A um subconjunto convexo do espaço vectorial real E. Uma função $f : A \to \mathbb{R}$ diz-se convexa se para quaisquer pontos $x, y \in A$ e $0 \leq \lambda \leq 1$ é

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

A função é **estritamente convexa** se a desigualdade estrita se verificar para $x \neq y$. A função $f(x)$ diz-se côncava (ou estritamente côncava) quando $(-f(x))$ é convexa (resp. estritamente convexa).

Da definição deduz-se sem dificuldade que a soma de funções convexas é uma função convexa. Por outro lado, se $c > 0$ e $f(x)$ é convexa, $cf(x)$ é obviamente convexa. Portanto: **qualquer combinação linear de funções convexas com coeficientes não-negativos é também uma função convexa.**

2.1.7 Transformações lineares

Definição 2.4 (Transformação linear).

> *Sejam E e F dois espaços vectoriais sobre o mesmo corpo K. Chama-se transformação (ou aplicação) linear de E em F a toda a aplicação $T : E \to F$ que satisfaz as condições*
>
> $$\begin{aligned} T(x+y) &= T(x)+T(y) \\ T(\lambda x) &= \lambda T(x) \end{aligned}$$
>
> *para quaisquer $x, y \in E$ e $\lambda \in K$.*

De acordo com a definição, as transformações lineares são, pois, aplicações que respeitam as operações dos espaços vectoriais envolvidos. As duas condições da definição podem, como é evidente, substituir-se pela condição equivalente

$$T(\lambda_1 x + \lambda_2 y) = \lambda_1 T(x) + \lambda_2 T(y)$$

que é frequentemente utilizada para caracterizar as transformações lineares. A sua aplicação repetida conduz à relação

$$T\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i T(x_i)$$

que traduz a propriedade fundamental das transformações lineares.

Exemplo 2.7. Seja E o espaço vectorial real dos polinómios $P(x)$ com coeficientes reais definidos em $[0, 1]$. A aplicação $D : E \to E$ definida pela forma

$$D(P) = \frac{dP}{dx}$$

é uma transformação linear de E sobre si mesmo visto que

$$\frac{d}{dx}(\lambda_1 P + \lambda_2 Q) = \lambda_1 \frac{dP}{dx} + \lambda_2 \frac{dQ}{dx}. \quad \blacktriangleleft$$

Exemplo 2.8. No espaço vectorial real $C[a,b]$ a aplicação $I : C[a,b] \to \mathbb{R}$ definida pela forma

$$I(f) = \int_a^b f(x)dx$$

é uma transformação linear de $C[a,b]$ em $\mathbb{R}$ visto que

$$\int_a^b [\lambda_1 f(x) + \lambda_2 g(x)]dx = \lambda_1 \int_a^b f(x)dx + \lambda_2 \int_a^b g(x)dx. \quad \blacktriangleleft$$

As transformações lineares, além de preservarem as operações, preservam também a origem e os negativos pois que

$$\begin{aligned} T(0) &= T(0 \cdot 0) = 0 \cdot T(0) = 0 \\ T(-x) &= T((-1)x) = (-1)T(x) = -T(x). \end{aligned}$$

Assim, numa transformação linear $T : E \to F$ a imagem (ou transformada) do vector zero de E é o vector zero de F e a imagem (ou transformada) do oposto dum vector de E é o oposto da imagem desse vector em F.

Proposição 2.1. *Numa transformação linear $T : E \to F$, para se obter a imagem de qualquer vector de E, basta conhecer as imagens dos vectores de uma base de E.*

Seja $\{e_1, e_2, \ldots, e_n\}$ uma base do espaço E. Então, para qualquer vector $x \in E$ existem escalares $\lambda_i \in K$ tais que

$$x = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n.$$

Como T é uma transformação linear tem-se

$$T(x) = \lambda_1 T(e_1) + \lambda_2 T(e_2) + \ldots + \lambda_n T(e_n)$$

o que mostra que, de facto, para se obter a imagem de $x \in E$ basta conhecer as imagens dos vectores de uma base de E. Quer dizer: uma transformação linear $T : E \to F$ fica completamente determinada pela maneira como transforma uma base de E.

Seja $\Im(E, F)$ o conjunto de todas as transformações lineares do espaço vectorial E no espaço vectorial F, ambos sobre o mesmo corpo K.

Sobre $\Im(E, F)$ podem definir-se, de maneira natural, uma operação de adição de duas transformações lineares e uma operação de multiplicação de uma transformação linear por um escalar pondo para quaisquer $T, U \in \Im(E, F)$ e $\lambda \in K$.

$$\begin{gathered} (T + U)(x) = T(x) + U(x) \\ (\lambda T)(x) = \lambda T(x). \end{gathered} \tag{2.1}$$

Como é fácil de verificar, $T + U$ e λT são também transformações lineares de E em F e as operações assim definidas conferem a $\Im(E, F)$ estrutura de espaço vectorial sobre o corpo K. Portanto:

Proposição 2.2. *As operações de adição de transformações lineares e de multiplicação de um escalar por uma transformação linear definidas pelas equações (2.1) conferem ao conjunto $\Im(E, F)$ das transformações lineares de E em F, ambos sobre o mesmo corpo K, estrutura de espaço vectorial sobre K.*

Um caso particular importante é aquele em que $F = K$, isto é, o caso das transformações lineares dum espaço E no corpo K dos escalares, as quais se designam formas lineares.

Definição 2.5 (**Forma linear**).

> *Chama-se **forma linear** a toda a transformação linear de um espaço vectorial no corpo do seus escalares.*

De acordo com a Proposição 2.2 o conjunto $\Im(E, F)$ de todas as formas lineares definidas sobre um espaço vectorial E é, ele próprio, um espaço vectorial, que se denomina **espaço dual de** E e se representa por E^*.

2.2 Espaços normados

2.2.1 Definição

Definição 2.6 (**Norma**).

> *Dado um espaço vectorial E (real ou complexo), chamase **norma** em E a uma aplicação, geralmente notada $x \to ||x||$, de E no conjunto $\mathbb{R}$ dos números reais, que verifica os seguintes axiomas:*

(i) $||x|| > 0$ *se* $x \neq 0$*;* $||0|| = 0$*;*

(ii) $||\lambda x|| = |y| \cdot ||x||$ *para todo o escalar* λ*;*

(iii) $||x + y|| \leq ||x|| + ||y||$.

A cada vector x do espaço vectorial E fica, assim, associado um número real não negativo, que se chama a **norma de** x e se representa por $||x||$. A aplicação $x \to ||x||$ pode também representar-se abreviadamente por $|| \cdot ||$.

Um espaço vectorial no qual esteja definida uma norma denomina-se **espaço vectorial normado** ou simplesmente **espaço normado**, isto é:

Definição 2.7 (**Espaço vectoria normado**).

Chama-se ***espaço vectorial normado*** *ao par* $(E, ||\cdot||)$ *constituído por um espaço vectorial* E *e uma norma* $||\cdot||$ *definida sobre* E.

Caso não haja dúvidas quanto à norma considerada, o espaço designa-se simplesmente por E.

Exemplo 2.9. Seja $\Re(X, \mathbb{R})$ o espaço vectorial das funções reais limitadas definidas sobre X. Neste espaço a aplicação definida pela forma

$$||f|| = \sup_{x \in X} |f(x)|$$

é uma norma, que se denomina **norma de convergência uniforme**. ◀

Exemplo 2.10. No espaço vectorial $C[a,b]$ das funções reais contínuas no intervalo $[a,b]$, além da norma do exemplo anterior, pode também definir-se a norma

$$||f|| = \int_a^b |f(x)|dx. \quad \blacktriangleleft$$

Exemplo 2.11. O espaço vectorial real $\mathbb{R}$ é naturalmente um espaço normado no qual a norma de um número x se define por $||x|| = |x|$. ◄

Proposição 2.3. *Num espaço vectorial normado $(E, ||\cdot||)$ a aplicação d definida pela forma*

$$d(x,y) = ||x-y||$$

para quaisquer $x, y \in E$, é uma métrica, denominada métrica induzida pela norma.

Demonstração. Com efeito, das propriedades que caracterizam uma norma, resulta para a função $d(x,y) = ||x-y||$:

(a) $d(x,y) = ||x-y|| > 0$ se $x \neq y$ e $d(x,x)+||x-x|| = ||0|| = 0$;

(b) $d(x,y) = ||x-y|| = ||y-x|| = d(y,x)$;

(c) $d(x,z) = ||x-z|| = ||(x-y)+(y-z)|| \leq$

$$\leq ||x-y|| + ||y-z|| = d(x,y) + d(y,z),$$

o que mostra que a aplicação $d : E \times E \to \mathbb{R}$ definida pela forma $d(x,y) = ||x-y||$ satisfaz os axiomas de uma métrica em E.

Por consequência: **todo o espaço normado E se pode converter num espaço métrico desde que se defina a distância entre dois elementos (vectores) de E por meio da igualdade**

$$d(x, y) = ||x - y||. \blacksquare$$

Por este motivo, todo o espaço normado será doravante sempre considerado um espaço métrico, aplicando-se-lhe por consequência toda a teoria exposta no Capítulo 1.

2.2.2 Espaços de Banach

Num espaço normado $(E, ||\cdot||)$ uma sucessão (x_n) de elementos de E diz-se sucessão de Cauchy se a todo o $\epsilon > 0$ corresponde um natural $n_0(\epsilon)$ tal que para quaisquer $p, q > n_0$ é

$$d(x_p, x_q) = ||x_p - x_q|| < \epsilon.$$

Tal como se viu no número 1.4.2 do Capítulo 1, toda a sucessão convergente num espaço normado E é uma sucessão de Cauchy, mas o recíproco é, em geral, inexacto.

Se, porém, acontecer que num dado espaço normado E toda a sucessão de Cauchy seja convergente, o espaço E diz-se então um **espaço de Banach**.

Definição 2.8 (Espaço de Banach).

> *Chama-se espaço de Banach a todo o espaço normado que é completo.*

Assim, $\mathbb{R}$ é um espaço de Banach; $\mathbb{Q}$, porém, não é.

2.2.3 Normas equivalentes

Definição 2.9 (Normas equivalentes).

Duas normas $x \to ||x||_1$ e $x \to ||x||_2$ sobre um espaço vectorial E dizem-se equivalentes quando existem números reais positivos h e k tais que, para qualquer $x \in E$, seja

$$||x||_1 \leq h||x||_2 \quad e \quad ||x||_2 \leq k||x||_1.$$

O interesse da noção de norma equivalente reside no facto de que certas propriedades dum espaço normado não se alteram quando se substitui a norma utilizada por outra equivalente.

As propriedades nessas condições, isto é, que não se alteram quando se substitui uma norma por outra que lhe seja equivalente, chamam-se **propriedades topológicas** dum espaço normado. No caso de espaços vectoriais de dimensão finita pode-se demonstrar a seguinte proposição:

Proposição 2.4. *Num espaço vectorial de dimensão finita todas as normas são equivalentes.*

2.2.4 Produto interno

É muitas vezes desejável introduzir num espaço vectorial uma outra operação, denominada **produto interno**, definida do seguinte modo:

Definição 2.10 (Produto interno).

Dado um espaço vectorial E, chama-se produto interno sobre E a uma aplicação de $E \times E$ em $\mathbb{R}$ que a cada par *de vectores x e y de E associa um número real, geralmente notado $x \cdot y$ e que satisfaz os seguintes axiomas:*

(i) ***Positividade:*** $x \cdot x > 0$, *excepto se* $x = 0$*;*

(ii) ***Simetria:*** $x \cdot y = y \cdot x$*;*

(iii) ***Linearidade:*** $(\alpha x + \beta y) \cdot z = \alpha(x \cdot z) + \beta(y \cdot z)$, $\forall \alpha, \beta \in K$.

Como sem dificuldade se pode verificar, num espaço com produto interno a aplicação que a cada vector x faz corresponder o número real não negativo $\sqrt{x \cdot x}$ define uma norma. Pondo

$$||x|| = \sqrt{x \cdot x}$$

o espaço vectorial converte-se então num espaço normado com norma induzida pelo produto interno .

Todo o espaço com produto interno pode, portanto, considerar-se um espaço normado. O recíproco não é, porém, verdadeiro, isto é, num espaço vectorial podem introduzir-se normas que não resultam de um produto interno.

Definição 2.11 (Espaço de Hilbert).

Chama-se espaço de Hilbert a um espaço com produto interno que é completo, isto é, a um espaço de Banach cuja norma é gerada por um produto interno.

2.2.5 Produtos de espaços normados

Sejam $(E_1, ||\cdot||_1)$ e $(E_2, ||\cdot||_2)$ dois espaços normados. No produto cartesiano $E = E_1 \times E_2$ pode introduzir-se uma qualquer das seguintes normas equivalentes:

$$\begin{aligned} ||(x,y)|| &= ||x||_1 + ||y||_2 \\ ||(x,y)|| &= \sqrt{(||x||_1)^2 + (||y||_2)^2} \\ ||(x,y)|| &= \max(||x||_1, ||y||_2) \end{aligned}$$

com $x \in E_1$, $y \in E_2$.

Deste modo o produto cartesiano de dois espaços normados converte-se num espaço normado que se denomina **espaço produto**. De forma análoga, se $(E_i, ||\cdot||_i)$, $i = 1, 2, \ldots, n$ for uma família finita de espaços normados, no produto cartesiano $E = E_1 \times E_2 \times \ldots \times E_n$ podem definir-se as normas equivalentes

$$\sum_{i=1}^{n} ||x_i||_i, \quad \sqrt{\sum_{i=1}^{n} (||x_i||_i)^2}, \quad \max_{i=1}^{n} ||x_i||_i$$

em que $x_i \in E_i$, $i = 1, 2, \ldots, n$.

No espaço produto resultante, os espaços $E_1, E_2, \ldots, E_n$ chamam-se os espaços coordenados e denomina-se **projecção** do conjunto produto E no espaço coordenado E_i a aplicação $P_i : E \to E_i$ tal que

$$P_i(x_1, x_2, \ldots, x_i, \ldots, x_n) = x_i.$$

Se os espaços coordenados forem espaços de Banach então o espaço produto é também de Banach. Para o provar considere-se, para maior simplicidade, o caso do produto de dois espaços de Banach E_1 e E_2.

Seja (x_n, y_n) uma sucessão de Cauchy de $E = E_1 \times E_2$ munido de uma das normas atrás indicadas. Então (x_n) é uma sucessão de Cauchy em E_1 e (y_n) uma sucessão de Cauchy de E_2 e, como estes espaços são completos, existem vectores $x \in E_1$ e $y \in E_2$ tais que $x_n \to x$ e $y_n \to y$.

Portanto, a sucessão $(x_n, y_n) \to (x, y)$ o que significa que o **produto cartesiano de dois espaços de Banach é um espaço de Banach, conclusão que se estende a um número finito de espaços**.

De forma análoga, verificam-se também nos produtos de espaços normados, como espaços métricos que são:

- *o produto de espaços normados compactos é compacto*;
- *o produto de espaços normados conexos é conexo*.

2.3 Exercícios

1. Mostre que num espaço vectorial é linearmente dependente qualquer sistema de vectores que:

   (a) contenha dois vectores iguais;

   (b) contenha o vector nulo.

2. Mostre que o espaço vectorial $C[a, b]$ das funções reais contínuas no intervalo $[a, b]$ é um subespaço vectorial $\mathcal{F}[a, b]$ do espaço das funções reais definidas em $[a, b]$.

3. Mostre que a intersecção de um número finito de subespaços dum espaço vectorial E é também um subespaço de E.

4. Mostre que no espaço vectorial $C[a, b]$ a aplicação de C em $\mathbb{R}$ definida pela forma

$$||f|| = \left[\int_a^b f^2(x)dx \right]^{\frac{1}{2}}$$

   é uma norma (chamada **norma da convergência em média quadrática**).

5. Mostre que em $C[a, b]$ a aplicação de $C \times C$ em $\mathbb{R}$ definida pela forma

$$f \cdot g = \int_a^b f(x)g(x)dx$$

   é um produto interno. Indique a norma por ele induzida.

6. Mostre que a norma dum espaço normado E é uma aplicação contínua de E em $\mathbb{R}$.

7. Mostre que $\mathbb{R}$, munido da norma $||x|| = |x|$, é um espaço de Banach.

8. Mostre que num espaço de Hilbert quaisquer vectores x e y verificam as relações:

   (a) $|x \cdot y| \leq ||x|| \cdot \|y||$ (desigualdade de Cauchy-Schwarz);

   (b) $||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2$ (regra do paralelogramo).

Capítulo 3

Espaço euclidiano $\mathbb{R}^n$

3.1 Espaço normado $\mathbb{R}^n$

Considere-se o conjunto $\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R} = \mathbb{R}^n$ cujo elementos são os énuplos $(x_1, x_2, \ldots, x_n)$ de n números reais segundo uma determinada ordem.

Se neste conjunto se definirem adição entre elementos do conjunto e multiplicação de um número real por um elemento do conjunto pela forma,

$$\begin{aligned}(x_1, \ldots, x_n) + (y_1, \ldots, y_n) &= (x_1 + y_1, \ldots, x_n + y_n), \\ \lambda \cdot (x_1, \ldots, x_n) &= (\lambda x_1, \ldots, \lambda x_n), \quad \lambda \in \mathbb{R},\end{aligned}$$

as operações assim definidas conferem ao conjunto $\mathbb{R}^n$ estrutura de espaço vectorial real. Os números reais são os **escalares** desse espaço e os elementos $(x_1, \ldots, x_n)$ os seus vectores, os quais se designarão abreviadamente por x ou (x_i).

Se se considerarem os n vectores linearmente independentes

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0), \\ e_2 &= (0, 1, \ldots, 0), \\ &\ldots \\ e_n &= (0, 0, \ldots, 1), \end{aligned}$$

verifica-se que qualquer vector $(x_1, x_2, \ldots, x_n)$ se exprime de forma única como combinação linear desses vectores:

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Os vectores $\{e_1, e_2, \ldots, e_n\}$ constituem portanto uma base do espaço vectorial $\mathbb{R}^n$, que é assim um espaço de dimensão n. Tal espaço é geralmente denominado espaço **cartesiano n–dimensional**.

No espaço cartesiano $\mathbb{R}^n$ pode agora definir-se uma nova operação, o chamado produto interno de dois vectores. Como é fácil de verificar, a aplicação de $\mathbb{R}^n \times \mathbb{R}^n$ em $\mathbb{R}^+$ que a dois vectores $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ de $\mathbb{R}^n$ associa o número real $x \cdot y$ [ou (x, y)] definido pela forma

$$x \cdot y = (x, y) = x_1 y_1 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

constitui um produto interno, que se denomina **produto interno natural ou euclidiano** de $\mathbb{R}^n$.

Se, sendo $x \neq 0$ e $y \neq 0$, é $x \cdot y = 0$ os vectores x e y dizem-se perpendiculares ou ortogonais.

Definido o produto interno euclidiano em $\mathbb{R}^n$ pode a partir dele definir-se uma **norma**.

De facto, tal como se referiu no parágrafo 2.2.4 do Capítulo anterior, a aplicação que a cada vector x faz corresponder o número real não negativo $\sqrt{x \cdot x}$ define uma norma em $\mathbb{R}^n$, chamada a **norma euclidiana**.

Pondo $||x|| = \sqrt{x \cdot x}$ o espaço cartesiano $\mathbb{R}^n$ converte-se pois num espaço vectorial normado, que se denomina o **espaço euclidiano** n–**dimensional** e se designa também por $\mathbb{R}^n$.

Tendo em atenção a definição do produto interno euclidiano, a norma euclidiana pode exprimir-se pela forma

$$||x|| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Observação. No espaço cartesiano pode introduzir-se outras normas distintas da euclidiana pela forma

$$\begin{aligned} ||x||_1 &= |x_1| + \ldots + |x_n| \\ ||x||_2 &= \max(|x_1| + \ldots + |x_n|). \end{aligned}$$

Estas normas não resultam todavia de um produto interno e por isso o espaço cartesiano munido de qualquer delas é um espaço normado mas não espaço euclidiano.

Por outro, como se pode observar (vd. 2.2.5. do Capitulo anterior), o espaço euclidiano $\mathbb{R}^n$ é afinal o espaço produto de n espaços normados todos iguais ao espaço normado $\mathbb{R}$.

Tal como acontece com qualquer espaço normado, o espaço euclidiano $\mathbb{R}^n$ converte-se num espaço métrico quando nele se toma para métrica, ou função distância , a métrica induzida pela norma, isto é, quando se define distância entre dois vectores por meio da igualdade

$$d(x, y) = ||x - y||$$

ou seja

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

A métrica assim definida denomina-se **métrica euclidiana ou métrica usual de** $\mathbb{R}^n$.

Uma vez que o espaço euclidiano $\mathbb{R}^n$ é simultaneamente um espaço vectorial e um espaço métrico, os seus elementos podem denominar-se indiferentemente **pontos** ou **vectores** desse espaço. Em qualquer dos casos, por uma questão de maior simplicidade na escrita e desde que isso não dê lugar a dúvidas os elementos representar-se-ão por uma só letra minúscula do alfabeto latino, que será encimado por uma seta quando se utilizar a designação "vector".

Assim, se se utilizar a designação "ponto", o elemento $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ representar-se-á simplesmente por x, escrevendo-se

$$x = (x_1, \ldots, x_n)$$

e

$$x_1, \ldots, x_n$$

denominam-se coordenadas do ponto.

Se $(x_1, \ldots, x_n)$ for considerado um "vector", então representar-se-á por $\overrightarrow{x}$, escrevendo-se $\overrightarrow{x} = (x_1, \ldots, x_n)$ e $x_1, \ldots, x_n$ são agora as componentes do vector.

3.2 Noções topológicas em $\mathbb{R}^n$

Com base na métrica definida no espaço euclidiano podem agora estender-se-lhe os conceitos anteriormente definidos para os espaços métricos.

Sejam $a = (a_1, \ldots, a_n)$ um ponto de $\mathbb{R}^n$ e r um número real positivo, $r > 0$.

Chama-se **bola aberta** de centro em a e raio r ao conjunto dos pontos de $\mathbb{R}^n$ situados a uma distância de a inferior a r, isto é, o conjunto dos pontos $x = (x_1, \ldots, x_n)$ tais que $||x - a|| < r$.

O conjunto dos pontos x para os quais $||x - a|| \leq r$ chama-se a **bola fechada** de centro a e raio r.

No caso de $\mathbb{R}^1$, a bola aberta de centro $a = (a)$ e raio r é o intervalo aberto com centro em a e comprimento $2r$; em $\mathbb{R}^2$ é o interior do círculo com centro em a e raio r; em $\mathbb{R}^3$ o interior da esfera com centro em a e raio r; para $n > 3$ não é possível dar qualquer interpretação geométrica e diz-se, geralmente, que se trata da **hiperesfera** de centro a e raio r.

Além de bola definem-se também intervalos n–dimensionais. Assim, se $a = (a_1, \ldots, a_n)$ e $b = (b_1, \ldots, b_n)$ são dois pontos distintos de $\mathbb{R}^n$ tais que $a_i < b_i$ para $i = 1, 2, \ldots, n$ chama-se **intervalo aberto** n–dimensional de extremidades a e b e representa-se por $]a, b[$ ao conjunto dos pontos $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ tais que é $a_i < x_i < b_i$, $i = 1, 2, \ldots, n$.

Tal conjunto pode considerar-se como o "produto cartesiano"

$$]a, b[=]a_1, b_1[\times]a_2, b_2[\times \ldots \times]a_n, b_n[.$$

Se for $a_i \leq x_i \leq b_i$ para todo o i, obtém-se o intervalo n–dimensional fechado $[a, b]$ produto de n intervalos fechados da recta real.

Das definições anteriores resulta que toda a bola com centro num dado ponto contém um intervalo que contém esse ponto e reciprocamente.

Um subconjunto A de $\mathbb{R}^n$ diz-se **conjunto aberto** se qualquer seu ponto é centro de uma bola aberta inteiramente contida em A. Como logo se reconhece toda a bola aberta é um conjunto aberto e tanto o conjunto vazio como o próprio $\mathbb{R}^n$ são abertos. Os abertos de $\mathbb{R}^n$ satisfazem as propriedades da Proposição 1.2 do Capítulo 1 e definem, portanto, uma topologia em $\mathbb{R}^n$.

Um subconjunto A de $\mathbb{R}^n$ diz-se **fechado** se o seu complementar é aberto. O conjunto vazio e $\mathbb{R}^n$ são pois dois fechados (e simultaneamente abertos como se viu anteriormente).

Um subconjunto A de $\mathbb{R}^n$ diz-se **limitado** se existe uma bola (ou um intervalo) que o contém.

Chama-se **vizinhança** de um ponto $a = (a_1, \ldots, a_n)$ de $\mathbb{R}^n$ a qualquer conjunto aberto que o contenha. Assim, qualquer bola aberta que contenha a é uma vizinhança de a. Em geral são vizinhanças deste tipo que utilizam e, em particular, a chamada vizinhança de raio $\epsilon > 0$ do ponto a que é a bola aberta com centro a e raio ϵ, isto é, o conjunto dos pontos $x = (x_1, x_2, \ldots, x_n)$ tais que

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 + \ldots + (x_n - a_n)^2 < \epsilon^2.$$

Tal vizinhança será designada por $V(a, \epsilon)$.

Se da vizinhança se excluir o ponto a, então a vizinhança denomina-se **vizinhança reduzida** de a de raio ϵ e escreve-se $V'(a, \epsilon)$.

Como toda a bola aberta com centro em a contém um intervalo aberto que contém a, e reciprocamente, em vez de bolas abertas podem tomar-se para vizinhanças dum ponto quaisquer intervalos abertos que o contenham. Assim, por exemplo, o intervalo $]a-\epsilon, a+\epsilon[$ constituído pelos pontos $(x_1, \ldots, x_n)$ tais que

$$\begin{array}{l} |x_1 - a_1| < \epsilon \\ |x_2 - a_2| < \epsilon \\ \ldots \\ |x_n - a_n| < \epsilon \end{array} \tag{3.1}$$

é uma vizinhança de $a = (a_1, a_2, \ldots, a_n)$ que contém a bola aberta $V(a, \epsilon)$.

Sejam A um subconjunto de $\mathbb{R}^n$ e $a = (a_1, a_2, \ldots, a_n)$ um ponto qualquer de $\mathbb{R}^n$. O ponto a diz-se:

Ponto interior a A quando existe uma vizinhança de a inteiramente contida em A;

Ponto exterior a A se é interior ao complementar de A;

Ponto fronteiro a A quando não é interior nem exterior a A;

Ponto aderente a A se não é exterior a A, isto é, quando é interior ou fronteiro a A;

Ponto de acumulação de A se qualquer vizinhança de a contém uma infinidade de pontos de A distintos do próprio a.

O conjunto de todos os pontos interiores de A forma o interior de A, designado por $int(A)$; o conjunto de pontos exteriores a A constitui o seu **exterior**, designado por $ext(A)$; o conjunto dos pontos fronteiriços é a **fronteira** de A, designada por $front(A)$; o conjunto dos pontos aderentes a A chama-se **fecho** ou **aderência** de

A, representa-se por $\overline{A}$; o conjunto dos pontos de acumulação de A chama-se o derivado de A e representada por A'.

Como está bem de ver, em $\mathbb{R}^n$ todos os pontos são interiores e não há pontos exteriores nem pontos fronteiros, isto é:

$$int(\mathbb{R}^n) = \mathbb{R}^n, \quad ext(\mathbb{R}^n) = \emptyset \quad \text{e} \quad front(\mathbb{R}^n) = \emptyset.$$

No caso do conjunto vazio, $int(\emptyset) = \emptyset$, $ext(\emptyset) = \mathbb{R}^n$ e $front(\emptyset) = \emptyset$.

Das definições anteriores decorre que:

(i) o fecho ou aderência de um conjunto A é o conjunto $\overline{A} = A \cup A'$;

(ii) um conjunto A é fechado se e só se coincide com o seu fecho, $A = \overline{A}$, isto é, contém todos os seus pontos de acumulação;

(iii) um subconjunto A é aberto se e só se coincide com o seu interior: $A = int(A)$.

Um subconjunto A de $\mathbb{R}^n$ diz-se **compacto** se de qualquer sua cobertura aberta se pode extrair uma subcobertura finita. Tal como na recta real, em $\mathbb{R}^n$, produto de n espaços idênticos a $\mathbb{R}$, todo o intervalo fechado $I = [a, b]$ é compacto, visto ser o produto de n intervalos fechados e portanto compactos de $\mathbb{R}$. Por outro lado, em $\mathbb{R}^n$ toda a sucessão (I_n) de intervalos e encaixados

$$I_1 \supset I_2 \supset \ldots \supset I_n \supset I_{n+1} \supset \ldots$$

tem intersecção não vazia:

$$\bigcap_{i=1}^{\infty} I_i \neq \emptyset.$$

Teorema 3.1 (**Teorema de Bolzano – Weierstrass**). *Todo o conjunto infinito e limitado de $\mathbb{R}^n$ tem pelo menos um ponto de acumulação.*

Demonstração. Seja A um conjunto infinito e limitado. Como A é limitado, existe um intervalo fechado I tal que $A \subset I$. Mas como I é completo então A é um subconjunto infinito dum conjunto compacto e, portanto, tem pelo menos um ponto de acumulação, como se pretendia provar. ■

Teorema 3.2 (**Heine – Borel**). *Todo o subconjunto fechado e limitado de $\mathbb{R}^n$ é compacto.*

Demonstração. Seja A um conjunto nessas condições. Como A é limitado, existe um intervalo fechado I que o contém, $A \subset I$. Então A é um subconjunto fechado dum conjunto compacto e portanto é compacto. ■

Por outro lado, facilmente se pode provar que:

Proposição 3.1. *Se um subconjunto A de $\mathbb{R}^n$ tem alguma das propriedades seguintes tem também as outras duas:*

(a) A é fechado e limitado;

(b) A é compacto;

(c) Todo o subconjunto infinito de A tem um ponto de acumulação em A.

Note-se que (b) e (c) são equivalentes em qualquer espaço métrico mas (a) em geral não implica (b) e (c), embora em $\mathbb{R}^n$ isso aconteça.

Um subconjunto A de $\mathbb{R}^n$ é **convexo** se não existe nenhuma partição de A em duas partes fechadas (ou abertas) não vazias.

Por outro lado, um subconjunto A de $\mathbb{R}^n$ diz-se **conexo por arcos** se dois quaisquer dos seus pontos podem ser ligados por um arco ou caminho inteiramente contido em A. Todo o **conjunto conexo** por arcos é **conexo**.

Como em $\mathbb{R}$ os conexos são o próprio $\mathbb{R}$ e os intervalos, segue-se que $\mathbb{R}^n$ e os intervalos de $\mathbb{R}^n$ são conexos (Capítulo 2 – 2.5).

Em $\mathbb{R}^n$ toda a sucessão de Cauchy é convergente. Portanto $\mathbb{R}^n$:

(a) Como espaço métrico é um **espaço completo**;

(b) Como espaço normado é um **espaço de Banach**;

(c) Como espaço com produto interno é um **espaço de Hilbert**.

Chama-se **domínio** a todo o subconjunto aberto, não vazio e conexo de $\mathbb{R}^n$.

Um domínio D diz-se **simplesmente conexo** se todo o ponto interior a qualquer contorno fechado inteiramente contido em D pertence também a D. Caso contrário diz-se **multiplamente conexo**. Assim, por exemplo, em $\mathbb{R}^2$ o conjunto dos pontos (x, y) que verifica a condição $x^2 + y^2 < 1$ é um domínio simplesmente conexo, mas o conjunto dos pontos que satisfazem simultaneamente as condições $x^2 + y^2 > 1$ e $x^2 + y^2 < 4$ é um domínio **duplamente conexo**.

3.3 Exercícios

1. Mostre que os vectores $\overrightarrow{i} = (1,0,0)$, $\overrightarrow{j} = (0,1,0)$ e $\overrightarrow{k} = (0,0,1)$ constituem uma base do espaço $\mathbb{R}^3$.

2. Prove que em $\mathbb{R}^n$ toda bola com centro num dado ponto contém um intervalo que contém esse ponto e reciprocamente.

3. Represente graficamente as bolas fechadas de $\mathbb{R}^2$ com centro no mesmo ponto a e com o mesmo raio $r > 0$ originados por cada uma das três normas indicadas no parágrafo 3.1.

4. Mostre que se x e y são dois vectores ortogonais de $\mathbb{R}^n$ então verificam a relação $||x + y||^2 = ||x||^2 + ||y||^2$ (Teorema de Pitágoras).

5. Mostre que se x e y são dois vectores não nulos de $\mathbb{R}^n$ existe um ângulo θ compreendido entre 0 e π tal que

$$\cos\theta = \frac{xy}{||x|| \cdot ||y||}.$$

6. Mostre que toda a bola de $\mathbb{R}^n$ é um conjunto: (a) convexo; (b) conexo.

7. Seja A um subconjunto de $\mathbb{R}^n$. Mostre que se qualquer subconjunto infinito de A tem um ponto de acumulação em A é:

   (a) fechado e limitado;

   (b) (b) compacto.

Capítulo 4

Funções reais de n variáveis reais

4.1 Funções de várias variáveis

Chama-se função real de n variáveis a toda a aplicação f de um conjunto D de $\mathbb{R}^n$ em $\mathbb{R}$

$$f : D \subset \mathbb{R}^n \to \mathbb{R}.$$

O conjunto D é o **domínio** ou **campo de existência** da função. Se $x = (x_1, x_2, \ldots, x_n)$ for um ponto genérico do domínio da função escreve-se

$$y = f(x_1, x_2, \ldots, x_n) \quad \text{ou} \quad y = f(x)$$

e diz-se que y (variável real) é função do ponto x ou das n variáveis reais $x_1, x_2, \ldots, x_n$.

A $x_1, \ldots, x_n$ chamam-se **variáveis independentes** ou **argumentos** e a y **variável dependente** ou **função**.

Ao conjunto dos valores de y obtidos quando x percorre o domínio da função, chama-se **contradomínio** da função.

Uma função fica definida quando é dado o seu domínio e indicada a lei que faz corresponder a cada ponto do domínio um valor y do contradomínio. Em geral essa lei é dada por uma expressão analítica, mas pode também ser dada de outro modo como, por exemplo, por uma tabela, uma relação de natureza física ou de natureza geométrica.

A mesma função pode, por vezes, ser definida por expressões analíticas diferentes, como é, por exemplo, o caso das expressões

$$\begin{aligned} f(x,y) &= (x+y)^2 \\ g(x,y) &= x^2+2xy+y^2 \end{aligned}$$

que definem a mesma função.

Por outro lado, uma mesma função pode ser definida por expressões analíticas distintas em diferentes partes do seu domínio, como é, por exemplo, o caso da função

$$f(x,y) = \begin{cases} x+y & \text{se} \quad x<y \\ 1 & \text{se} \quad x=y \\ x-y & \text{se} \quad x>y. \end{cases}$$

Por vezes, define-se uma função apenas pela sua expressão analítica sem se indicar o seu domínio. Subentende-se, então, que o domínio é o conjunto dos pontos para os quais são possíveis, no campo real, as operações indicadas na expressão analítica.

Exemplo 4.1. O domínio da função $z = +\sqrt{1 - x^2 - y^2}$ é o conjunto dos pontos (x, y) de $\mathbb{R}^2$ tais que:

$$1 - x^2 - y^2 \geq 0 \quad \text{ou} \quad x^2 + y^2 \leq 1.$$

Trata-se pois, do círculo com centro na origem e raio igual a 1 (Figura 4.1). ◄

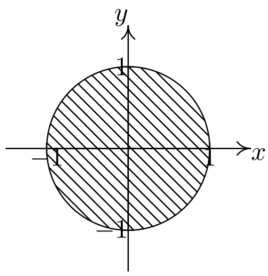

Figura 4.1: Domínio da função $z = +\sqrt{1 - x^2 - y^2}$

Exemplo 4.2. O domínio da função $z = \ln(x + y)$ é o conjunto dos pontos (x, y) tais que: $x + y > 0$ ou $y > -x$. É, portanto, o semi-plano situado para a direita da recta $y = -x$ (com exclusão da recta) (Figura 4.2). ◄

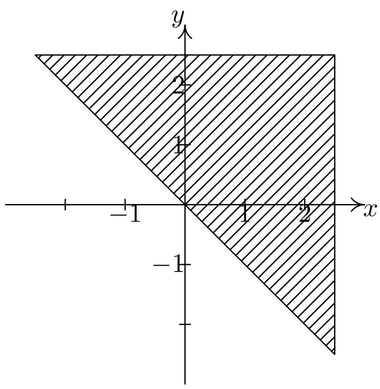

Figura 4.2: Domínio da função $z = \ln(x + y)$

4.2 Gráfico duma função

Dada a função $f : D \subset \mathbb{R}^n \to \mathbb{R}$, chama-se gráfico da função ao conjunto dos pontos $(x_1, x_2, \ldots, x_n, y)$ de $\mathbb{R}^{n+1}$, com $(x_1, x_2, \ldots, x_n) \in D$.

Se $n = 1$, o gráfico é um conjunto de pontos de $\mathbb{R}^2$ que, uma vez fixado um referencial cartesiano, pode ser visualizado geometricamente por uma curva do plano. Se $n = 2$, o gráfico é um conjunto de $\mathbb{R}^3$ que se pode também, instituído um referencial cartesiano, visualizar geometricamente por uma superfície do espaço a três dimensões. Se, porém, for $n > 2$, não é possível obter qualquer visualização geométrica, dado o carácter tridimensional do espaço ordinário.

4.3 Funções implícitas e funções multívocas

Embora seja satisfeita por certos pares (x, y) de números reais, a equação $x^2 + y^2 = 1$ não representa uma função, visto que, se x for variável independente, então a cada x tal que $|x| \leq 1$ corresponderiam dois valores de y o que contradiz a definição. O que acontece é que na equação dada estão implicitamente definidas, não uma função, mas várias, nomeadamente

$$y = f(x) = +\sqrt{1 - x^2} \quad (|x| \leq 1)$$

ou

$$y = g(x) = -\sqrt{1 - x^2} \quad (|x| \leq 1)$$

ou

$$y = h(x) = \begin{cases} +\sqrt{1-x^2}, & 0 \leq x \leq 1 \\ \\ -\sqrt{1-x^2}, & -1 \leq x \leq 0. \end{cases}$$

De forma análoga, a fórmula $y =$ arcsen x, $-1 \leq x \leq 1$ não define uma função, visto que, para cada valor de x, há uma infinidade de ângulos cujo *seno* é x. Neste caso, diz-se que a fórmula define uma **função multívoca** .

4.4 Limites

4.4.1 Definição de limite

Definição 4.1.

> *Seja $f(x) = f(x_1, \ldots, x_n)$ uma função de domínio D definida em todos os pontos duma vizinhança do ponto $a = (a_1, \ldots, a_n)$ excepto, possivelmente, no próprio ponto a. Diz-se que $f(x)$ tende para o limite ℓ ao tender x para a, quando a todo o $\epsilon > 0$ corresponde um $\delta = \delta(\epsilon) > 0$ tal que seja $|f(x) - \ell| < \epsilon$ para todos os $x \in D$ tais que $0 < ||x - a|| < \delta$.*

A definição dada equivale a afirmar que, para toda a sucessão de pontos de D convergente para a, a correspondente sucessão dos valores da função nesses pontos converge para o número ℓ.

Se a definição é satisfeita escreve-se $\lim_{x \to a} f(x) = \ell$ ou

$$\lim_{\substack{x_1 \to a_1 \\ x_2 \to a_2 \\ \cdots \\ x_n \to a_n}} f(x_1, \ldots, x_n) = \ell.$$

A definição pode traduzir-se em termos de vizinhanças dizendo que a toda a vizinhança $V(\ell, \epsilon)$ de ℓ corresponde uma vizinhança reduzida $V'(a, \delta)$ de a tal que, quando x percorre $V'(a, \delta)$ segundo pontos do domínio D da função, os correspondentes valores da função estão contidos na vizinhança $V(\ell, \epsilon)$, isto é

$$x \in V'(a, \delta) \cap D \Longrightarrow f(x) \in V(\ell, \epsilon).$$

Como a própria definição o indica, não é necessário que o ponto a pertença ao domínio da função; basta que seja um seu ponto de acumulação.

Essencial sim é que a desigualdade $|f(x) - \ell| < \epsilon$ se verifique sempre, seja qual for o modo como x tende para a dentro do domínio da função.

Como é evidente, o ponto x pode tender para a ao longo de uma curva e na vizinhança há uma infinidade de curvas que conduzem de x a a. Ora, para que o limite exista, é necessário que, qualquer que seja o percurso seguido, o valor final obtido seja sempre o mesmo. Se, variando o caminho seguido, se obtêm valores distintos, a função não tem limite.

Exemplo 4.3. A função $f(x, y) = \dfrac{x - y}{x + y}$ $(x \neq -y)$ não tem limite na origem. Com efeito, se (x, y) tender para $(0, 0)$ ao longo do eixo das abcissas é

$$\lim_{x \to 0} f(x, 0) = \lim_{x \to 0} \frac{x}{x} = 1;$$

mas, se $(x, y) \to (0, 0)$ ao longo do eixo das ordenadas, tem-se

$$\lim_{y\to 0} f(0, y) = \lim_{y\to 0} \frac{-y}{y} = -1.$$

Como estes limites são distintos, a função não tem limite quando (x, y) tende para $(0, 0)$. ◀

Exemplo 4.4. Seja a função $f(x, y) = \dfrac{xy}{x^2 + y^2}$ cujo domínio é todo o plano excepto o ponto $(0, 0)$ e verifiquemos se a função tem limite na origem. Se $(x, y) \to (0, 0)$ ao longo do eixo das abcissas tem-se $\lim\limits_{x\to 0} f(x, 0) = 0$; se $(x, y) \to (0, 0)$ ao longo do eixo das ordenadas tem-se $\lim\limits_{y\to 0} f(0, y) = 0$.

Os limites calculados ao longo dos eixos coordenados são iguais, mais isso não permite, todavia, concluir que a função tem limite na origem, pois há uma infinidade de outras formas de (x, y) convergir para a origem. Se, por exemplo, a convergência se fizer da recta $y = mx$, tem-se

$$\lim_{x\to 0} f(x, mx) = \frac{m}{1 + m^2}$$

o que mostra que qualquer valor de m conduz a um resultado diferente. Portanto, a função não tem limite na origem. ◀

4.4.2 Limites sucessivos

Quando se diz que $f(x) = f(x_1, \ldots, x_n)$ tende para o limite ℓ ao tender $x = (x_1, \ldots, x_n)$ para $a = (a_1, \ldots, a_n)$, pressupõe-se, na definição 4.1, que as variáveis independentes tendem todas simultaneamente para os respectivos limites. Por esse motivo muitas vezes se escreve

$$\lim_{\substack{x_1 \to a_1 \\ x_2 \to a_2 \\ \cdots \\ x_n \to a_n}} f(x_1, \ldots, x_n) = \ell,$$

em vez de $\lim_{x \to a} f(x) = \ell$.

Ocorre, porém, frequentemente ser necessário calcular o limite da função em relação a uma das variáveis, considerando as demais constantes, e reiterar em seguida o processo em relação a cada uma das restantes variáveis. O limite assim calculado chama-se **limite sucessivo** (ou reiterado) da função. Se a ordem de cálculo dos sucessivos limites tiver sido, por exemplo, a mesma das coordenadas do ponto x escreve-se então

$$\lim_{x_n \to a_n} \ldots \lim_{x_2 \to a_2} \cdot \lim_{x_1 \to a_1} f(x_1, x_2, \ldots, x_n)$$

e de forma análoga nos demais casos.

Uma mudança na ordem de passagem ao limite pode alterar o resultado final pelo que qualquer modificação da ordem deve ser cuidadosamente considerada. Se os diversos limites sucessivos forem iguais diz-se que há permutabilidade dos limites sucessivos. Todavia, como indicam os Exemplos 4.6 e 4 a seguir apresentados, a permutabilidade dos limites sucessivos não implica a existência do limite, nem, por outro lado, a existência deste implica a permutabilidade.

Exemplo 4.5. A função $f(x, y) = \dfrac{x^2}{x^2 + y^2}$ não tem limite na origem. Todavia

$$\lim_{x \to 0} \lim_{y \to 0} \frac{x^2}{x^2 + y^2} = 1; \quad \lim_{y \to 0} \lim_{x \to 0} \frac{x^2}{x^2 + y^2} = 0. \quad \blacktriangleleft$$

Exemplo 4.6. Como se viu atrás, a função $f(x, y) = \dfrac{xy}{x^2 + y^2}$ não tem limite quando $(x, y) \to (0, 0)$. No entanto os limites sucessivos existem e são iguais

$$\lim_{x\to 0}\lim_{y\to 0} f(x, y) = \lim_{y\to 0}\lim_{x\to 0} f(x, y) = 0. \quad \blacktriangleleft$$

Exemplo 4.7. Considere-se a função

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

A função tem limite quando $(x, y) \to (0, 0)$: $\lim\limits_{\substack{x\to 0\\ y\to 0}} f(x) = 0.$

Todavia não existem os limites sucessivos

$$\lim_{x\to 0}\lim_{y\to 0} f(x, y) \quad \text{e} \quad \lim_{y\to 0}\lim_{x\to 0} f(x, y). \blacktriangleleft$$

4.4.3 Limite infinito e limite quando $||x|| \to \infty$

A definição de limite dada em 4.1 admite as seguintes generalizações.

Definição 4.2 (Limite infinito).

*Diz-se que a função $f(x)$ tem **limite infinito** ao tender x para a quando a todo o número real $N > 0$ corresponde* *um $\delta > 0$ tal que é $|f(x)| > N$ sempre que $||x - a|| < \delta$.*

Nestas condições escreve-se $\lim\limits_{x \to a} f(x) = \infty$.

Definição 4.3 (**Limite finito**).

*Diz-se que $f(x)$ tem **limite finito** ℓ quando $||x|| \to \infty$, se a todo o $\epsilon > 0$ corresponde um número $N > 0$ tal que é $|f(x) - \ell| < \epsilon$ sempre que $||x|| > N$.*

Escreve-se então, $\lim\limits_{||x|| \to \infty} f(x) = \ell$.

4.4.4 Operações com limites

Os teoremas sobre os limites da soma, do produto e do quociente de funções reais de uma variável real permanecem válidos para as funções de n variáveis como é fácil de verificar. Assim, se $f(x)$ e $g(x)$ são duas funções definidas no mesmo domínio D de $\mathbb{R}^n$, a é ponto de acumulação de D e $\lim\limits_{x \to a} f(x) = \ell$, $\lim\limits_{x \to a} g(x) = v$, então:

(a) $\lim\limits_{x \to a} (f + g) = \ell + v$

(b) $\lim\limits_{x \to a} (f \cdot g) = \ell \cdot v$

(c) $\lim\limits_{x \to a} \dfrac{f}{g} = \dfrac{\ell}{v}$ se $v \neq 0$.

4.5 Continuidade

4.5.1 Continuidade num ponto

Definição 4.4.

> *Seja $f(x) = f(x_1, \ldots, x_n)$ uma função de domínio D definida em todos os pontos duma vizinhança do ponto $a = (a_1, \ldots, a_n)$. Diz-se que $f(x)$ é contínua no ponto a quando*

$$\lim_{x \to a} f(x) = f(a).$$

A definição anterior é equivalente à seguinte: diz-se que $f(x) = f(x_1, \ldots, x_n)$ é contínua no ponto a do seu domínio quando a todo o $\epsilon > 0$ corresponde um $\delta = \delta(\epsilon) > 0$ tal que seja

$$|f(x) - f(a)| < \epsilon$$

sempre que $x \in D$ e $||x - a|| < \delta$.

De acordo com a definição é, portanto, necessário que se verifiquem as três condições seguintes:

(a) Existir $f(a)$;

(b) Existir $\lim_{x \to a} f(x)$;

(c) $\lim_{x \to a} f(x) = f(a)$.

Caso não se verifique alguma destas condições a função diz-se **descontínua no ponto** a. Se a descontinuidade resultar da não observância da condição (c), ela pode ser eliminada alterando o valor da função no ponto de modo a coincidir com o valor do limite da função nesse ponto, pelo que se diz que se trata de uma **descontinuidade removível**. Quando, porém a descontinuidade resulta da não verificação da condição (b) a descontinuidade diz-se **descontinuidade essencial**.

Exemplo 4.8. A função $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$

é descontínua no ponto $(0,0)$.

De facto, como se viu no Exemplo 4.4 do parágrafo 4.4.1, não existe $\lim\limits_{(x,y)\to(0,0)} f(x,y)$. A **descontinuidade é essencial**. ◄

Exemplo 4.9. A função $f(x,y) = \begin{cases} \dfrac{xy}{(x^2+y^2)^{1/2}}, & \text{se } x \neq 0,\ y \neq 0 \\ \\ 1, & \text{se } x = 0,\ y = 0 \end{cases}$

é descontínua no ponto $(0,0)$, visto que $\lim\limits_{\substack{x\to 0\\ y\to 0}} f = 0 \neq f(0,0) = 1$.

A descontinuidade pode, porém, ser removida, bastando para isso redefinir a função nesse ponto de modo a ser $f(0,0) = 0$. ◄

Exemplo 4.10. A função $z = \dfrac{4}{x^2+y^2}$ é obviamente descontínua no ponto $(0,0)$ visto nem sequer estar definida nesse ponto. ◄

4.5.2 Continuidade num conjunto

Uma função diz-se contínua num conjunto (do seu domínio) se e só se for contínua em todos os pontos desse conjunto.

4.5.3 Continuidade uniforme

Se uma função $f(x)$ é contínua num dado conjunto S, então isso significa que, em cada ponto x_0 do conjunto, qualquer que seja o número $\epsilon > 0$ escolhido é sempre possível fazer-lhe corresponder um $\delta > 0$ de tal modo que seja $|f(x) - f(x_0)| < \epsilon$ sempre que $||x - x_0|| < \delta$. O valor de δ depende, geralmente, não só do ϵ escolhido, mas também do ponto x_0 considerado. Pode, todavia, acontecer que escolhido um $\epsilon > 0$ seja possível determinar um mesmo $\delta > 0$ válido para todos os pontos do conjunto. Quando isto acontece a função diz-se **uniformemente contínua** no conjunto. Isto é:

Definição 4.5.

Seja $f(x)$ uma função definida num conjunto de $\mathbb{R}^n$. Diz-se que a função é uniformemente contínua nesse conjunto quando para todo $\epsilon > 0$ existe um $\delta > 0$ (dependente apenas de ϵ) tal que, qualquer que seja o par de pontos x', x'' de S tais que

$$||x' - x''|| < \delta, \quad \textit{seja} \quad |f(x') - f(x'')| < \epsilon.$$

Da definição resulta que se uma função é uniformemente contínua num conjunto é contínua nesse conjunto. A recíproca, porém, não é valida.

4.5.4 Propriedades das funções contínuas

As funções reais contínuas de n variáveis reais gozam de propriedades análogas às das funções de uma só variável real. As respectivas demonstrações ou foram feitas nos capítulos anteriores ou são análogas às das funções de uma variável, pelo que nos limitemos a enuncia-las.

Teorema 4.1. *Se $f(x)$ e $g(x)$ são funções contínuas no ponto a, também a sua soma, o seu produto e o seu quociente são funções contínuas em a, salvo, no caso de quociente, se $g(a) = 0$.*

Teorema 4.2. *Se $f(x)$ é contínua em a e $g(t)$ contínua em $t_0 = f(a)$ então a função composta gof é contínua em a.*

Teorema 4.3. *Se $f(x)$ é contínua em a e $f(x) \neq 0$, existe uma vizinhança de a na qual a função mantém o mesmo sinal que tem em a.*

Teorema 4.4 (Teorema de Weierstrass). *Função contínua num conjunto compacto (limitado e fechado) é limitada nesse conjunto e admite aí um máximo e um mínimo .*

Teorema 4.5 (Teorema de Bolzano). *Se $f(x)$ é contínua num conjunto conexo e toma aí dois valores k e l então toma também qualquer valor entre k e ℓ.*

Teorema 4.6. *Função contínua num conjunto compacto é uniformemente contínua nesse conjunto.*

De acordo com o estudo feito nos capítulos anteriores, tanto o Teorema 4.4 como o Teorema 4.5 admitem formulação mais simples.

Assim, como em $\mathbb{R}^n$ e, portanto, também em $\mathbb{R}^1 = \mathbb{R}$, os conjuntos compactos são os limitados e fechados, o enunciado de 4.4 pode ser substituído pelo seguinte: **as funções contínuas transformam conjuntos compactos em conjuntos compactos de $\mathbb{R}$.**

Por outro lado, como em $\mathbb{R}$ os conjuntos conexos são os intervalos, o Teorema 4.5 pode enunciar-se: **as funções contínuas transformam conjuntos conexos de $\mathbb{R}^n$ em conjuntos conexos de $\mathbb{R}$.**

4.6 Exercícios

1. Determine o domínio de cada uma das funções:

   (a) $f(x, y) = \ln y^2 + \sqrt{1 - x^2}$;

   (b) $z = \text{ arcsen } (x + y)$;

   (c) $z = \dfrac{1}{\sqrt{xy}}$.

2. Determine os seguintes limites caso existam:

   (a) $\lim\limits_{\substack{x \to 5 \\ y \to 0}} \dfrac{\text{tg } (xy)}{y}$;

   (b) $\lim\limits_{\substack{x \to \infty \\ y \to \infty}} \dfrac{x^4 + y^4}{10^{x+y}}$;

   (c) $\lim\limits_{\substack{x \to 0 \\ y \to 0}} \dfrac{x}{x + y}$.

3. Determine os pontos de descontinuidade das funções:

   (a) $f(x, y) = \text{ sen } \dfrac{1}{xy}$;

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

   (c) $f(x, y) = \dfrac{1}{xyz}$.

4. Demonstre a continuidade em todo o plano xoy da função $f(x, y) = ax + by + c$.

5. Estude a continuidade em (0,0) da função com a definição

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

6. Mostre que toda a função racional fraccionária de $(x_1, x_2, \ldots, x_n)$ é contínua nos pontos que não anulam o denominador.

7. Demonstre que se $f(x_1, \ldots, x_n)$ é contínua num ponto $a = (a_1, \ldots, a_n)$ interior ao seu domínio e se $f(a) \neq 0$, então existe uma vizinhança de a na qual a função não se anula.

Capítulo 5

Derivação

5.1 Derivadas parciais

Definição 5.1.

Seja $a = (a_1, \ldots, a_n)$ *um ponto interior do domínio da função* $f(x) = f(x_1, \ldots, x_n)$. *Se se atribuírem a todas as variáveis excepto* x_i *os valores constantes* $a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$, *a função dada converte-se numa função de uma só variável* x_i. *E esta função de uma única variável admitir derivada no ponto* $x_i = a_i$, *isto é, se existir*

$$\lim_{\Delta x_i \to 0} \frac{f(a_1, \ldots, a_i + \Delta x_i, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{\Delta x_i},$$

então a este limite chama-se a ***derivada parcial da função*** *em ordem a* x_i *no ponto* a *e designa-se por qualquer das formas*

$$\frac{\partial f}{\partial x_i}(a), \quad f'_{x_i}(a), \quad (D_{x_i} f)_a.$$

De forma análoga se definem as derivadas parciais em ordem às demais variáveis no ponto a.

Da definição 5.1 resulta que as derivadas parciais, se existirem, obedecem às regras de derivação utilizadas para as funções de uma variável.

Se, por exemplo, a função for uma função de duas variáveis, $z = f(x, y)$, a sua derivada parcial em ordem a x no ponto (a, b) é

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

caso o limite exista.

De forma análoga, a derivada parcial de $f(x, y)$ em ordem a y no ponto (a, b) é o limite, quando existe

$$f'_y(a, b) = \lim_{k \to 0} \frac{f(a, b + k) - f(a, b)}{k}.$$

Ao contrário do que acontece com as funções de uma só variável, em que a existência de derivada finita num ponto implica a continuidade da função nesse ponto, no caso das funções de mais de uma variável o facto de existirem as derivadas parciais de 1ª ordem não implica que a função seja contínua nesse ponto.

É, por exemplo, o que acontece com a função

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

que, como se viu no Exemplo 4.8 de 4.5.1 do Capítulo 4, é descontínua em (0,0), mas todavia tem derivadas parciais nesse ponto: $f'_x(0,0)$ e $f'_y(0,0)$.

Como adiante se verá, a continuidade de $f(x)$ num ponto a só se pode deduzir da existência das derivadas parciais em a se essas derivadas forem contínuas nesse ponto.

5.2 Derivação sucessiva

Definição 5.2.

> *Se a função f admitir derivada parcial em ordem a x_i, num dado conjunto S do seu domínio, então a cada ponto de S pode associar-se um número real, a derivada parcial da função em ordem a x_i nesse ponto. Fica assim definida em S uma nova função que se chama função derivada parcial em ordem a x_i e se representa por*
>
> $$\frac{\partial f}{\partial x_i}, \quad f'_{x_i}, \quad D_{x_i} f.$$

De modo semelhante se definem as funções derivadas parciais de f em ordem às demais variáveis independentes $x_1, \ldots, x_n$.

Pode porém acontecer que as novas funções assim definidas admitam também derivadas parciais. Assim, por exemplo, se a função $\dfrac{\partial f}{\partial x_i}$ admitir derivada parcial em ordem a x_k no ponto a, essa derivada chama-se **segunda derivada** parcial (ou derivada parcial de 2ª ordem) de f em ordem primeiro a x_i e depois a x_k no ponto a, a qual se designa por qualquer das formas

$$\frac{\partial^2 f}{\partial x_i \partial x_k}(a), \quad f''_{x_i x_k}(a), \quad (D^2_{x_i x_k} f)_a.$$

Se f admitir derivadas parciais de 2ª ordem em relação a x_i e x_k em todos os pontos dum conjunto S então fica definida nesse conjunto uma nova função que se chama **função segunda derivada** parcial em ordem a x_i e x_k e representa-se por

$$\frac{\partial^2 f}{\partial x_i \partial x_k}, \quad f''_{x_i x_k}, \quad D^2_{x_i x_k} f.$$

De forma análoga se definem as derivadas parciais de 2ª ordem em relação a qualquer outro par de variáveis independentes (ou duas vezes em relação a uma mesma variável).

A partir das derivadas parciais de segunda ordem definem-se as derivadas de 3ª ordem e sucessivamente as de ordem superior.

Exemplo 5.1. Determinar as derivadas parciais de 2ª ordem da função $z = f(x, y) = e^{xy} + \operatorname{sen} x$.

Resolução. A derivada parcial em ordem a x é

$$\frac{\partial z}{\partial x} = y e^{xy} + \cos x,$$

e a derivada parcial em ordem a y é

$$\frac{\partial z}{\partial y} = x e^{xy}.$$

Derivando agora cada uma destas funções em ordem a x e a y obtêm-se as derivadas de 2ª ordem pretendidas

$$\frac{\partial^2 z}{\partial x \partial x} = \frac{\partial^2 z}{\partial x^2} = y^2 e^{xy} - \operatorname{sen} x; \quad \frac{\partial^2 z}{\partial x \partial y} = e^{xy} + xye^{xy};$$
$$\frac{\partial^2 z}{\partial y \partial x} = e^{xy} + xye^{xy}; \quad \frac{\partial^2 z}{\partial y \partial y} = \frac{\partial^2 z}{\partial y^2} = x^2 e^{xy}. \quad \blacktriangleleft$$

Como facilmente se pode verificar, uma função de duas variáveis que seja n vezes derivável admite 2^n derivadas parciais de ordem n.

Na prática, porém, muitas delas são iguais entre si, como o mostra o exemplo anterior, em que

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$$

isto é, são iguais as derivadas de 2ª ordem em relação a x e y, independentemente da ordem de derivação. A partir deste exemplo pode-se perguntar em que condições é que a ordem de derivação é indiferente.

No caso das derivadas de 2ª ordem de funções de 2 variáveis a resposta é dada pelo teorema seguinte:

Teorema 5.1 (Teorema de Schwarz). *Se as derivadas f'_x, f'_y e f''_{xy} existirem numa vizinhança do ponto (a, b) e a 2ª derivada f''_{xy} for contínua nesse ponto, então também existe f''_{yx} nesse ponto e é $f''_{xy}(a, b) = f''_{yx}(a, b)$.*

Demonstração. De acordo com a definição de derivada, $f''_{yx}(a,b)$, se existir, será dada pela expressão seguinte, caso os limites existam

$$
\begin{aligned}
f''_{yx}(a,b) &= \lim_{h\to 0} \frac{f'_y(a+h,b) - f'_y(a,b)}{h} = \\
&= \lim_{h\to 0} \left\{ \frac{1}{h} \left[\lim_{k\to 0} \frac{f(a+h,b+k) - f(a+h,b)}{k} - \right.\right. \\
&\qquad \left.\left. - \lim_{k\to 0} \frac{f(a,b+k) - f(a,b)}{k} \right] \right\}
\end{aligned}
\tag{5.1}
$$

Com efeito, pondo

$$\Delta = [f(a+h,b+k) - f(a+h,b)] - [f(a,b+k) - f(a,b)]$$

a expressão anterior toma a forma

$$f''_{xy}(a,b) = \lim_{h\to 0} \frac{1}{h} \cdot \lim_{k\to 0} \frac{1}{k} \cdot \Delta.$$

O limite em relação a k existe visto que, por hipótese, existe f'_y numa vizinhança de (a,b). Então, como existe o limite em k o limite sucessivo

$$\lim_{h\to 0} \frac{1}{h} \cdot \lim_{k\to 0} \frac{\Delta}{k}$$

existirá caso exista o limite

$$\lim_{\substack{h\to 0 \\ k\to 0}} \frac{\Delta}{hk}$$

e, ambos terão então o mesmo valor.

Pondo $\mu(x) = f(x, b+k) - f(x,b)$ e aplicando o teorema dos acréscimos finitos para funções de uma variável real, tem-se

$$\begin{aligned} \Delta &= \mu(a+b) - \mu(a) = h\mu'(a+\theta_1 h) = \\ &= h[f'_x(a+\theta_1 h, b+k) - f'_x(a+\theta_1 h, b)] = \\ &= hkf''_{xy}(a+\theta_1 h, b+\theta_2 k) \end{aligned}$$

com $0 < \theta_1 < 1$ e $0 < \theta_2 < 1$.

Como f''_{xy} é, por hipótese, contínua em (a,b), vem

$$\lim_{\substack{h\to 0 \\ k\to 0}} \frac{\Delta}{hk} = \lim_{\substack{h\to 0 \\ k\to 0}} f''_{xy}(a+\theta_1 h, b+\theta_2 k) = f''_{xy}(a,b)$$

o que prova que existe $f''_{yx}(a,b)$ e é $f''_{yx}(a,b) = f''_{xy}(a,b)$. ■

Se as derivadas de $f(x,y)$ de ordem superior à segunda forem contínuas, o teorema anterior aplica-se a essas derivadas. Assim, se as derivadas f'''_{xxy}, f'''_{xyx} e f'''_{yxx} existirem e forem contínuas, então são iguais entre si.

O teorema extende-se a funções de mais de 2 variáveis. Assim, se a função $f(x_1, \ldots, x_n)$ pode ser derivada k_1 vezes em ordem a x_1, k_2 vezes em ordem a $x_2, \ldots,$ k_n vezes em ordem a x_n e essas derivadas são contínuas, então é indiferente a ordem segundo a qual se fazem as derivações e o seu valor comum pode ser representado por

$$\frac{\partial^{k_1+k_2+\ldots+k_n} f}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}}.$$

5.3 Funções diferenciáveis

Definição 5.3.

Uma função $f(x_1, \ldots, x_n)$ *diz-se* ***diferenciável*** *num ponto interior* $a = (a_1, \ldots, a_n)$ *do seu domínio se existe uma vizinhança desse ponto na qual seja*

$$\begin{aligned} & f(a_1 + \Delta x_1, \ldots, a_n + \Delta x_n) - f(a_1, \ldots, a_n) = \\ & = k_1 \Delta x_1 + \ldots k_n \Delta x_n + \epsilon_1 \Delta x_1 + \ldots + \epsilon_n \Delta x_n \end{aligned} \tag{5.2}$$

em que $k_1, \ldots, k_n$ *são números reais independentes de* $\Delta x_1, \ldots, \Delta x_n$ *e* $\epsilon_1, \ldots, \epsilon_n$ *são funções de* $\Delta x_1, \ldots, \Delta x_n$ *que tendem para zero quando*

$$\Delta x = (\Delta x_1, \ldots, \Delta x_n) \to (0, \ldots, 0).$$

Da definição resulta imediatamente do seguinte teorema.

Teorema 5.2. *Se* $f(x)$ *é diferenciável no ponto* a *então é contínua nesse ponto e admite nele derivadas parciais que são*

$$f'_{x_1}(a) = k_1, \; f'_{x_2}(a) = k_2, \ldots, f'_{x_n}(a) = k_n.$$

Demonstração. Com efeito, da igualdade (5.2) que traduz a diferenciabilidade da função resulta imediatamente que

$$\lim_{\substack{\Delta x_1 \to 0 \\ \cdots \\ \Delta x_n \to 0}} f(a_1 + \Delta x_1, \ldots, a_n + \Delta x_n) = f(a_1, \ldots, a_n)$$

o que prova que a continuidade da função no ponto $a = (a_1, \ldots, a_n)$.

No que se refere à diferenciabilidade, para a derivada em ordem a x_i, por exemplo, tem-se

$$\begin{aligned} f'_{x_i}(a) &= \lim_{\Delta x_i \to 0} \frac{f(a_1, \ldots, a_i + \Delta x_i, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{\Delta x_i} \\ &= \lim_{\Delta x_i \to 0} \frac{k_i \Delta x_i + \epsilon_i \Delta x_i}{\Delta x_i} = \\ &= k_i \end{aligned}$$

tal como se pretendia provar. ■

O teorema põe em evidência que a diferenciabilidade é uma condição mais forte do que a simples existência de derivadas parciais, pois que, enquanto esta não garante, como se viu anteriormente, a continuidade da função, a diferenciabilidade implica tanto a continuidade como a existência das derivadas parciais.

De acordo com o teorema anterior a expressão da diferenciabilidade pode, portanto, tomar a forma

$$\begin{aligned} f(a_1 + \Delta x_1, \ldots, a_n + \Delta x_n) - f(a_1, \ldots, a_n) = \\ = f'_{x_1}(a)\Delta x_1 + \ldots + f'_{x_n}(a)\Delta x_n + \epsilon_1 \Delta x_1 + \ldots + \epsilon_n \Delta x_n \end{aligned}$$

ou ainda, uma vez que $a = (a_1, \ldots, a_n)$ e $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$, tem-se:

$$f(a + \Delta x) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\Delta x_i + \sum_{i=1}^{n} \epsilon_i \Delta x_i.$$

O recíproco do teorema anterior é, todavia, falso. Isto é, o facto de uma função ser contínua e admitir derivadas parciais de primeira ordem num dado ponto não implica que a função seja diferenciável nesse ponto. É o que acontece, por exemplo, com a função

$$f(x,y) = +\sqrt{|xy|}$$

que é contínua e tem derivadas parciais na origem e todavia não é diferenciável nesse ponto.

O teorema seguinte, que se enuncia sem demonstração estabelece, porém, uma condição suficiente de diferenciabilidade.

Teorema 5.3. *É condição suficiente para que uma função de n variáveis seja diferenciável num dado ponto que nesse ponto admita derivadas parciais de primeira ordem e $n-1$ dessas derivadas sejam contínuas.*

Embora suficiente, esta condição não é todavia necessária. Por exemplo, a função

$$f(x,y) = \begin{cases} (x+y)^2 \operatorname{sen} \dfrac{1}{x+y}, & \text{se } (x,y) \neq (0,0) \\ \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$$

tem ambas as primeiras derivadas parciais descontínuas na origem e, no entanto, é diferenciável nesse ponto.
Diz-se que uma função f é **continuamente derivável** ou de classe C^1 num ponto (ou num conjunto) quando admite derivadas parciais contínuas nesse ponto (ou conjunto). Em geral, diz-se que a função f é da classe C^n num dado conjunto quando nele admite derivadas parciais contínuas até a ordem n.

Do Teorema 5.3 decorre que **toda a função de classe C^n é diferenciável** (e portanto contínua).

5.4 Diferencial de uma função

Como se viu no parágrafo anterior, se uma função f é diferenciável num dado ponto a do seu domínio, então existe uma vizinhança desse ponto na qual se verifica a relação

$$f(a + \Delta x) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\Delta x_i + \sum_{i=1}^{n} \epsilon_i \Delta x_i.$$

O primeiro membro desta igualdade representa o acréscimo sofrido pela função f ao passar do ponto a para $a + \Delta x$ e denomina-se simplesmente **acréscimo da função** entre a e $a + \Delta x$.

Tal como o indica a própria igualdade, esse acréscimo é numericamente igual à soma de duas partes

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\Delta x_i + \sum_{i=1}^{n} \epsilon_i \Delta x_i.$$

Como os ϵ_i são, por definição, infinitésimos simultâneos como os Δx_i, então, para valores suficientemente pequenos de Δx (e portanto dos Δx_i) a parte $\sum_{i=1}^{n} \epsilon_i \Delta x_i$ pode tornar-se tão pequena quanto se queira. Isso significa que, para valores suficientemente pequenos de $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$, a expressão

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\Delta x_i \tag{5.3}$$

dá boas aproximações do valor do acréscimo da função, isto é

$$f(a + \Delta x) - f(a) \approx \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\Delta x_i. \tag{5.4}$$

A expressão 5.3 denomina-se o **diferencial total** da função f no ponto a e costuma designar-se por df, escrevendo-se

$$df(a) = \frac{\partial f}{\partial x_1}(a)\Delta x_1 + \ldots + \frac{\partial f}{\partial x_n}(a)\Delta x_n.$$

Como para as funções

$$g_i(x_1, \ldots, x_i, \ldots, x_n) = x_i, \quad i = 1, 2, \ldots, n$$

é

$$dg_i = dx_i = \Delta x_i,$$

a expressão anterior pode escrever-se (abstraindo do ponto a para maior simplicidade)

$$df = \frac{\partial f}{\partial x_1}dx_1 + \ldots + \frac{\partial f}{\partial x_n}dx_n.$$

E pondo $\Delta f = f(a + \Delta x) - f(a)$, a expressão 5.4 pode portanto escrever-se $\Delta f \approx df$ o que traduz o facto já referido de que o diferencial df é o **valor principal** do acréscimo Δf da função (subentende-se que é num mesmo ponto e para os mesmos acréscimos das variáveis independentes).

Exemplo 5.2. Seja $f(x, y) = \dfrac{x}{y}$. Então, o diferencial total da função num ponto genérico é

$$df = \frac{1}{y}dx - \frac{x}{y^2}dy \quad \text{ou} \quad df = y^{-1}dx - xy^{-2}dy. \quad \blacktriangleleft$$

Exemplo 5.3. Seja $z = x^2 - y^2$. Então, $dz = 2xdx - 2ydy$. ◀

5.5 Diferenciais de ordem superior

Seja $z = f(x, y)$ uma função de duas variáveis que admite derivadas parciais contínuas até uma determinada ordem $n > 1$. O diferencial da função é neste caso

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

e depende tanto das variáveis x e y como dos parâmetros dx e dy.

Se se fixarem dx e dy, o diferencial dz passa a ser função apenas de x e y. Ao diferencial desta função chama-se **diferencial de 2ª ordem** da função $z = f(x, y)$ e representa-se por

$$d^2z = d(dz).$$

Assim, tem-se então que,

$$\begin{aligned}
d^2z &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right)dx + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right)dy = \\
&= \left(\frac{\partial^2 f}{\partial x^2}dx + \frac{\partial^2 f}{\partial y \partial x}dy\right)dx + \left(\frac{\partial^2 f}{\partial x \partial y}dx + \frac{\partial^2 f}{\partial y^2}dy\right)dy = \\
&= \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2
\end{aligned}$$

onde dx^2 e dy^2 designam, respectivamente, $(dx)^2$ e $(dy)^2$ e não $d(x^2)$ e $d(y^2)$.

De forma análoga se define o **diferencial de 3ª ordem**

$$d^3z = d(d^2z),$$

obtendo-se

$$d^3z = \frac{\partial^3 f}{\partial x^3}dx^3 + 3\frac{\partial^3 f}{\partial x^2 \partial y}dx^2dy + 3\frac{\partial^3 f}{\partial x \partial y^2}dxdy^2 + \frac{\partial^3 f}{\partial y^3}dy^3.$$

Agora consideremos que os diferencias $d^4z,\ d^5z, \ldots,\ d^kz$ definem-se de modo análogo e que pode verificar-se por indução que as expressões se podem obter da igualdade

$$d^kz = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^k z$$

onde o operador

$$\left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^k$$

é uma potência simbólica em cujo desenvolvimento um produto da forma $\left(\frac{\partial}{\partial x}\right)^r \cdot \left(\frac{\partial}{\partial y}\right)^s$, com $r+s=k$, significa $\frac{\partial^k}{\partial x^r \partial y^s}$.

Assim, as definições anteriores generalizam-se às funções com mais de duas variáveis independentes. Donde, se for $z = f(x_1, \ldots, x_n)$, o diferencial de ordem k da função é

$$d^kz = \left(\frac{\partial}{\partial x_1}dx_1 + \ldots + \frac{\partial}{\partial x_n}dx_n\right)^k z.$$

5.6 Derivadas e diferenciais de funções compostas

Teorema 5.4. *Seja $F(t)$ a função que resulta de compor $w = f(x,y)$ com $x = x(t)$ e $y = f(t)$. Se $x(t)$ e $y(t)$ são diferenciáveis no ponto t_0 e $w = f(x,y)$ é diferenciável no correspondente ponto $(x_0,y_0) = (x(t_0), y(t_0))$, então a função composta $F(t) = f(x(t), y(t))$ tem também derivada em t_0 que é dada pela expressão*

$$F'(t_0) = \frac{dw}{dt}(t_0) = \frac{\partial f}{\partial x}(x_0,y_0) \cdot \frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(x_0,y_0) \cdot \frac{dy}{dt}(t_0).$$

Demonstração. Sejam

$$\Delta x = x(t_0 + \Delta t) - x(t_0) \quad \text{e} \quad \Delta y = y(t_0 + \Delta t) - y(t_0).$$

Então, tem-se

$$\begin{aligned} \Delta w &= F(t_0 + \Delta t) - F(t_0) = \\ &= f(x(t_0 + \Delta t), y(t_0 + \Delta t)) - f(x(t_0), y(t_0)) = \\ &= f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0). \end{aligned}$$

Mas como, por hipótese, $f(x,y)$ é diferenciável em (x_0,y_0), vem:

$$F(t_0+\Delta t)-F(t_0) = \frac{\partial f}{\partial x}(x_0,y_0)\cdot\Delta x+\frac{\partial f}{\partial y}(x_0,y_0)\cdot\Delta y+\epsilon_1\Delta x+\epsilon_2\Delta y =$$

$$= \frac{\partial f}{\partial x}(x_0,y_0)[x(t_0 + \Delta t) - x(t_0)] + \frac{\partial f}{\partial y}(x_0,y_0)[y(t_0 + \Delta t) - y(t_0)]+$$
$$+\epsilon_1[x(t_0 + \Delta t) - x(t_0)] + \epsilon_2[y(t_0 + \Delta t) - y(t_0)].$$

Dividindo ambos os membros da igualdade por Δt e tomando limites quando $\Delta t \to 0$, tem-se,

$$F'(t_0) = \frac{\partial f}{\partial x}(x_0, y_0) \cdot \frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{dy}{dt}(t_0)$$

visto que, se $x(t)$ e $y(t)$ são diferenciáveis em t_0, são contínuas nesse ponto e, portanto, quando $\Delta t \to 0$ também Δx, $\Delta y \to 0$, por consequência, $\epsilon_1, \epsilon_2 \to 0$. ■

Em geral, abstraindo dos pontos considerados, a regra de derivação dada pelo teorema traduz-se abreviadamente pela fórmula

$$\frac{dw}{dt} = \frac{dF}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Raciocinando de forma análoga, a regra pode estender-se ao caso da composição de uma função de n variáveis $w = f(x_1, \ldots, x_n)$ com n funções de uma variável t,

$$x_1 = x_1(t),\ x_2 = x_2(t),\ \ldots,\ x_n = x_n(t)$$

desde que sejam satisfeitas as condições do Teorema 5.4.

A expressão da derivada é então

$$\frac{dw}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \ldots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}.$$

Exemplo 5.4. Determinar a derivada da função composta de $w = x^2y + xy^3$ com $x = t^2 + 3t$ e $y = e^t$ sen t.

Resolução. Como

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$

calculando as derivadas envolvidas na fórmula obtém-se

$$\frac{dw}{dt} = (2xy + y^3)(2t + 3) + (x^2 + 3xy^2)(e^t \text{ sen } t + e^t \cos t). \quad \blacktriangleleft$$

Teorema 5.5. *Seja $F(r, s)$ a função que resulta da composição da função $w = f(x, y, z)$, definida e diferenciável num domínio D de $\mathbb{R}^3$, com as funções $x = g_1(r, s)$, $y = g_2(r, s)$, $z = g_3(r, s)$ definidas e diferenciáveis num domínio D_1 de $\mathbb{R}$ e que a cada ponto de D_1 fazem corresponder um ponto de D. Então a função composta tem derivadas parciais em relação às variáveis r e s que são dadas pelas expressões*

$$\frac{\partial w}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial r}$$

$$\frac{\partial w}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s}.$$

Demonstração. Com efeito, se $w = f(x, y, z)$ é diferenciável, tem-se

$$\Delta w = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z + \epsilon_1\Delta x + \epsilon_2\Delta y + \epsilon_3\Delta z$$

com $\epsilon_1, \epsilon_2, \epsilon_3 \to 0$ quando $\Delta x, \Delta y, \Delta z \to 0$.

Ora, dividindo ambos os membros da igualdade por Δr tem-se:

$$\frac{\Delta w}{\Delta r}=\frac{\partial f}{\partial x}\frac{\Delta x}{\Delta r}+\frac{\partial f}{\partial y}\frac{\Delta y}{\Delta r}+\frac{\partial f}{\partial z}\frac{\Delta z}{\Delta r}+\epsilon_1\frac{\Delta x}{\Delta r}+\epsilon_2\frac{\Delta y}{\Delta r}+\epsilon_3\frac{\Delta z}{\Delta r}$$

e tomando limites quando $\Delta r\to 0$

$$\frac{\partial w}{\partial r}=\frac{\partial f}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial r}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial r}.$$

De forma análoga pode se deduzir a fórmula correspondente à derivada $\dfrac{\partial w}{\partial s}$. ■

A regra traduzida pelo teorema anterior (Teorema 5.5), conhecida por "**Regra da Cadeia**", estende-se ao caso mais geral da composta de uma função de n variáveis $w=f(x_1,x_2,\ldots,x_n)$ definida e diferenciável num domínio D de $\mathbb{R}^n$ com n funções de k variáveis independentes $x_i=g_i(r_1,r_2,\ldots,r_k)$, $i=1,2,\ldots,n$, definidas e diferenciáveis num domínio D_1 de $\mathbb{R}^k$ e tais que fazem corresponder a cada ponto de D_1 um ponto de D, de modo que

$$w=f(g_1,g_2,\ldots,g_n)=F(r_1,r_2,\ldots,r_k).$$

Nestas condições, o mesmo raciocínio feito anteriormente conduz às expressões

$$\begin{aligned}
\frac{\partial w}{\partial r_1} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial r_1}+\frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial r_1}+\ldots+\frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial r_1}\\
\ldots & \ldots\ \ldots\\
\frac{\partial w}{\partial r_k} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial r_k}+\frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial r_k}+\ldots+\frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial r_k}.
\end{aligned}$$

Ora se se multiplicar a primeira destas equações por dr_1, a segunda por dr_2 e, de uma forma geral, a de ordem i por dr_i $(i = 1, 2, \ldots, n)$ e em seguida somar se as equações obtidas resulta

$$\frac{\partial w}{\partial r_1} dr_1 + \ldots + \frac{\partial w}{\partial r_k} dr_k = dw = \frac{\partial f}{\partial x_1} \left(\frac{\partial x_1}{\partial r_1} dr_1 + \ldots + \frac{\partial x_1}{\partial r_k} dr_k \right) +$$

$$+ \ldots + \frac{\partial f}{\partial x_n} \left(\frac{\partial x_n}{\partial r_1} dr_1 + \ldots + \frac{\partial x_n}{\partial r_k} dr_k \right) = \frac{\partial f}{\partial x_1} dx_1 + \ldots + \frac{\partial f}{\partial x_n} dx_n,$$

o que mostra que o diferencial de w em ordem às variáveis finais $r_1, r_2, \ldots, r_k$ se pode obter formando o diferencial de w em ordem às variáveis intermediárias $x_1, x_2, \ldots, x_n$ e exprimindo em seguida os diferenciais destas (isto é, $dx_1, dx_2, \ldots, dx_n$) em função das variáveis finais.

Exemplo 5.5. Sendo

$$w = x^2 + y^2 + z^2 \text{ e } x = r^2 + s^2, \; y = r^2 + s, z = rs,$$

determinar $\dfrac{\partial w}{\partial r}$ e $\dfrac{\partial w}{\partial s}$

Resolução.

$$\frac{\partial w}{\partial r} = (2x)(3r^2) + (2y)(2r) + (2z)(s);$$

$$\frac{\partial w}{\partial s} = (2x)(2s) + (2y)(1) + (2z)(r). \blacktriangleleft$$

5.7 Transformações. Campos escalares e campos vectoriais

5.7.1 Definições

Até agora se consideraram aplicações do tipo $f : D \subset \mathbb{R}^n \to \mathbb{R}$, isto é, aplicações que a cada ponto de $\mathbb{R}^n$ fazem corresponder um escalar (número real). Tais aplicações denominam-se aplicações escalares e diz-se então que f define um **campo escalar**.

No caso mais geral em que se consideram dois espaços euclidianos $\mathbb{R}^n$ e $\mathbb{R}^k$, com $n > 1$ e $k > 1$, as aplicações $f : D \subset \mathbb{R}^n \to \mathbb{R}^k$ designam-se genericamente por **transformações** (do conjunto $D \subset \mathbb{R}^n$ com valores em $\mathbb{R}^k$).

Como os elementos de $\mathbb{R}^n$ e de $\mathbb{R}^k$ se podem considerar quer como pontos quer como vectores, diversas situações podem ocorrer. Assim:

- se f aplica pontos de $\mathbb{R}^n$ em pontos de $\mathbb{R}^k$ designa-se por **transformação pontual**;
- se f aplica vectores de $\mathbb{R}^n$ em vectores de $\mathbb{R}^k$, designa-se por **transformação vectorial**;
- se f aplica pontos de $\mathbb{R}^n$ em vectores de $\mathbb{R}^k$, designa-se por **campo vectorial**.

Sendo este último caso o que se irá doravante considerar.

Se $\overrightarrow{f} : D \subset \mathbb{R}^n \to \mathbb{R}^k$ é um campo vectorial, então a imagem através de $\overrightarrow{f}$ de cada ponto $x = (x_1, x_2, \ldots, x_n) \in D$ é um vector de $\mathbb{R}^k$ de componentes $f_1(x), f_2(x), \ldots, f_k(x)$. Sendo $\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_k}$

uma base ortonormada de $\mathbb{R}^k$ tem-se:

$$\overrightarrow{f}(x) = f_1(x)\overrightarrow{e}_1 + \ldots + f_k(x)\overrightarrow{e}_k.$$

A definição do campo vectorial $\overrightarrow{f}$ exige, portanto, a definição de k funções reais de n variáveis reais tais que,

$$\begin{array}{rcl} y_1 & = & f_1(x_1, \ldots, x_n) \\ y_2 & = & f_2(x_1, \ldots, x_n) \\ & & \ldots \\ y_k & = & f_k(x_1, \ldots, x_n). \end{array}$$

E o campo vectorial pode, pois, designar-se por

$$(y_1, \ldots, y_k) = (f_1(x), \ldots, f_k(x)) \quad \text{ou} \quad \overrightarrow{y} = \overrightarrow{f}(x)$$

ou ainda

$$\overrightarrow{f} = (f_1, \ldots, f_k).$$

Para estas funções os conceitos de limite, de continuidade, de derivada e de integral definem-se como habitualmente.

Por outro lado, é fácil de verificar que uma função vectorial $\overrightarrow{f}$ é contínua, derivável ou integrável num dado domínio se e só se o forem todas as suas componentes $f_1, f_2, \ldots, f_k$.

5.7.2 O vector simbólico ∇ (nabla ou del)

No espaço normado $\mathbb{R}^n$ chama-se operador **nabla** ou **del** ao vector simbólico,

$$\nabla = \overrightarrow{e}_1 \frac{\partial}{\partial x_1} + \overrightarrow{e}_2 \frac{\partial}{\partial x_2} + \ldots + \overrightarrow{e}_n \frac{\partial}{\partial x_n}.$$

No caso particular de $\mathbb{R}^3$ o vector escreve-se,

$$\nabla = \overrightarrow{i} \frac{\partial}{\partial x} + \overrightarrow{j} \frac{\partial}{\partial y} + \overrightarrow{k} \frac{\partial}{\partial z}.$$

5.7.3 Gradiente

Se $f(x_1, x_2, \ldots, x_n)$ é uma função escalar de classe C^1 num dado domínio D de $\mathbb{R}^n$, chama-se **gradiente** da função (ou do campo escalar por ela definido) num ponto genérico $x \in D$ ao vector

$$\text{grad} f = \frac{\partial}{\partial x_1} \overrightarrow{e}_1 + \frac{\partial}{\partial x_2} \overrightarrow{e}_2 + \ldots + \frac{\partial}{\partial x_n} \overrightarrow{e}_n.$$

Deste modo, a cada ponto de D fica a corresponder um vector e por isso se pode dizer que o gradiente é um operador que transforma campos escalares em campos vectoriais.

Como se pode ver (vd. 5.7.2)

$$\text{grad} f = \nabla f.$$

Se $f(x_1, \ldots, x_n)$ e $g(x_1, \ldots, x_n)$ definem dois campos escalares de classe C^1 num mesmo domínio D de $\mathbb{R}^n$ é fácil de verificar que:

(a) grad $(f+g) =$ grag $f +$ grad g

(b) grad $(k \cdot f) = k \cdot$ grad f

(c) grad $(f \cdot g) = f$grad $g + g$grad f

(d) grad $\left(\dfrac{f}{g}\right) = \dfrac{g\text{grad } f - f\text{grad } g}{g^2} \quad (g \neq 0)$.

As duas primeiras propriedades mostram que o gradiente é um operador linear.

5.7.4 Divergência

Seja agora o campo vectorial $\overrightarrow{f} : D \subset \mathbb{R}^n \to \mathbb{R}^n$. Se

$$\overrightarrow{f} = (f_1, f_2, \ldots, f_n)$$

é de classe C^1, chama-se **divergência** de $\overrightarrow{f}$ num ponto genérico de D, e designa-se por $div\ \overrightarrow{f}$, ao escalar

$$div\, \overrightarrow{f} = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \ldots + \frac{\partial f_n}{\partial x_n}.$$

Trata-se, portanto, de um operador que transforma campos vectoriais em campos escalares . Utilizando o vector simbólico nabla tem-se

$$div\, \overrightarrow{f} = \nabla \cdot \overrightarrow{f}.$$

Se for $div \overrightarrow{f} = 0$ em todo o campo, este diz-se um **campo solenoidal**.

A divergência goza das seguintes propriedades:

- $div\ (\overrightarrow{f} + \overrightarrow{g}) = div \overrightarrow{f} + div \overrightarrow{g}$
- $div\ (k \overrightarrow{f}) = k\ div \overrightarrow{f}$.

Trata-se, portanto, dum operador linear.

Exemplo 5.6. Calcular a divergência do campo vectorial

$$\overrightarrow{v} = x^2 \overrightarrow{i} - xy \overrightarrow{j} + xyz \overrightarrow{k}.$$

Resolução. De acordo com a definição tem-se

$$div \overrightarrow{v} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 2x - x + xy = x + xy. \blacktriangleleft$$

5.7.5 Rotacional

No caso particular em que $\overrightarrow{f} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$, e $\overrightarrow{f}$ é de classe C^1, chama-se **rotacional** do campo $\overrightarrow{f} = (f_1, f_2, f_3)$ num ponto genérico de D ao vector definido pela forma

$$rot \overrightarrow{f} = \nabla \times \overrightarrow{f} = \begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ f_1 & f_2 & f_3 \end{vmatrix}.$$

O rotacional é pois um operador que transforma campos vectoriais em novos campos vectoriais.

Se $\overrightarrow{f}$ e $\overrightarrow{g}$ definem dois campos vectoriais de classe C^1 num mesmo domínio D de $\mathbb{R}^3$, então:

(a) $rot\ (\overrightarrow{f} + \overrightarrow{g}) = rot\,\overrightarrow{f} + rot\,\overrightarrow{g}$

(b) $rot\ (k \cdot \overrightarrow{f}) = k\ rot\,\overrightarrow{f}$,

o que indica que o rotacional é um operador linear.

Se for $rot\,\overrightarrow{f} = 0$ em todo o campo então este diz-se **irrotacional**.

Se grad$f$ for de classe C^1 num domínio D de $\mathbb{R}^3$ é $rot(\text{grad}f) = 0$.

Se $\overrightarrow{f} = (f_1, f_2, f_3)$ for de classe C^2 então $div(rot\,\overrightarrow{f}) = 0$.

Exemplo 5.7. Determinar o rotacional do campo

$$\overrightarrow{v} = 2xyz\,\overrightarrow{i} + x^2z\,\overrightarrow{j} + x^2y\,\overrightarrow{k}.$$

Resolução. De acordo com a definição, tem-se:

$$rot\,\overrightarrow{v} = \begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 2xyz & x^2z & x^2y \end{vmatrix} = 0. \quad \blacktriangleleft$$

5.7.6 Laplaciano

Definição 5.4.

*Chama-se **Laplaciano** ou operador de Laplace ao operador*

$$\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2}.$$

O Laplaciano pode também exprimir-se em função do operador nabla pela forma

$$\nabla^2 = \nabla \cdot \nabla.$$

Donde, dada uma função real $f(x_1, \ldots, x_n)$ de classe C^2 chama-se Laplaciano da função à expressão

$$\nabla^2 f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}.$$

Como facilmente se pode verificar é

$$\nabla^2 f = \nabla \cdot \nabla f = div(\text{grad } f).$$

A equação $\nabla^2 f = 0$ denomina-se **Equação de Laplace** .

Se num dado domínio for $\nabla^2 f = 0$, diz-se que f é **função harmónica** ou que satisfaz a equação de Laplace nesse domínio.

Se em vez de um campo escalar se considerar um campo vectorial $\overrightarrow{f} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$, de componentes (f_1, f_2, f_3), o laplaciano de $\overrightarrow{f}$ define-se como sendo o campo de componentes $(\nabla^2 f_1, \nabla^2 f_2, \nabla^2 f_3)$.

5.7.7 Jacobiano

Se $\overrightarrow{f} : D \subset \mathbb{R}^n \to \mathbb{R}^n$, com $\overrightarrow{f} = (f_1, f_2, \ldots, f_n)$ chama-se **Jacobiano** da função vectorial $\overrightarrow{f}$, ou do sistema de funções $f_1, f_2, \ldots, f_n$ e designa-se por

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)}$$

ao valor do determinante

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)} = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{vmatrix}$$

5.8 Derivada direccional

Dada uma função $f(x) = f(x_1, x_2, \ldots, x_n)$ de domínio D, sejam $a = (a_1, a_2, \ldots, a_n)$ um ponto interior de D,

$$\overrightarrow{v} = (v_1, v_2, \ldots, v_n) = (\cos\alpha_1, \cos\alpha_2, \ldots, \cos\alpha_n)$$

um vector unitário e t um escalar tal que o ponto $(a + t\overrightarrow{v})$ pertence a D:

- se em f se substituir x por $a + t\overrightarrow{v}$ obtêm-se uma função de t

$$h(t) = f(a + t\overrightarrow{v}) = f(a_1 + tv_1, \ldots, a_n + tv_n).$$

- Se esta função admitir derivada em ordem a t no ponto $t = 0$, a essa derivada chama-se a derivada da função $f(x)$ no ponto a segundo a direcção $\overrightarrow{v}$ ou, simplesmente, segundo o vector $\overrightarrow{v}$.

Isto é:

Definição 5.5 (Derivada direccional).

Chama-se derivada da função $f(x)$ no ponto a na direcção dum vector unitário $\overrightarrow{v}$ e representa-se por

$$f'_{\overrightarrow{v}}(a) \quad ou \quad \frac{\partial f}{\partial \overrightarrow{v}}(a),$$

ao limite seguinte, caso exista

$$\lim_{t\to 0} \frac{f(a + t\overrightarrow{v}) - f(a)}{t} = f'_{\overrightarrow{v}}(a) = \frac{\partial f}{\partial \overrightarrow{v}}(a).$$

Como se pode ver, a relação da definição pode também escrever-se:

$$\lim_{t\to 0} \frac{f(a_1 + tv_1, \ldots, a_n + tv_n) - f(a_1, \ldots, a_n)}{t} = f'_{\overrightarrow{v}}(a).$$

Se $\overrightarrow{v}$ for um dos versores dos eixos coordenados, por exemplo,

$$\overrightarrow{v} = \overrightarrow{e}_i = (0, \ldots, 0, 1, \ldots, 0),$$

então a derivada direccional segundo esse vector é:

$$\begin{aligned} f'_{\overrightarrow{e_i}}(a) &= \lim_{t\to 0} \frac{f(a_1,\ldots,a_{i-1},a_i+t,a_{i+1},\ldots,a_n) - f(a_1,\ldots,a_n)}{t} \\ &= \frac{\partial f}{\partial x_i}(a) \end{aligned}$$

o que mostra que a derivada segundo o versor do eixo dos x_i coincide com a derivada parcial em ordem a x_i.

Situação análoga ocorre com as derivadas segundo os versores dos demais eixos coordenados, o que permite concluir que **as derivadas parciais não são mais do que casos particulares das derivadas direccionais**.

Teorema 5.6. *Se $f(x) = f(x_1,\ldots,x_n)$ é diferenciável no ponto $a = (a_1,\ldots,a_n)$ então tem derivada nesse ponto segundo qualquer vector unitário*

$$\overrightarrow{v} = (v_1, v_2, \ldots, v_n) = (\cos\alpha_1, \cos\alpha_2, \ldots, \cos\alpha_n)$$

dada pela expressão

$$f'_{\overrightarrow{v}} = \frac{\partial f}{\partial \overrightarrow{v}} = \frac{\partial f}{\partial x_1}\cos\alpha_1 + \frac{\partial f}{\partial x_2}\cos\alpha_2 + \ldots + \frac{\partial f}{\partial x_n}\cos\alpha_n,$$

em que as derivadas parciais se supõem calculadas no ponto a.

Demonstração. De facto, atendendo à diferenciabilidade da função,

$$\lim_{t\to 0}\frac{f(a_1+t\cos\alpha_1,\ldots,a_n+t\cos\alpha_n)-f(a_1,\ldots,a_n)}{t}=$$

$$=\lim_{t\to 0}\frac{t\left(\dfrac{\partial f}{\partial x_1}\cos\alpha_1+\ldots+\dfrac{\partial f}{\partial x_n}\cos\alpha_n+\epsilon_1\cos\alpha_1+\ldots+\epsilon_n\cos\alpha_n\right)}{t}=$$

$$=\frac{\partial f}{\partial x_1}\cos\alpha_1+\frac{\partial f}{\partial x_2}\cos\alpha_2+\ldots+\frac{\partial f}{\partial x_n}\cos\alpha_n.\ \blacksquare$$

Se $\overrightarrow{v}$ não for um vector unitário, chama-se **derivada direccional** segundo $\overrightarrow{v}$ à derivada segundo o versor de $\overrightarrow{v}$. Nesse caso, se $\overrightarrow{v}=(v_1,v_2,\ldots,v_n)$ tem-se:

$$vers\ \overrightarrow{v}=\frac{\overrightarrow{v}}{||\overrightarrow{v}||}=\left(\frac{v_1}{||\overrightarrow{v}||},\frac{v_2}{||\overrightarrow{v}||},\ldots,\frac{v_n}{||\overrightarrow{v}||}\right)$$

e portanto, segue-se que,

$$\begin{aligned} f'_{vers\overrightarrow{v}} &= \frac{\partial f}{\partial x_1}\frac{v_1}{||\overrightarrow{v}||}+\ldots+\frac{\partial f}{\partial x_n}\frac{v_n}{||\overrightarrow{v}||} \\ &= \frac{1}{||\overrightarrow{v}||}\left(\frac{\partial f}{\partial x_1}v_1+\ldots+\frac{\partial f}{\partial x_n}v_n\right) \end{aligned}$$

de onde resulta

$$\frac{\partial f}{\partial x_1}v_1+\ldots+\frac{\partial f}{\partial x_n}v_n=||\overrightarrow{v}||f'_{vers\overrightarrow{v}}.$$

Exemplo 5.8. Calcular a derivada de $f(x, y, z) = x^2 + y^2 + z^2$ no ponto $(1, 2, -3)$ segundo o vector $\overrightarrow{v} = \overrightarrow{i} + 4\overrightarrow{j} - \overrightarrow{k}$.

Resolução. O vector dado não é unitário, pelo que deve ser substituído pelo seu versor, isto é, pelo vector unitário

$$\frac{\overrightarrow{v}}{||\overrightarrow{v}||} = \frac{\overrightarrow{i} + 4\overrightarrow{j} - \overrightarrow{k}}{\sqrt{18}}.$$

Os co-senos directores do vector dado são, portanto,

$$\cos\alpha_1 = \frac{1}{\sqrt{18}}; \quad \cos\alpha_2 = \frac{4}{\sqrt{18}}; \quad \cos\alpha_3 = \frac{-1}{\sqrt{18}}.$$

Então,

$$f'_{vers\overrightarrow{v}}(1, 2, -3) = \frac{2}{\sqrt{18}} + \frac{16}{\sqrt{18}} + \frac{6}{\sqrt{18}} = 4\sqrt{2}. \quad \blacktriangleleft$$

Como é fácil de verificar, a expressão do Teorema 5.6 é igual ao produto escalar do gradiente da função pelo vector unitário $\overrightarrow{v}$, o que permite interpretar a derivada direccional duma função como sendo o produto escalar do seu gradiente pelo vector unitário que determina a direcção segundo a qual se calcula a derivada, isto é

$$f'_{\overrightarrow{v}}(a) = \nabla f(a) \cdot \overrightarrow{v}.$$

Se θ for o ângulo formado pelos vectores gradiente e $\overrightarrow{v}$ (Figura 5.1) tem-se portanto

$$f'_{\overrightarrow{v}}(a) = ||\nabla f(a)|| \cdot ||\overrightarrow{v}|| \cos\theta = ||\nabla f(a)|| \cos\theta.$$

Desta relação conclui-se que a derivada direccional terá valor máximo quando $\theta = 0$, o que significa que a **direcção segundo a qual a taxa de variação da função f (que é dada por $f'_{\overrightarrow{v}}(a)$) é**

máxima é a da direcção do gradiente de f. É claro que a derivada direccional muda de sinal sempre que numa dada direcção se mudar de sentido, isto é,

$$f'_{(-\overrightarrow{v})}(a) = -f'_{\overrightarrow{v}}(a)$$

e se $\overrightarrow{v}$ não for unitário

$$f'_{vers(-\overrightarrow{v})}(a) = -f'_{vers\overrightarrow{v}}(a).$$

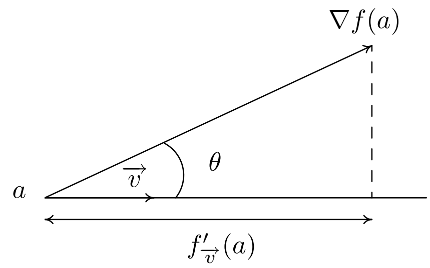

Figura 5.1:

5.9 Derivadas direccionais de ordem superior

Se a função f for diferenciável num dado conjunto, então, escolhido um vector unitário $\overrightarrow{v}$, é possível definir nesse conjunto uma nova função que a cada ponto do conjunto faz corresponder a derivada de f nesse ponto segundo o vector $\overrightarrow{v}$. A função assim definida chama-se função derivada direccional de f segundo $\overrightarrow{v}$ e representa-se por

$$f'_{\overrightarrow{v}} \quad \text{ou} \quad \frac{\partial f}{\partial \overrightarrow{v}},$$

escrevendo-se

$$f'_{\overrightarrow{v}} = \frac{\partial f}{\partial x_1}\cos\alpha_1 + \frac{\partial f}{\partial x_2}\cos\alpha_2 + \ldots + \frac{\partial f}{\partial x_n}\cos\alpha_n.$$

Se f for de classe C^2, então esta função é, por seu turno diferenciável, já que as derivadas parciais de f são, nessa hipótese, contínuas. Nessas condições, a função $f'_{\overrightarrow{v}}$ admite derivada segundo o vector $\overrightarrow{v}$, a qual se denomina segunda derivada de f segundo o vector $\overrightarrow{v}$ e se representa por

$$f''_{\overrightarrow{v}}, \quad \frac{\partial^2 f}{\partial \overrightarrow{v}^2} \quad \text{ou} \quad f^{(2)}_{\overrightarrow{v}}.$$

De acordo com o exposto em 5.8. depois do Exemplo 5.8, será:

$$f''_{\overrightarrow{v}} = \nabla f'_{\overrightarrow{v}} \cdot \overrightarrow{v}.$$

Efectuando o produto escalar tem-se

$$\begin{aligned} f''_{\overrightarrow{v}} &= \frac{\partial f'_{\overrightarrow{v}}}{\partial x_1}\cos\alpha_1 + \frac{\partial f'_{\overrightarrow{v}}}{\partial x_2}\cos\alpha_2 + \ldots + \frac{\partial f'_{\overrightarrow{v}}}{\partial x_n}\cos\alpha_n = \\ &= \left(\frac{\partial}{\partial x_1}\cos\alpha_1 + \frac{\partial}{\partial x_2}\cos\alpha_2 + \ldots + \frac{\partial}{\partial x_n}\cos\alpha_n\right)^2 f = \\ &= (\nabla \cdot \overrightarrow{v})^2 f; \end{aligned}$$

com as convenções habituais acerca das potências simbólicas.

De forma análoga se definem as derivadas de ordem superior à segunda.

Por indução pode mostrar-se que, se f é de classe C^n numa vizinhança do ponto a, então admite nesse ponto derivada direccional de ordem n segundo o vector $\overrightarrow{v}$ dada pela expressão,

$$f^{(n)}_{\overrightarrow{v}} = (\nabla \cdot \overrightarrow{v})^n f.$$

Como é fácil de verificar

$$f^{(n)}_{-\overrightarrow{v}}(a) = (-1)^n f^{(n)}_{\overrightarrow{v}}(a).$$

5.10 Funções homogéneas

Uma função $f(x) = f(x_1, \ldots, x_n)$ diz-se homogénea de grau de homogeneidade α quando verifica a condição,

$$f(tx_1, tx_2, \ldots, tx_n) = t^\alpha f(x_1, x_2, \ldots, x_n),$$

para α constante e todos os valores $t,\ x_1, x_2, \ldots, x_n$ tais que

$$(x_1, x_2, \ldots, x_n) \quad \text{e} \quad (tx_1, tx_2, \ldots, tx_n)$$

pertencem ao domínio da função.

Assim, por exemplo, a função

$$f(x, y) = (x^2 + y^2) \text{ arcsen } \left(\frac{y}{x}\right) \quad (x \neq 0)$$

é homogénea do grau 2 pois que,

$$f(tx, ty) = t^2(x^2 + y^2) \text{ arcsen } \left(\frac{y}{x}\right) = t^2 f(x, y).$$

Se se derivarem ambos os membros da relação da definição em ordem à variável x_i tem-se

$$t \cdot \frac{\partial f(tx)}{\partial x_i} = t^\alpha \cdot \frac{\partial f(x)}{\partial x_i}$$

ou seja:

$$\frac{\partial f(tx)}{\partial x_i} = t^{\alpha-1} \cdot \frac{\partial f(x)}{\partial x_i};$$

o que mostra que as derivadas parciais de primeira ordem de uma função homogénea de grau α são funções homogéneas de grau $\alpha - 1$. A aplicação reiterada desta propriedade leva à conclusão de que as derivadas parciais de ordem k de uma função homogénea de grau α são funções homogéneas de grau $\alpha - k$.

Teorema 5.7 (Teorema de Euler). *Se $f(x_1, \ldots, x_n)$ é homogénea de grau α então verifica a identidade*

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \ldots + x_n \frac{\partial f}{\partial x_n} = \alpha f(x_1, \ldots, x_n)$$

(denominada identidade de Euler) em todo o ponto em que seja diferenciável.

Demonstração. Fazendo $y_1 = tx_1, y_2 = tx_2, \ldots, y_n = tx_n$, visto f ser homogénea de grau, tem-se: α

$$f(y_1, y_2, \ldots, y_n) = t^\alpha f(x_1, x_2, \ldots, x_n)$$

em que o primeiro membro é função composta de $f(y_1, y_2, \ldots, y_n)$ com $y_1 = tx_1, \ldots, y_n = tx_n$. Donde, derivando a igualdade em ordem a t vem:

$$\frac{\partial f}{\partial y_1}\frac{dy_1}{dt} + \ldots + \frac{\partial f}{\partial y_n}\frac{dy_n}{dt} = \alpha t^{\alpha-1} f(x_1, \ldots, x_n)$$

ou seja

$$x_1 \frac{\partial f}{\partial y_1} + \ldots + x_n \frac{\partial f}{\partial y_n} = \alpha t^{\alpha-1} f(x_1, \ldots, x_n).$$

Para $t = 1$ resulta pois

$$x_1 \frac{\partial f}{\partial x_1} + \ldots + x_n \frac{\partial f}{\partial x_n} = \alpha f(x_1, \ldots, x_n)$$

visto então ser $y_1 = x_1, y_2 = x_2, \ldots, y_n = x_n$. ■

Pode também demonstrar-se que, reciprocamente, numa função diferenciável que verifique a identidade de Euler é homogénea de grau α. Se na relação que traduz a homogeneidade de f

$$f(tx_1, tx_2, \ldots, tx_n) = t^\alpha f(x_1, x_2, \ldots, x_n)$$

se fizer $t = \dfrac{1}{x_1}$, obtêm-se

$$f\left(1, \frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\right) = \frac{1}{x_1^\alpha} f(x_1, x_2, \ldots, x_n)$$

que se pode também escrever

$$f(x_1, x_2, \ldots, x_n) = x_1^\alpha g\left(\frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\right)$$

expressão que traduz outra propriedade importante das funções homogéneas.

Exemplo 5.9. Verificar se

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

é homogénea.

Resolução. Como

$$f(tx, yt) = t\frac{x^2}{\sqrt{x^2 + y^2}} = tf(x, y)$$

a função é homogénea de grau 1. ◀

5.11 Exercícios

1. Calcule as derivadas parciais de primeira ordem das funções

   (a) $z = x^y$;

   (b) $z = x$ sen xy.

2. Calcule as derivadas parciais de 2ª ordem das funções

   (a) $z = x + y + \dfrac{xy}{x-y}$;

   (b) $z = xe^y$.

3. Verifique que a função $z = \ln(e^x + e^y)$ é uma solução particular da equação $\dfrac{\partial z}{dx} + \dfrac{\partial z}{dy} = 1$.

4. Calcule o diferencial total das funções

   (a) $z = \operatorname{arctg}\left(\dfrac{x+y}{1-xy}\right)$;

   (b) $u = (x^2 + y^2 + z^2)^{1/2}$.

5. Calcule o diferencial de 2ª ordem das funções

   (a) $z = x + xy$;

   (b) $z = x$ sen ${}^2 y$.

6. Sendo $u = e^{ax+by}$, calcule $d^n u$.

7. Calcule as derivadas das funções compostas:

   (a) $z = e^{u-2v}$, $u =$ sen x, $v = x^3$;

(b) $z = \operatorname{arctg}\left(\dfrac{x}{y}\right)$, $x = u \operatorname{sen} v$, $y = u\cos v$;

(c) $u = \ln \operatorname{sen} \dfrac{x}{\sqrt{y}}$, $x = 3t^2$, $y = \sqrt{t^2+1}$;

8. Mostre que se $u = f(x,y)$ e $x = r\cos\theta$ e $y = r \operatorname{sen}\theta$, então

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2.$$

9. Sendo $f(x,y,z) = x^2yz^3$ e $\overrightarrow{v} = xz\overrightarrow{i} - y^2\overrightarrow{j} + 2x^2y\overrightarrow{k}$, calcule:

   (a) ∇f;

   (b) $\nabla \cdot \overrightarrow{v}$;

   (c) $\nabla \times \overrightarrow{v}$;

   (d) $div(f\overrightarrow{v})$.

10. Prove que:

   (a) $\nabla \cdot (\overrightarrow{u} + \overrightarrow{v}) = \nabla \cdot \overrightarrow{u} + \nabla \cdot \overrightarrow{v}$;

   (b) $rot(\text{grad} f) = 0$.

11. Calcule as derivadas direccionais das funções seguintes nos pontos e direcções indicadas:

   (a) $f(x,y,z) = 2x^2 - y^2 + z^2$ em $(1,2,3)$ e na direcção da recta que passa por $(1,2,3)$ e $(3,5,0)$;

   (b) $f(x,y) = e^x\cos y$ em $(0,0)$ na direcção que forma um ângulo de 60 graus com o eixo OX;

   (c) $f(x,y) = 2x - 3y$ em (1,1) ao longo da curva $y = x^2$ no sentido de x crescente;

   (d) $f(x,y) = \sqrt{x^2+y^2}$ num ponto genérico e segundo o raio vector $\overrightarrow{v} = x\overrightarrow{i} + y\overrightarrow{j}$.

12. Mostre que as funções seguintes são harmónicas:

    (a) $e^x \cos y$;

    (b) $\ln \sqrt{x^2 + y^2}$.

13. Verifique a fórmula de Euler para as funções:

    (a) $z = e^{x/y}$;

    (b) $z = \sqrt{x^2 + xy + y^2}$.

Capítulo 6

Teorema dos acréscimos finitos. Teorema de Taylor

6.1 Teorema dos acréscimos finitos

Para maior simplicidade na exposição, começar-se-á por considerar funções de apenas duas variáveis, isto é, funções reais definidas em $\mathbb{R}^2$.

Teorema 6.1. *Se $f(x,y)$ é diferenciável numa vizinhança do ponto (a,b), então, para qualquer ponto $(a+h,b+k)$ dessa vizinhança é:*

$$f(a+h,b+k)-f(a,b) = h\frac{\partial f}{\partial x}(a+\theta h,b+\theta k)+k\frac{\partial f}{\partial y}(a+\theta h,b+\theta k)$$

com $0<\theta<1$.

Demonstração. Pondo $x=a+ht$, $y=b+kt$, $t\in[0,1]$, $f(x,y)$ será função composta de t através de x e y

$$f(x,y)=f(a+ht,b+kt)=F(t) \quad \text{com} \quad t\in[0,1].$$

Como f é diferenciável na vizinhança considerada e x e y são funções contínuas e deriváveis de t, então $F(t)$ cumpre as condições do teorema dos acréscimos finitos (ou Lagrange) para funções de uma só variável real em $[0,1]$.

Portanto, $F(1) - F(0) = F'(\theta)$ com $0 < \theta < 1$.

Ora, como

$$F(1) = f(a+h, b+k), \quad F(0) = f(a,b)$$

e

$$F'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y},$$

tem-se, fazendo $t = \theta$

$$f(a+h, b+k) - f(a,b) = h\frac{\partial f}{\partial x}(a+\theta h, b+\theta k) + k\frac{\partial f}{\partial y}(a+\theta h, b+\theta k)$$

como se pretendia provar. ■

A extensão a $\mathbb{R}^n$ é imediata e traduz-se pelo enunciado do teorema seguinte:

Teorema 6.2. *Se $f(x) = f(x_1, \ldots, x_n)$ é diferenciável numa vizinhança do ponto*

$$a = (a_1, \ldots, a_n) \quad e \quad a+h = (a_1+h_1, a_2+h_2, \ldots, a_n+h_n)$$

é um ponto dessa vizinhança, então existe, no segmento que une a e $a+h$, um ponto $a+\theta h$ $(0 < \theta < 1)$ tal que:

$$f(a+h) - f(a) = h_1\frac{\partial f}{\partial x_1}(a+\theta h) + h_2\frac{\partial f}{\partial x_2}(a+\theta h) +$$

$$+ \ldots + h_n \frac{\partial f}{\partial x_n}(a + \theta h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a + \theta h).$$

Demonstração. Compondo f com $x_i = a_i + h_i t$ $(i = 1, 2, \ldots, n)$, $0 < t < 1$, obtém-se a função de t:

$$F(t) = f(a + ht) = f(a_1 + h_1 t, \ldots, a_n + h_n t)$$

definida no intervalo $[0, 1]$. Como tanto f como as componentes $x_i = a_i + h_i t$ são direfenciáveis, a função composta $F(t)$ é contínua e derivável em $[0, 1]$ e pode-se-lhe, portanto, aplicar o teorema dos acréscimos finitos (Teorema 6.3) de Lagrange , isto é:

$$F(1) - F(0) = F'(\theta) \quad \text{com} \quad 0 < \theta < 1.$$

Como

$$\begin{aligned} F(1) &= f(a + h), \\ F(0) &= f(a) \quad \text{e} \\ F'(t) &= h_1 \frac{\partial f}{\partial x_1}(a + ht) + \ldots + h_n \frac{\partial f}{\partial x_n}(a + ht), \end{aligned}$$

fazendo $t = \theta$ vem

$$f(a + h) - f(a) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a + \theta h). \quad \blacksquare$$

Donde, do teorema dos acréscimos finitos resulta imediatamente que:

Corolário 6.1. *Uma função com derivadas parciais de primeira ordem todas nulas num dado domínio é constante nesse domínio.*

Corolário 6.2. *Duas funções f e g com derivadas parciais de primeira ordem finitas e iguais num dado domínio diferem apenas por uma constante (nesse domínio).*

Corolário 6.3. *Uma função f é contínua num ponto a quando numa vizinhança desse ponto tiver derivadas parciais de primeira ordem limitadas.*

6.2 Teorema de Taylor

Em $\mathbb{R}^2$ o teorema de Taylor tem a formulação seguinte:

Teorema 6.3. *Se $f(x,y)$ e todas as suas derivadas parciais até à ordem $n+1$ são contínuas numa vizinhança do ponto (a,b), então o valor da função em qualquer ponto $(a+h, b+k)$ pertencente a essa vizinhança é dada pela fórmula*

$$f(a+h,b+k) = f(a,b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a,b) +$$

$$+\frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a,b) + \frac{1}{3!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^3 f(a,b) +$$

$$+ \ldots + \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a,b) +$$

$$+\frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n+1} f(a+\theta h, b+\theta k); \quad 0 < \theta < 1,$$

que se denomina ***fórmula de Taylor*** .

Demonstração. Pondo $x = a + ht$, $y = b + kt$, com $t \in [0, 1]$, tem-se

$$f(x, y) = f(a + ht, b + kt) = F(t), \quad t \in [0, 1].$$

Dadas as hipóteses do teorema, a função $F(t)$ pode exprimir-se em $[0, 1]$ através da fórmula da Mac-Laurin e portanto

$$\begin{aligned} F(1) &= F(0) + F'(0) + \frac{1}{2!}F''(0) + \ldots + \\ &+ \frac{1}{n!}F^{(n)}(0) + \frac{1}{(n+1)!}F^{(n+1)}(\theta) \end{aligned} \tag{6.1}$$

com $0 < \theta < 1$.

Como

$$\begin{aligned} F'(t) &= h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} \\ F''(t) &= \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f \\ &\cdots \\ F^{(n)}(t) &= \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f \end{aligned}$$

e, fazendo $t = 0$ e substituindo em (6.1) vem a expressão pretendida

$$f(a + h, b + k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a, b) +$$

$$+\frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \ldots + \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a, b) +$$

$$+\frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{n+1} f(a+\theta h, b+\theta k),$$

em que $0 < \theta < 1$.

O último termo da fórmula denomina-se o **resto de Lagrange** e designa-se abreviadamente por R_n. ■

A generalização da fórmula de Taylor para funções de mais de duas variáveis é imediata e traduz-se pelo teorema seguinte, cuja demonstração é análoga à anterior, pelo que se deixa como exercício.

Teorema 6.4. *Se $f(x_1, \ldots, x_n)$ é de classe C^{n+1} numa vizinhança do ponto $a = (a_1, \ldots . a_n)$ e $a + \overrightarrow{h}$, com $\overrightarrow{h} = (h_1, \ldots, h_n)$ pertence a essa vizinhança, então*

$$f(a+\overrightarrow{h}) = f(a) + \left(h_1\frac{\partial}{\partial x_1}+\ldots+h_n\frac{\partial}{\partial x_n}\right) f(a) +$$

$$+\frac{1}{2!}\left(h_1\frac{\partial}{\partial x_1}+\ldots+h_n\frac{\partial}{\partial x_n}\right)^2 f(a) + \ldots +$$

$$+\frac{1}{n!}\left(h_1\frac{\partial}{\partial x_1}+\ldots+h_n\frac{\partial}{\partial x_n}\right)^n f(a) +$$

$$+\frac{1}{(n+1)!}\left(h_1\frac{\partial}{\partial x_1}+\ldots+h_n\frac{\partial}{\partial x_n}\right)^{n+1} f(a+\theta\overrightarrow{h}), \quad 0<\theta<1$$

*que se denomina **Fórmula de Taylor** com resto de Lagrange.*

A fórmula de Taylor pode escrever-se de forma mais sintética utilizando derivadas direccionais. De facto, como foi visto no ponto 5.8 – Capítulo 5,

$$h_1\frac{\partial f}{\partial x_1}+\ldots+h_n\frac{\partial f}{\partial x_n}=||\overrightarrow{h}||f'_{vers\,\overrightarrow{h}},$$

e de uma forma geral

$$\left(h_1\frac{\partial f}{\partial x_1}+\ldots+h_n\frac{\partial f}{\partial x_n}\right)^n=||\overrightarrow{h}||^n f^{(n)}_{vers\,\overrightarrow{h}},$$

então, substituindo na expressão do Teorema 6.4 obtém-se:

$$f(a+\overrightarrow{h})=f(a)+||\overrightarrow{h}||f'_{vers\,\overrightarrow{h}}(a)+\frac{||\overrightarrow{h}||^2}{2!}f''_{vers\,\overrightarrow{h}}(a)+\ldots+$$

$$+\frac{||\overrightarrow{h}||^n}{n!}f^{(n)}_{vers\,\overrightarrow{h}}(a)+\frac{||\overrightarrow{h}||^{n+1}}{(n+1)!}f^{(n+1)}_{vers\,\overrightarrow{h}}(a+\theta\overrightarrow{h}).$$

No caso de ser $a=(0,0,\ldots,0)$ a fórmula denomina-se **Fórmula de Mac-Laurin.**

Se $f(x)$ for de classe C^∞ (isto é, indefinidamente derivável) e $\lim\limits_{n\to\infty} R_n=0$, a fórmula de Taylor converte-se na **Série de Taylor** (ou de Mac-Laurin se $a=0$).

Exemplo 6.1. Determinar o desenvolvimento pela fórmula de Taylor da função

$$f(x,y,z)=x^2+y^2+z^2+2xy-yz-4x-3y-z+4$$

numa vizinhança do ponto $(1,1,1)$.

Resolução. Como as derivadas parciais são todas nulas a partir da terceira ordem, é

$$f(1+h_1, 1+h_2, 1+h_3) = f(1,1,1)+\left(h_1\frac{\partial}{\partial x}+h_2\frac{\partial}{\partial y}+h_3\frac{\partial}{\partial z}\right)f(1,1,1)+$$

$$+\frac{1}{2!}\left(h_1\frac{\partial}{\partial x}+h_2\frac{\partial}{\partial y}+h_3\frac{\partial}{\partial z}\right)^2 f(1,1,1) = h_1^2+h_2^2+h_3^2+2h_1h_2-h_2h_3.$$

Fazendo:

$$\begin{aligned} 1+h_1 &= x \\ 1+h_2 &= y \\ 1+h_3 &= z, \end{aligned}$$

tem-se:

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

6.3 Exercícios

1. Determine o desenvolvimento de Taylor das seguintes funções nas condições indicadas

   (a) $f(x,y) = \cos x \cos y$, relativamente ao ponto $(0,0)$ e até o 3º termo;

   (b) $f(x,y) = \dfrac{1}{1+x+y}$, relativamente a $(0,0)$ e até o 15º termo;

   (c) $f(x,y) = x^y$, relativamente a $(1,0)$ e até o 3º termo.

2. Determine o desenvolvimento de Mac-Laurin das seguintes funções até ao termo de segunda ordem

   (a) $f(x,y) = \sqrt{1-x^2-y^2}$;

   (b) $f(x,y) = e^x \text{ sen } y$.

3. Determine o desenvolvimento de Mac-Laurin de

$$f(x,y) = x \text{ sen } y + y \text{ sen } x.$$

4. Determine o desenvolvimento de Mac-Laurin de

$$f(x,y) = \ln(1+xy).$$

5. Determine o desenvolvimento da função $f(x,y) = e^x \cos y$ em série de potências de h e k.

Capítulo 7

Extremos relativos

7.1 Extremos em pontos interiores

Seja $f : D \subset \mathbb{R}^n \to \mathbb{R}$ e a um ponto interior de D.

Definição 7.1.

> *Diz-se que a função f tem um **máximo relativo** (ou local) no ponto a quando existe uma vizinhança de a na qual se verifica a relação $f(x) \leq f(a)$ para todo o x de D situado nessa vizinhança.*

Definição 7.2.

> *Diz-se que a função f tem um **mínimo relativo** (ou local) em a se existe uma vizinhança de a na qual se verifica a relação $f(a) \leq f(x)$ para todo o $x \in D$ situado na vizinhança.*

Das definições 7.1 e 7.2 decorre que há um **extremo relativo** (máximo ou mínimo) no ponto a se e só se a diferença $f(x) - f(a)$ não muda de sinal numa vizinhança de a.

Na exposição que se segue estudar-se-á a existência de extremos relativos de funções a que seja aplicável o **Teorema de Taylor**.

Teorema 7.1. *É condição necessária (mas não suficiente) para que uma função diferenciável $f(x) = f(x_1, \ldots, x_n)$ tenha um extremo relativo num ponto interior $a = (a_1, \ldots, a_n)$ que sejam nulas nesse ponto todas as suas primeiras derivadas parciais.*

Demonstração. Para que haja um extremo em a a diferença $f(a+\overrightarrow{h}) - f(a)$ deve ter sinal constante nalguma vizinhança de a. Se, para um dado $\overrightarrow{h}$, fosse $f'_{vers\,\overrightarrow{h}}(a) \neq 0$ então, de acordo com o Teorema de Taylor,

$$f(a+\overrightarrow{h}) - f(a) = ||\overrightarrow{h}||f'_{vers\,\overrightarrow{h}}(a+\theta\overrightarrow{h}), \quad \text{com} \quad 0 < \theta < 1. \tag{7.1}$$

Como f é diferenciável no ponto a, a sua derivada segundo o versor de $\overrightarrow{h}$ é contínua em a e portanto

$$f'_{vers\,\overrightarrow{h}}(a+\theta\overrightarrow{h}) = f'_{vers\,\overrightarrow{h}}(a) + \alpha,$$

em que $\lim_{||\overrightarrow{h}|| \to 0} \alpha = 0$. A igualdade (7.1) toma então a forma

$$f(a+\overrightarrow{h}) - f(a) = ||\overrightarrow{h}||[f'_{vers\,\overrightarrow{h}}(a) + \alpha)].$$

Como α tende para zero quando $||\overrightarrow{h}|| \to 0$, é possível escolher $\overrightarrow{h}$ por forma a que $|\alpha| < |f'_{vers\,\overrightarrow{h}}(a)|$ e, consequentemente, o sinal de $f(a+\overrightarrow{h}) - f(a)$ seja o de $f'_{vers\,\overrightarrow{h}}(a)$.

Mas então, como a derivada direccional muda de sinal sempre que numa dada direcção se mudar de sentido, isto é

$$f'_{vers(-\overrightarrow{h})}(a) = -f'_{vers\,\overrightarrow{h}}(a),$$

não haverá vizinhança de a na qual $f(a + \overrightarrow{h}) - f(a)$ tenha sinal constante a não ser que $f'_{vers-\overrightarrow{h}}(a) = 0$ para todo o $\overrightarrow{h}$.

Logo, para haver extremo, deve necessariamente ser

$$\frac{\partial f}{\partial x_i}(a) = 0 \quad \text{para} \quad i = 1, 2, \ldots, n. \ \blacksquare$$

Os pontos em que se verifica a condição do teorema denominam-se **pontos críticos** ou **pontos de estacionaridade** da função e é a eles que se circunscreve o estudo da existência ou não de extremos. Para isso é, porém, necessário recorrer a condições adicionais que permitam esclarecer o que de facto ocorre num dado ponto crítico.

Teorema 7.2. *Se a é um ponto de estacionaridade da função $f(x_1, \ldots, x_n)$ e $m > 1$ é a ordem da primeira derivada direccional de f que não é identicamente nula em a, então:*

(a) *Se m é ímpar não há extremo em a;*

(b) *Se m é par e $f^{(m)}_{vers\,\overrightarrow{h}}(a)$ é positiva (resp. negativa) a função tem um mínimo (resp. máximo) em a;*

(c) *Se m é par mas $f^{(m)}_{vers\,\overrightarrow{h}}(a)$ não tem sinal determinado não há extremo em a;*

(d) *Se m é par mas $f^{(m)}_{vers\,\overrightarrow{h}}(a) \geq 0$ (ou ≤ 0) nada se pode concluir por esta via (caso duvidoso).*

Demonstração. De acordo com a hipótese, a expressão da **Fórmula de Taylor** para o ponto a é

$$f(a+\overrightarrow{h})-f(a)=\frac{||\overrightarrow{h}||^m}{m!}f^{(m)}_{vers\,\overrightarrow{h}}(a+\theta\overrightarrow{h}),\quad \text{com}\quad 0<\theta<1.$$

Mas um raciocínio análogo ao utilizado no teorema anterior permite concluir que

$$f^{(m)}_{vers\,\overrightarrow{h}}(a+\theta\overleftarrow{h})=f^{(m)}_{vers\,\overrightarrow{h}}(a)+\alpha$$

em que α tende para zero quando $||\overrightarrow{h}||\to 0$. A expressão da fórmula de Taylor pode pois escrever-se

$$f(a+\overrightarrow{h})-f(a)=\frac{||\overrightarrow{h}||^m}{m!}\left[f^{(m)}_{vers\,\overrightarrow{h}}(a)+\alpha\right] \qquad (7.2)$$

com $\lim\limits_{||\overrightarrow{h}||\to 0}\alpha=0$, o que mostra que, para $||\overrightarrow{h}||$ suficientemente pequeno, o sinal de $f(a+\overrightarrow{h})-f(a)$ é o de $f^{(m)}_{vers\,\overrightarrow{h}}(a)$. É com base neste facto que se pode agora proceder ao estudo dos extremos.

(a) Se m for ímpar, a derivada direccional não mantém o mesmo sinal para todo o $\overrightarrow{h}$ visto que

$$f^{(m)}_{vers(-\overrightarrow{h})}(a)=(-1)^m f^{(m)}_{vers\,\overrightarrow{h}}(a).$$

Por consequência $f(a+\overrightarrow{h})-f(a)$ não tem sinal constante em nenhuma vizinhança de a e, portanto, não há extremo em a.

(b) Se m é par e $f^{(m)}_{vers\,\overrightarrow{h}}(a)>0$ a expressão (7.2) mostra que $f(a+\overrightarrow{h})>f(a)$ e portanto $f(a)$ é um mínimo relativo . De forma análoga se conclui que sendo m par e $f^{(m)}_{vers\,\overrightarrow{h}}(a)<0$ há um máximo em a.

(c) Se m é par mas a derivada direccional de ordem m tem sinal determinado para todo o $\overrightarrow{h}$, a relação (7.2) mostra que o mesmo sucede com $f(a+\overrightarrow{h})-f(a)$ e, portanto, não pode haver extremo em a.

(d) Se m é par mas $f^{(m)}_{vers\,\overrightarrow{h}}(a) \geq 0$ (ou ≤ 0) não há garantia de que para direcções vizinhas daquelas em que a derivada se anula (chamadas direcções singulares) o sinal da derivada direccional não permite pois esclarecer a situação, pelo que terá de se proceder ao estudo directo do comportamento da função nas vizinhanças de a. ■

Exemplo 7.1. Estudar a existência de extremos locais da função

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

Resolução. Os pontos críticos são as soluções do sistema

$$\begin{cases} \dfrac{\partial f}{\partial x} = 2x - 2 = 0 \\[2ex] \dfrac{\partial f}{\partial y} = 4y = 0 \end{cases}$$

Da resolução do sistema conclui-se que há apenas um ponto crítico, ponto interior $(1,0)$. Designando por ϵ_1 e ϵ_2 as componentes do versor dum vector genérico $\overrightarrow{h}$, isto é, $vers\,\overrightarrow{h} = (\epsilon_1, \epsilon_2)$, as derivadas da função no ponto $(1,0)$ segundo o versor de $\overrightarrow{h}$ são

$$\begin{aligned} f'_{vers\,\overrightarrow{h}}(1,0) &= 0\epsilon_1 + 0\epsilon_2 = 0 \\[2ex] f''_{vers\,\overrightarrow{h}}(1,0) &= 2\epsilon_1^2 + 4\epsilon_2^2 > 0. \end{aligned}$$

A primeira derivada direccional não nula no ponto $(1,0)$ é de ordem par ($m=2$) e tem valor positivo. Portanto a função tem um mínimo no ponto $(1,0)$ de valor $f(1,0)=0$. ◄

Exemplo 7.2. Estudar a existência de extremos locais da função

$$f(x,y) = 2x^2 + y^3 - 2xy^2.$$

Resolução. Da resolução do sistema

$$\begin{cases} \dfrac{\partial f}{\partial x} = 4x - 2y^2 = 0 \\[2ex] \dfrac{\partial f}{\partial y} = 3y^2 - 4xy = 0 \end{cases}$$

conclui-se haver dois pontos críticos, $(0,0)$ e $\left(\dfrac{9}{8}, \dfrac{3}{2}\right)$.

- Em $(0,0)$, é

  $$f'_{vers\,\overrightarrow{h}}(0,0) = 0 \quad \text{e} \quad f''_{vers\,\overrightarrow{h}}(0,0) = 4\epsilon_1^2 \geq 0$$

  e trata-se, portanto do caso duvidoso.

- Em $\left(\dfrac{9}{8}, \dfrac{3}{2}\right)$ é

  $$f''_{vers\,\overrightarrow{h}}\left(\frac{9}{8}, \frac{3}{2}\right) = 4\epsilon_1^2 - 12\epsilon_1\epsilon_2 + \frac{9}{2}\epsilon_2^2 = 4 + \frac{1}{2}\epsilon_2^2 - 12\epsilon_1\epsilon_2$$

  que não tem sinal determinado (basta ver, por exemplo, o que se passa nas direcções (0,1) e $\left(\dfrac{\sqrt{2}}{2}, \dfrac{\sqrt{2}}{2}\right)$. Logo, não há extremo. ◀

Na prática assume particular importância na prática o caso em que a função é apenas de duas variáveis e $m = 2$ num ponto crítico (a,b). Nesse caso o estudo da existência de extremo em (a,b) resume-se ao estudo do sinal de

$$f^{(2)}_{vers\,\overrightarrow{h}}(a,b) = \frac{\partial^2 f}{\partial x^2}(a,b)\epsilon_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a,b)\epsilon_1\epsilon_2 + \frac{\partial^2 f}{\partial y^2}(a,b)\epsilon_2^2$$

em que ϵ_1 e ϵ_2 são as componentes do versor de $\overrightarrow{h}$.

Pondo $\dfrac{\partial^2 f}{\partial x^2}(a,b) = r$, $\dfrac{\partial^2 f}{\partial x \partial y}(a,b) = s$ e $\dfrac{\partial^2 f}{\partial y^2}(a,b) = t$ (o que se denomina notação de Monge) a expressão anterior pode escrever-se mais simplesmente

$$f^{(2)}_{vers\,\overrightarrow{h}}(a,b) = r\epsilon_1^2 + 2s\epsilon_1\epsilon_2 + t\epsilon_2^2.$$

Como é evidente,

$$\begin{aligned} rf^{(2)}_{vers\,\overrightarrow{h}}(a,b) &= r^2\epsilon_1^2 + 2rs\epsilon_1\epsilon_2 + s^2\epsilon_2^2 - s^2\epsilon_2^2 + rt\epsilon_2^2 \\ &= (r\epsilon_1 + s\epsilon_2)^2 + (rt - s^2)\epsilon_2^2. \end{aligned} \tag{7.3}$$

Pondo, na expressão (7.3), $\Delta = rt - s^2$, é agora fácil proceder ao estudo da existência de extremo no ponto (a,b). Assim:

(a) Se $\Delta = rt - s^2 > 0$, o segundo membro de (7.3) é positivo e portanto:

- $r > 0 \Longrightarrow f^{(2)}_{vers\,\overrightarrow{h}}(a,b) > 0 \Longrightarrow$ há mínimo em (a,b);
- $r < 0 \Longrightarrow f^{(2)}_{vers\,\overrightarrow{h}}(a,b) < 0 \Longrightarrow$ há máximo em (a,b).
- O caso $r = 0$ nunca se verifica, pois que então seria $\Delta = -s^2 > 0$, o que é absurdo.

Um raciocínio análogo referido a t permitiria concluir que há mínimo se $t > 0$ e máximo se $t < 0$.

(b) Se $\Delta = rt - s^2 < 0$, com $r \neq 0$, por exemplo $r > 0$, a segunda derivada direccional não tem sinal constante e portanto não pode haver extremo. De facto, para as direcções em que $\epsilon_2 = 0$ tem-se

$$rf^{(2)}_{vers\,\overrightarrow{h}}(a,b) = (r\epsilon_1 + s\epsilon_2)^2 > 0,$$

e portanto a derivada direccional é positiva, enquanto que para as direcções em que $r\epsilon_1 + s\epsilon_2 = 0$ é

$$rf^{(2)}_{vers\,\overrightarrow{h}}(a,b) = (rt - s^2)\epsilon_2^2 < 0,$$

e a derivada é negativa. Conclusão análoga se pode tirar quando $r = 0$ com $t \neq 0$ e quando $r = t = 0$.

Por consequência, quando $\Delta < 0$ **não há extremo.**

(c) Se $\Delta = rt - s^2 = 0$, tem-se

$$rf^{(2)}_{vers\,\overrightarrow{h}}(a,b) = (r\epsilon_1 + s\epsilon_2)^2$$

e a segunda derivada direccional tem o sinal de r, excepto nas direcções singulares tais que $r\epsilon_1 + s\epsilon_2 = 0$, nas quais se anula. Trata-se, pois, do **caso duvidoso.**

A análise anterior permite portanto concluir que uma função de duas variáveis $f(x, y)$ só tem extremo nos pontos críticos em que é

$$rt - s^2 = \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0$$

e que esse extremo é um mínimo se for $\dfrac{\partial^2 f}{\partial x^2} > 0$ (ou $\dfrac{\partial^2 f}{\partial y^2} > 0$) e um máximo se $\dfrac{\partial^2 f}{\partial x^2} < 0$ (ou $\dfrac{\partial^2 f}{\partial y^2} < 0$).

Nos pontos críticos em que

$$\frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 < 0$$

não há extremo.

Se $rt - s^2 = 0$ nada se pode concluir e há que recorrer a outra forma de análise.

Exemplo 7.3. Resolver por este método o problema anterior Exemplo 7.1.

Resolução. Como, $\dfrac{\partial^2 f}{\partial x^2} = 2$, $\dfrac{\partial^2 f}{\partial y^2} = 4$ e $\dfrac{\partial^2 f}{\partial x \partial y} = 0$ é $\Delta = rt - s^2 = 2 \cdot 4 - 0 = 8 > 0$. E, como, $\dfrac{\partial^2 f}{\partial x^2} > 0$, segue-se que a função tem um mínimo no ponto $(1, 0)$. ◀

O estudo anteriormente feito incidiu apenas na determinação de extremos em pontos interiores do domínio da função. Mas se esse domínio for um conjunto fechado , é evidente que pode também acontecer que a função admita extremo em pontos da fronteira.

Para esses pontos a análise pode fazer-se recorrendo, tal como no caso dos pontos interiores, à fórmula de Taylor, mas tendo em atenção que apenas se pode considerar derivadas segundo vectores situados no interior do domínio.

7.2 Extremos Condicionados

É frequente nas aplicações práticas ser necessário determinar os extremos de uma função quando as suas variáveis estão sujeitas a certas condições, chamadas condições suplementares.

Nesse caso o problema deixa de ser um problema comum de determinação de extremos e passa a constituir o que se denomina um problema de **extremos condicionados** (ou sujeitos a condições adicionais).

Dada, por exemplo, a função $w = f(x, y, z)$, suponhamos que as suas variáveis x, y, z estão "ligadas" pelas duas condições

$$\varphi_1(x, y, z) = 0 \quad \text{e} \quad \varphi_2(x, y, z) = 0.$$

Vejamos como determinar condições necessárias para que nessas condições a função f tenha um extremo relativo num ponto (x_0, y_0, z_0) do seu domínio.

Se nesse ponto há um extremo, o diferencial de f deve ser nulo, visto que aí a função não cresce nem decresce. Portanto

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = 0.$$

Por outro lado, uma vez que $\varphi_1 = 0$ e $\varphi_2 = 0$, também

$$\begin{aligned} d\varphi_1 &= \frac{\partial \varphi_1}{\partial x}dx + \frac{\partial \varphi_1}{\partial y}dy + \frac{\partial \varphi_1}{\partial z}dz = 0 \\ d\varphi_2 &= \frac{\partial \varphi_2}{\partial x}dx + \frac{\partial \varphi_2}{\partial y}dy + \frac{\partial \varphi_2}{\partial z}dz = 0. \end{aligned}$$

Multiplicando estas equações por λ_1 e λ_2, respectivamente, se as somarmos membro a membro com a equação $df = 0$ vem:

$$\begin{aligned} d(f + \lambda_1\varphi_1 + \lambda_2\varphi_2) &= \left(\frac{\partial f}{\partial x} + \lambda_1\frac{\partial \varphi_1}{\partial x} + \lambda_2\frac{\partial \varphi_2}{\partial x}\right)dx + \\ &+ \left(\frac{\partial f}{\partial y} + \lambda_1\frac{\partial \varphi_1}{\partial y} + \lambda_2\frac{\partial \varphi_2}{\partial y}\right)dy + \\ &+ \left(\frac{\partial f}{\partial z} + \lambda_1\frac{\partial \varphi_1}{\partial z} + \lambda_2\frac{\partial \varphi_2}{\partial z}\right)dz = 0. \end{aligned}$$

Como os diferenciais dx, dy, dz são arbitrários, esta soma só é nula se o forem separadamente os seus coeficientes, isto é

$$\begin{cases} \dfrac{\partial f}{\partial x} + \lambda_1\dfrac{\partial \varphi_1}{\partial x} + \lambda_2\dfrac{\partial \varphi_2}{\partial x} = 0 \\ \dfrac{\partial f}{\partial y} + \lambda_1\dfrac{\partial \varphi_1}{\partial y} + \lambda_2\dfrac{\partial \varphi_2}{\partial y} = 0 \\ \dfrac{\partial f}{\partial z} + \lambda_1\dfrac{\partial \varphi_1}{\partial z} + \lambda_2\dfrac{\partial \varphi_2}{\partial z} = 0. \end{cases}$$

Associando a estas 3 equações as 2 equações vinculares $\varphi_1 = 0$, $\varphi_2 = 0$, obtém-se um sistema de 5 equações a 5 incógnitas x, y, z, λ_1, λ_2. Ora, resolvendo este sistema, os valores de x, y e z que o satisfazem correspondem às coordenadas dos pontos em que a função pode admitir extremos, visto que foram obtidas a partir da condição $df = 0$.

O método exposto denomina-se **método dos multiplicadores de Lagrange**.

Como se pode observar, o método corresponde à determinação de condições necessárias para que um ponto seja ponto de **extremo não condicionado** da função

$$V = f + \lambda_1\varphi_1 + \lambda_2\varphi_2.$$

No caso geral em que se tem uma função $W = f(x_1, x_2, \ldots, x_n)$ cujas variáveis $x_1, x_2, \ldots, x_n$ estão ligadas por $m < n$ condições

$$\begin{array}{rcl} \varphi_1(x_1, \ldots, x_n) & = & 0 \\ \varphi_2(x_1, \ldots, x_n) & = & 0 \\ \ldots & & \\ \varphi_m(x_1, \ldots, x_n) & = & 0, \end{array} \tag{7.4}$$

pode mostrar-se, raciocinando de forma análoga, que os possíveis extremos da função sujeitos a essas condições existem nos pontos em que há possíveis extremos **não condicionados** da função

$$V = f + \lambda_1\varphi_1 + \lambda_2\varphi_1 + \ldots + \lambda_m\varphi_m$$

onde $\lambda_1, \ldots, \lambda_m$ são os chamados multiplicadores de Lagrange. Tais pontos são as soluções do sistema de $m + n$ equações constituído pelas equações

$$\frac{\partial v}{\partial x_i} = 0, \quad i = 1, 2, \ldots, n;$$

e pelas equações das condições (7.4).

Exemplo 7.4. Determinar os extremos da função $f(x,y) = x+2y$ sujeitos à condição $x^2 + y^2 = 5$.

Resolução. Seja $V(x) = x + 2y + \lambda(x^2 + y^2 - 5)$. Os pontos críticos desta função são as soluções do sistema

$$\begin{cases} \dfrac{\partial V}{\partial x} = 1 + 2\lambda x = 0, \\[2ex] \dfrac{\partial V}{\partial y} = 2 + 2\lambda y = 0, \\[2ex] x^2 + y^2 = 5. \end{cases}$$

Resolvendo o sistema obtêm-se as soluções

$$\begin{cases} x = 1 \\ y = 2 \\ \lambda = -\dfrac{1}{2} \end{cases} \quad \text{e} \quad \begin{cases} x = -1 \\ y = -2 \\ \lambda = \dfrac{1}{2} \end{cases}$$

Como: $\dfrac{\partial^2 v}{\partial x^2} = 2\lambda$, $\dfrac{\partial^2 v}{\partial y^2} = 2\lambda$, $\dfrac{\partial^2 v}{\partial x \partial y} = 0$, tem-se:

(a) No ponto $(1,2)$:

$$\frac{\partial^2 V}{\partial x^2} \cdot \frac{\partial^2 V}{\partial y^2} - \left(\frac{\partial^2 V}{\partial x \partial y}\right)^2 = 4\lambda^2 = 4\left(-\frac{1}{2}\right)^2 = 1 > 0$$

e

$$\frac{\partial^2 v}{\partial x^2} = 2\lambda = 2\left(-\frac{1}{2}\right) = -1 < 0,$$

e portanto há um máximo em $(1,2)$;

(b) No ponto $(-1,-2)$:

$$\frac{\partial^2 V}{\partial x^2}\cdot\frac{\partial^2 v}{\partial y^2}-\left(\frac{\partial^2 V}{\partial x\partial y}\right)^2=4\lambda^2=4\cdot\left(\frac{1}{2}\right)^2=1>0$$

e

$$\frac{\partial^2 V}{\partial x^2}=2\lambda=2\cdot\left(\frac{1}{2}\right)=1>0,$$

e portanto há um mínimo em $(-1,-2)$. ◀

7.3 Exercícios

1. Determine os extremos locais das funções:

   (a) $z = x^3 + 8y^3 - 6xy + 1$;

   (b) $z = x^{-1} + y^{-1} - xy$;

   (c) $z = 3x^3y - x^2y^2 + x$;

   (d) $z = e^{\frac{x}{2}}(x + y^2)$;

   (e) $z = (2ax - x^2)(2bx - y^2)$.

2. Determine os extremos das funções cujas variáveis estão sujeitos às condições indicadas:

   (a) $z = x^2 + y^2$, $x + y = 1$;

   (b) $z = e^{xy}$, $x + y = a$;

   (c) $z = xy$, $x^2 + y^2 \leq 1$;

   (d) $u = xyz$, $x + y - z = 3$, $x - y - z = 8$;

   (e) $u = x + yz$, $x^2 + y^2 + z^2 \leq 1$

3. Determine as dimensões do cilindro de superfície total $S = 6\pi$ cujo volume é máximo.

4. Determine o paralelipípedo rectangular de volume V de superfície mínima.

5. Determine a distância mínima do ponto $(-1, 5)$ à parábola $y^2 = x$.

6. Determine o triângulo de área máxima que se pode inscrever numa circunferência de raio dado.

7. De entre os triângulos com o mesmo perímetro, determine o que tem área máxima.

8. Determine a distância mínima da origem ao plano

$$x + y + z = 2.$$

Capítulo 8

Funções implícitas

8.1 Funções definidas por uma equação

É frequente, no *Cálculo* e na *Geometria Analítica*, a equação de uma curva plana ser dada, não na forma $y = f(x)$, mas sim na forma $F(x, y) = 0$. É, por exemplo, o caso da circunferência com centro na origem e raio r cuja equação é geralmente assim escrita:

$$x^2 + y^2 - r^2 = 0.$$

Neste caso, para cada valor de x no intervalo $-r < x < r$ obtêm-se dois valores de y dados por

$$y = \pm\sqrt{r^2 - x^2}.$$

Quer dizer, há duas funções, $y = \pm\sqrt{r^2 - x^2}$, que satisfazem a equação $x^2 + y^2 - r^2 = 0$. Diz-se então que esta equação define implicitamente y como função de x. Porém, nem sempre uma equação de forma $F(x, y) = 0$ define y como uma função de x. É o que acontece, por exemplo, com a equação $F(x, y) = x^2 + y^2 + 1 = 0$, pois não há nenhum par (x, y) em $\mathbb{R}^2$ que satisfaça a equação.

De forma análoga, também a equação de uma superfície é frequentemente apresentada na forma $F(x, y, z) = 0$ e não na forma $z = f(x, y)$. Assim, por exemplo, se na equação

$$2x - y + z - 1 = 0$$

se substituir a variável z pela função

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

a equação é satisfeita e diz-se, por isso, que a equação dada define implicitamente z como função de x e y. Já, porém, a equação

$$x^2 + y^2 + z^2 + 1 = 0$$

não define z como função de x e y, visto que não há nenhum termo (x, y, z) de números reais que a satisfaça.

Considerações semelhantes podem ser feitas para mais variáveis. Assim, diz-se que uma equação $F(x_1, x_2, \ldots, x_n; u) = 0$ define implicitamente u como função das demais variáveis se existe uma função $u = f(x_1, x_2, \ldots, x_n)$ tal que

$$F(x_1, x_2, \ldots, x_n;\ f(x_1, x_2, \ldots, x_n)) = 0.$$

Nos exemplos anteriores foi fácil verificar se as equações definiam ou não implicitamente uma das variáveis como função da outra ou outras, para o que bastou resolver a equação em ordem à variável considerada dependente, isto é, explicitar a chamada função implícita.

Tal explicitação não é, todavia, sempre possível, pelo que a teoria das funções implícitas tem por objectivo estudar as suas propriedades sem necessidade de as explicitar.

A primeira questão que se levanta é a das chamadas **condições**

de existência, isto é, a de saber em que condições é que uma dada equação define (implicitamente) uma das suas variáveis como função das restantes. O teorema que a seguir se enuncia responde a essa questão.

Teorema 8.1 (Teorema de existência). *Seja dada a equação*

$$F(x_1, x_2, \ldots, x_n; u) = 0.$$

Se F é uma função definida num conjunto aberto D de $\mathbb{R}^{n+1}$ tal que:

(a) Existe um ponto $P = (a_1, a_2, \ldots, a_n; u_0)$ em D no qual a equação é satisfeita, isto é, em que $F(a_1, a_2, \ldots, a_n; u_0) = 0$;

(b) É de classe C^1 numa vizinhança do ponto P e

(c) $\dfrac{\partial F}{\partial u}(x_1, x_2, \ldots, x_n; u) \neq 0$ nessa vizinhança;

então existe em $\mathbb{R}^n$ uma vizinhança de $(a_1, a_2, \ldots, a_n)$ na qual está definida uma e uma só função $u = f(x_1, x_2, \ldots, x_n)$ que verifica nessa vizinhança a equação dada.

Demonstração. Com efeito, se

$$\frac{\partial F}{\partial u}(a_1, a_2, \ldots, a_n; u_0) \neq 0,$$

a função $F = (a_1, a_2, \ldots, a_n; u)$, de uma única variável u, é estritamente crescente ou decrescente no ponto $u = u_0$ e como, por hipótese, é contínua nesse ponto, existe um número positivo δ tal que:

(a) $F = (a_1, \ldots, a_n; u_0 - \delta)$ e

(b) $F = (a_1, a_2, \ldots, a_n; u_0 + \delta)$;

têm sinais contrários. Pela mesma razão, existe uma vizinhança do ponto $a = (a_1, \ldots, a_n)$ na qual (a) e (b) conservam os seus sinais. Sendo assim, a função $F = (x_1, \ldots, x_n; u)$ admite entre $u_0 - \delta$ e $u_0 + \delta$ pelo menos uma raiz real da forma $u = f(x_1, \ldots, x_n)$ que na vizinhança de a verifica a condição

$$F = (x_1, \ldots, x_n; f(x_1, \ldots, x_n)) = 0.$$

Esta raiz (função) é única pois que, se houvesse duas funções $u_1 \neq u_2$ para as quais a função F se anulasse na vizinhança considerada do ponto $a = (a_1, \ldots, a_n)$, então, de acordo com o teorema de Rolle, existiria entre u_1 e u_2 um ponto c na qual

$$\frac{\partial F}{\partial u}(x_1, x_2, \ldots, x_n; c) = 0$$

o que contradiz a hipótese (c) do teorema. ■

O teorema tem, como se vê, carácter local pois assegura a existência da função implícita apenas numa vizinhança do ponto $(a_1, a_2, \ldots, a_n)$.

Observação. Pode também demonstrar-se que, nas condições do teorema, a função implícita $u = f(x_1, x_2, \ldots, x_n)$ é de classe C^1 (e portanto diferenciável) no ponto $(a_1, a_2, \ldots, a_n)$.

Exemplo 8.1. Verificar se a equação

$$x^2 + y^2 + z^2 - 1 = 0$$

define z como função implícita de x e y numa vizinhança do ponto $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)$.

Resolução. A função do primeiro membro da equação é de classe C^1 e o ponto dado satisfaz a equação. Todavia nesse ponto é $\dfrac{\partial F}{\partial z} = 0$. Logo nada garante que a equação defina z como função implícita de x e y em alguma vizinhança do ponto considerado. ◀

Averiguada, através do teorema anterior, a existência de função implícita, é possível determinar também as suas derivadas sem necessidade de explicitar a função, tal como poderá ver-se em seguida.

(a) *Função* $y = f(x)$ *definida implicitamente pela equação* $F(x, y) = 0$.

Compondo $w = F(x, y)$ com $y = f(x)$ obtém-se a função de uma só variável $w = F(x, f(x))$. Se tanto $F(x, y)$ como $f(x)$ são diferenciáveis numa vizinhança do ponto em que se pretende calcular a derivada , tem-se, aplicando à função composta a regra de cadeia

$$\frac{dw}{dx} = \frac{\partial F}{\partial x} \cdot \frac{dx}{dx} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx}.$$

Mas, como é $w = F(x, y) = 0$, então $\dfrac{dw}{dx} = 0$ e por consequência

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx} = 0$$

de onde resulta a expressão pretendida da derivada da função implícita $y = f(x)$

$$\frac{dy}{dx} = -\frac{\dfrac{\partial F}{dx}}{\dfrac{\partial F}{dy}} \quad (\text{se } \frac{\partial F}{dy} \neq 0).$$

(b) *Função* $z = f(x, y)$ *definida implicitamente pela equação* $F(x, y, z) = 0$.

Tal como no caso anterior, se $F(x, y, z)$ e $f(x, y)$ são diferenciáveis, aplicando a regra da cadeia à composta de $w = F(x, y, z)$ com $z = f(x, y)$ e $x = x$, $y = y$ obtém-se:

$$\frac{\partial w}{\partial x} = \frac{\partial F}{\partial x} \cdot \frac{dx}{dx} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial x} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{\partial z}{\partial x}.$$

Como

$$w = F(x, y, z) = 0, \quad \text{é} \quad \frac{\partial w}{\partial x} = 0$$

e portanto,

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial x} = 0$$

de onde resulta:

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} \quad \left(\text{se } \frac{\partial F}{\partial z} \neq 0\right).$$

De forma análoga se obtém a expressão da derivada em ordem a y

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} \quad \left(\text{se } \frac{\partial F}{\partial z} \neq 0\right).$$

(c) *Caso geral: função* $u = f(x_1, x_2, \ldots, x_n)$ *definida implicitamente pela equação* $F(x_1, x_2, x_n; u) = 0$.

De forma análoga à dos casos anteriores, se as funções envolvidas são diferenciáveis, aplicando a regra de cadeia à composta de $w = F(x_1, \ldots, x_n; u)$ com $u = f(x_1, x_2, \ldots, x_n)$ e tendo em conta que $w = 0$, a derivada em ordem a x_i é

$$\begin{aligned}\frac{\partial w}{\partial x_i} &= \frac{\partial F}{\partial x_1} \cdot \frac{dx_1}{dx_i} + \ldots + \frac{\partial F}{\partial x_i} \cdot \frac{dx_i}{dx_i} + \ldots + \frac{\partial F}{\partial u} \cdot \frac{\partial u}{\partial x_i} = \\ &= \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} \cdot \frac{\partial u}{\partial x_i} = 0\end{aligned}$$

de onde resulta

$$\frac{\partial u}{\partial x_i} = -\frac{\dfrac{\partial F}{\partial x_i}}{\dfrac{\partial F}{\partial u}} \quad \left(\text{se } \frac{\partial F}{\partial u} \neq 0\right)$$

para a expressão da derivada da função implícita em relação à variável x_i $(i = 1, 2, \ldots, n)$.

Exemplo 8.2. Determinar $\dfrac{dy}{dx}$, sendo y definida implicitamente por $F(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$.

Resolução: $\dfrac{dy}{dx} = -\dfrac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = -\dfrac{2Ax + By + D}{Bx + 2Cy + E}$. ◀

Exemplo 8.3. Determinar $\dfrac{\partial z}{\partial x}$ e $\dfrac{\partial z}{\partial y}$, sendo z definida implicitamente por $x^2y + y^2x + xz = 0$.

Resolução: $\dfrac{\partial z}{\partial x} = -\dfrac{2xy + z}{y^2 + x}$; $\quad \dfrac{\partial z}{\partial y} = -\dfrac{x^2 + 2yz}{y^2 + x}$. ◀

8.2 Funções definidas implicitamente por um sistema de equações

Em vez de serem definidas por uma única equação as funções implícitas podem também sê-lo por um sistema de equações simultâneas. Assim, por exemplo, se o sistema

$$\begin{cases} F_1(x,y,u,v) = 0 \\ \\ F_2(x,y,u,v) = 0 \end{cases}$$

puder ser resolvido em ordem a u e v, isso significa que o sistema define implicitamente duas funções de x e y tais que: $u = u(x,y)$ e $v = v(x,y)$.

As condições em que um sistema de equações pode definir implicitamente funções são estabelecidas no teorema seguinte que se enuncia sem demonstração.

Teorema 8.2 (**Teorema de existência**). *Considere-se o sistema de n equações a $m+n$ variáveis*

$$\begin{cases} f_1(x_1,\ldots,x_m;\ y_1,\ldots,y_n) = 0 \\ f_2(x_1,\ldots,x_m;\ y_1,\ldots,y_n) = 0 \\ \qquad\cdots \\ f_n(x_1,\ldots,x_m;\ y_1,\ldots,y_n) = 0 \end{cases} \tag{8.1}$$

em que $f_1, f_2, \ldots, f_n$ são funções definidas num aberto de $\mathbb{R}^{m+n}$. Se se verificarem as condições:

(a) Existe um ponto $p = (a_1,\ldots,a_m;\ b_1,\ldots,b_n)$ no qual as equações do sistema são satisfeitas;

(b) As funções $f_1, f_2, \ldots, f_n$ *são de classe* C^1 *numa vizinhança do ponto p;*

(c) O Jacobiano $\dfrac{\partial(f_1, f_2, \ldots, f_n)}{\partial(y_1, y_2, \ldots, y_n)}$ *é diferente de zero no ponto p.*

Então, existe em $\mathbb{R}^m$ *uma vizinhança do ponto* $a = (a_1, a_2, \ldots, a_m)$ *na qual ficam definidas implicitamente n funções*

$$\begin{aligned} y_1 &= g_1(x_1, \ldots, x_m) \\ y_2 &= g_2(x_1, \ldots, x_m) \\ y_n &= g_n(x_1, \ldots, x_m) \end{aligned} \tag{8.2}$$

que se tomam em $(a_1, \ldots, a_m)$ *os valores*

$$b_i = g_i(a_1, \ldots, a_m), \quad i = 1, \ldots, n;$$

satisfazem as equações do sistema dado e são de classe C^1 *na referida vizinhança de a.*

Estabelecida a existência das funções implícitas y_i $(i = 1, \ldots, n)$ (através do teorema anterior), é possível determinar as suas derivadas parciais $\dfrac{\partial y_i}{\partial x_k}$ no ponto $(a_1, \ldots, a_n)$ sem necessidade de as funções serem explicitadas. Com efeito, derivando em ordem a x_k as equações do sistema (8.1) após composição das funções $f_1, f_2, \ldots, f_n$ com as funções em (8.2), obtém-se o sistema:

$$\begin{cases} \dfrac{\partial f_1}{\partial x_k} + \dfrac{\partial f_1}{\partial y_1} \cdot \dfrac{\partial y_1}{\partial x_k} + \ldots + \dfrac{\partial f_1}{\partial y_n} \cdot \dfrac{\partial y_n}{\partial x_k} = 0 \\ \\ \ldots \\ \\ \dfrac{\partial f_n}{\partial x_k} + \dfrac{\partial f_n}{\partial y_1} \cdot \dfrac{\partial y_1}{\partial x_k} + \ldots + \dfrac{\partial f_n}{\partial y_n} \cdot \dfrac{\partial y_n}{\partial x_k} = 0 \end{cases} \tag{8.3}$$

que resolvido pela regra de Cramer conduz às expressões pretendidas das derivadas das funções $y_1, y_2, \ldots, y_n$ em ordem a x_k

$$\frac{\partial y_i}{\partial x_k} = -\frac{\dfrac{\partial(f_1, f_2, \ldots, f_n)}{\partial(y_1, \ldots, y_{k-1}, x_k, y_{k+1}, \ldots, y_n)}}{\dfrac{\partial(f_1, f_2, \ldots, f_n)}{\partial(y_1, y_2, \ldots, y_n)}} \tag{8.4}$$

com $i = 1, 2, \ldots, n$.

Por exemplo, a derivada em ordem a x_1 da função y_2 é

$$\frac{\partial y_2}{\partial x_1} = -\frac{\dfrac{\partial(f_1, f_2, \ldots, f_n)}{\partial(y_1, x_2, y_3, \ldots, y_n)}}{\dfrac{\partial(f_1, f_2, \ldots, f_n)}{\partial(y_1, y_2, \ldots, y_n)}}.$$

Das anteriores expressões logo se conclui que o sistema (8.3) só é possível se

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(y_1, y_2, \ldots, y_n)} \neq 0.$$

Exemplo 8.4. Determine $\dfrac{\partial u}{\partial x}$ e $\dfrac{\partial v}{\partial x}$, sendo u e v funções de x e y definidas implicitamente pelo sistema de equações

$$\begin{cases} u^2 - v^2 + 2x = 0, \\ \\ uv - y = 0. \end{cases}$$

Resolução.

Método 1. Derivando as equações do sistema em ordem a x obtém-se

$$2u \cdot \frac{\partial u}{\partial x} - 2v \cdot \frac{\partial v}{\partial x} + 2 = 0 \quad \text{e} \quad v \cdot \frac{\partial u}{\partial x} + u \cdot \frac{\partial v}{\partial x} = 0.$$

Resolvendo o sistema em ordem a $\dfrac{\partial u}{\partial x}$ e $\dfrac{\partial v}{\partial x}$ tem-se,

$$\frac{\partial u}{\partial x} = -\frac{u}{u^2+v^2}, \quad \text{e} \quad \frac{\partial v}{\partial x} = \frac{v}{u^2+v^2}.$$

Método 2. Pondo $f_1 = u^2 - v^2 + 2x$ e $f_2 = uv - y$ as expressões das derivadas podem obter-se directamente da fórmula (8.4):

$$\frac{\partial u}{\partial x} = -\frac{\dfrac{\partial(f_1,f_2)}{\partial(x,v)}}{\dfrac{\partial(f_1,f_2)}{\partial(u,v)}} = -\frac{u}{u^2+v^2}; \quad \frac{\partial v}{\partial x} = -\frac{\dfrac{\partial(f_1,f_2)}{\partial(u,x)}}{\dfrac{\partial(f_1,f_2)}{\partial(u,v)}} = \frac{v}{u^2+v^2}$$

desde que $u^2 + v^2 \neq 0$. ◀

Exemplo 8.5. Verificar se o sistema de equações

$$\begin{cases} f_1(x,y,z) = 9xy + z^2 = 0 \\ \\ f_2(x,y,z) = xz - 3y^2 = 0 \end{cases}$$

define implicitamente y e z como funções de x numa vizinhança do ponto $(1,-1,3)$.

Resolução. O sistema é verificado no ponto $(1,-1,3)$ e as funções f_1 e f_2 de classe C^1 em qualquer vizinhança desse ponto. Como, além disso, nesse ponto é

$$\frac{\partial(f_1,f_2)}{\partial(y,z)} = -27 \neq 0,$$

o sistema define de facto y e z como funções de x na vizinhança do ponto considerado. ◀

8.3 Extremos de funções definidas implicitamente

A teoria exposta no Capítulo 7 para a determinação dos extremos de funções explícitas estende-se sem dificuldade às funções definidas implicitamente.

Comecemos por considerar o caso de uma função $y = f(x)$ definida implicitamente pela equação $F(x, y) = 0$ num intervalo em que

$$\frac{\partial F}{\partial y} \neq 0.$$

Para que a função implícita tenha um extremo num ponto interior desse intervalo deverá ser

$$\frac{df}{dx} = y' = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = 0.$$

Como $\dfrac{\partial F}{\partial y} \neq 0$ no intervalo, então num ponto em que haja extremo deverá ser necessariamente $\dfrac{\partial F}{\partial x} = 0$.

Portanto, os pontos críticos da função implícita são os pontos que satisfazem o sistema de equações

$$\begin{cases} F(x, y) = 0 \\[2ex] \dfrac{\partial F}{\partial x}(x, y) = 0. \end{cases}$$

Determinados os pontos críticos, para se averiguar se há ou não extremo e qual a sua natureza, calcula-se a segunda derivada da função da função implícita nesses pontos. Ora, atendendo a que $\dfrac{\partial F}{\partial x} = 0$, a expressão da segunda derivada é

$$y'' = -\frac{\dfrac{\partial^2 F}{\partial x^2}}{\dfrac{\partial F}{\partial y}}.$$

Se $y'' > 0$ há um mínimo no ponto crítico considerado e se $y'' < 0$ há um máximo, tal como no caso das funções explícitas. Se for $y'' = 0$ há que recorrer a derivadas de ordem superior.

Exemplo 8.6. Determinar os extremos da função $y = f(x)$ definida implicitamente pela equação $F(x, y) = x^3 + y^3 - 3xy = 0$.

Resolução. Os possíveis pontos de extremo são as soluções do sistema

$$\begin{cases} x^3 + y^3 - 3xy = 0 \\ \\ 3x^2 - 3y = 0 \end{cases}$$

isto é, os pontos $(0, 0)$ e $(\sqrt[3]{2}, \sqrt[3]{4})$. O primeiro não serve pois que nele é $\dfrac{\partial F}{\partial y} = 0$. No segundo ponto o valor da segunda derivada é $y'' = -2$ e, portanto, a função tem um máximo nesse ponto.◀

No caso mais geral em que uma equação

$$F(x_1, x_2, \ldots, x_n; u) = 0$$

define implicitamente u como função de $x_1, x_2, \ldots, x_n$, de forma análoga à do caso anterior se conclui que (supondo $\dfrac{\partial F}{\partial u} \neq 0$) os pontos críticos são as soluções do sistema

$$\begin{cases} F = 0 \\ \\ \dfrac{\partial F}{\partial x_1} = 0 \\ \quad \cdots \\ \dfrac{\partial F}{\partial x_n} = 0. \end{cases}$$

Uma vez determinados os pontos críticos, o estudo da existência de extremos nesses pontos faz-se em seguida da mesma forma que no caso das funções definidas não implicitamente.

Exemplo 8.7. Verificar se tem extremo no ponto $(0, 0, 1)$, sendo z função de x e y definida implicitamente pela equação

$$F = x^6 + y^6 + z^6 + x^2 + y^2 + z^2 - 2 = 0.$$

Resolução. A equação dada define z como função de x e y numa vizinhança do ponto $(0, 0, 1)$ visto que nesse ponto é

$$\frac{\partial F}{\partial z}(0, 0, 1) = 8 \quad \text{e} \quad F(0, 0, 1) = 0.$$

Os pontos críticos da função são as soluções do sistema

$$\begin{cases} F = 0 \\ \\ \dfrac{\partial F}{\partial x} = 0 \\ \\ \dfrac{\partial F}{\partial y} = 0. \end{cases}$$

Como facilmente se pode verificar o ponto $(0,0,1)$ é uma das soluções do sistema. Calculando as segundas derivadas de z nesse ponto tem-se

$$r = \frac{\partial^2 z}{\partial x^2} = -\frac{\dfrac{\partial^2 F}{\partial x^2}}{\dfrac{\partial F}{\partial z}} = -\frac{2}{8} = -\frac{1}{4}; \quad s = \frac{\partial^2 z}{\partial x \partial y} = -\frac{\dfrac{\partial^2 F}{\partial x \partial y}}{\dfrac{\partial F}{\partial z}} = 0$$

e $\quad t = \dfrac{\partial^2 z}{\partial y^2} = -\dfrac{1}{4}$.

Portanto no ponto considerado é

$$rt - s^2 = \frac{1}{16} > 0,$$

o que significa que há um extremo. Como $r = t = -\dfrac{1}{4} < 0$, esse extremo é um máximo. ◄

8.4 Inversão das transformações pontuais

Seja $f : D_1 \subset \mathbb{R}^n \to D_2 \subset \mathbb{R}^n$ a transformação pontual definida pelas funções

$$\begin{array}{rcl} y_1 & = & f_1(x_1, x_2, \dots, x_n) \\ y_2 & = & f_2(x_1, x_2, \dots, x_n) \\ & \cdots & \\ y_n & = & f_n(x_1, x_2, \dots, x_n) \end{array} \tag{8.5}$$

definidas e de classe C^1 em D_1. Fazendo, para $i = 1, 2, \dots, n$

$$f_i(x_1, x_2, \dots, x_n) - y_i = g_i(x_1, x_2, \dots, x_n, y_i) \tag{8.6}$$

as equações que definem a transformação tomam a forma

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n, y_1) &= 0 \\ g_2(x_1, x_2, \ldots, x_n, y_2) &= 0 \\ \ldots \\ g_n(x_1, x_2, \ldots, x_n, y_n) &= 0 \end{aligned} \tag{8.7}$$

sendo os primeiros membros também de classe C^1. Então, de (8.6) resulta

$$\frac{\partial f_i}{\partial x_k} = \frac{\partial g_i}{\partial x_k} \quad (i, k = 1, 2, \ldots, n)$$

e portanto será,

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)} = \frac{\partial(g_1, g_2, \ldots, g_n)}{\partial(x_1, x_2, \ldots, x_n)}.$$

Se este Jacobiano for diferente de zero num dado ponto $a \in D_1$, então, o Teorema 8.2 das funções implícitas garante que a transformação considerada define $x_1, x_2, \ldots, x_n$ como funções de $y_1, y_2, \ldots, y_n$ numa vizinhança do transformado $f(a)$ de a.

Quer dizer: se o jacobiano fôr diferente de zero a transformação é bijectiva, pois faz corresponder a cada ponto $(y_1, y_2, \ldots, y_n)$ de D_2 o ponto $(x_1, x_2, \ldots, x_n)$ de D_1 de que aquele é o transformado.

Nestas condições diz-se que a transformação é invertível ou que tem uma **inversa**.

Se se designar a transformação inicialmente considerada por $Y = f(X)$, a sua inversa, se existir, designa-se por

$$X = f^{-1}(Y).$$

O Teorema 8.2 permite pois afirmar que se as funções (8.5) que definem a transformação são de classe C^1 numa vizinhança do ponto $a \in D_1$ e o seu jacobiano não é nulo, então existe, univocamente determinada, uma transformação inversa numa vizinhança do ponto $(y_1, y_2, \ldots, y_n)$ correspondente a $(x_1, x_2, \ldots, x_n)$.

8.5 Dependência funcional

Definição 8.1.

Dadas n funções de n variáveis

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n) \\ y_2 &= f_2(x_1, x_2, \ldots, x_n) \\ &\ldots \\ y_n &= f_n(x_1, x_2, \ldots, x_n), \end{aligned} \tag{8.8}$$

diz-se que elas são ***funcionalmente dependentes*** *(ou interdependentes) num dado conjunto D de $\mathbb{R}^n$ quando entre elas se pode eliminar as variáveis $x_1, \ldots, x_n$ de modo a obter-se pelo menos uma relação da forma $g(y_1, y_2 \ldots, y_n) = 0$ para todos os valores*

$$(x_1, x_2, \ldots, x_n) \in D.$$

Se não existir tal relação, as funções dizem-se ***funcionalmente independentes****.*

Exemplo 8.8. Sejam, por exemplo, as funções

$$y_1 = \frac{x_2}{x_1} \quad \text{e} \quad y_2 = \sqrt{1 - \frac{x_2^2}{x_1^2}}.$$

Substituindo a primeira na segunda vem $y_2 = \sqrt{1 - y_1^2}$ de onde resulta (elevando ao quadrado e transpondo os termos para o primeiro membro)

$$y_1^2 + y_2^2 - 1 = 0,$$

o que indica que as funções dadas são funcionalmente dependentes. ◀

Exemplo 8.9. De forma análoga, as funções $y_1 = \operatorname{sen} x$ e $y_2 = \cos x$ são funcionalmente dependentes para qualquer x, já que verificam sempre a relação: $y_1^2 + y_2^2 - 1 = 0$. ◀

O teorema seguinte, que se enuncia sem demonstração, estabelece uma condição necessária e suficiente para a dependência funcional.

Teorema 8.3. *Se as funções (8.8) são de classe C^1 num aberto D de $\mathbb{R}^n$, uma condição necessária e suficiente para que sejam funcionalmente dependentes em D é que o seu jacobiano seja identicamente nulo em D, isto é,*

$$\frac{\partial(y_1, y_2, \dots, y_n)}{\partial(x_1, x_2, \dots, x_n)} = 0.$$

Exemplo 8.10. Verificar se são funcionalmente dependentes as funções $u = xy - xz$, $v = yz - xy$ e $w = xz - yz$.

Resolução. Como é

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} y - z & x & -x \\ -y & z - x & y \\ z & -z & x - y \end{vmatrix} = 0,$$

as funções dadas são funcionalmente dependentes. ◀

8.6 Exercícios

1. Se $x^3 + y^3 = 3xy$, calcule $\dfrac{d^2y}{dx^2}$.

2. Calcule $\dfrac{\partial v}{\partial y}$, sendo v a função de x e y definida implicitamente pelo sistema
$$\begin{cases} xu^2 + v = y^3 \\ \\ 2yu - xv^3 = 4x \end{cases}$$

3. Se $u^2 - v = 3x + y$ e $u - 2v^2 = x - 2y$, calcule $\dfrac{\partial u}{\partial x}$ e $\dfrac{\partial v}{\partial y}$.

4. Verifique se as funções $f = xy + yz + xz$, $g = x^2 + y^2 + z^2$ e $h = x + y + z$ são funcionalmente dependentes e, no caso afirmativo, determine a relação entre elas.

5. Dada a transformação $x = u - 2v$, $y = 2u + v$ determine as equações da transformação inversa.

6. Verifique que as funções seguintes são funcionalmente dependentes:

   (a) $u = x^2+y^2+z^2$, $v = x^2-y^2+2z$ e $w = 4y^2+2z^2-4z+8$;

   (b) $u = \ln x + \ln y$ e $v = \cos xy$.

7. Determine os extremos da função de x definida implicitamente pela equação:

   (a) $x^2 - y^2 + x^2y - x = 0$;

   (b) $x^4 + y^4 - 4xy + 2 = 0$.

8. Determine os extremos da função de x e y definida implicitamente pela equação $x^2 + y^2 + z^2 = 1$, supondo $z > 0$.

Capítulo 9

Curvas e superfícies

9.1 Curvas contínuas

Intuitivamente uma curva pode ser identificada com a trajectória de um ponto que se move no espaço segundo uma lei com um grau de liberdade. Esta noção intuitiva foi precisada por Jordan ao definir curva (contínua) como sendo todo o lugar geométrico de pontos do espaço cujas coordenadas, num dado sistema de eixos cartesianos, são funções contínuas de um parâmetro real t definidas num dado intervalo de $\mathbb{R}$.

Abreviadamente pode, portanto, dizer-se que curva, segundo Jordan, é toda a imagem contínua de um segmento da recta real.

Assim, em $\mathbb{R}^3$ chama-se curva contínua ao conjunto C de pontos (x, y, z) determinados pelas funções contínuas

$$x = x(t), \quad y = y(t), \quad z = z(t), \tag{9.1}$$

com t, parâmetro real, a variar num intervalo $[a, b] \subset \mathbb{R}$. Generalizando, curva contínua de $\mathbb{R}^n$ é o conjunto C dos pontos $(x_1, \ldots, x_n)$

definidos por n funções contínuas

$$x_i = x_i(t), \quad i = 1, 2, \ldots, n, \quad a \leq t \leq b.$$

As equações (9.1) chamam-se **equações paramétricas da curva** e definem a chamada **representação paramétrica** da curva.

Se $t = f(s)$ for uma aplicação contínua reversível do intervalo $[c, d]$ no intervalo $[a, b]$ então as relações

$$\begin{aligned} x &= x[f(s)] = \varphi_1(s) \\ y &= y[f(s)] = \varphi_2(s), \quad c \leq s \leq d; \\ z &= z[f(s)] = \varphi_3(s) \end{aligned}$$

representam a mesma curva.

De uma forma geral, se se considerar que duas funções contínuas

$$F_1(t)\ definida\ \ em \quad [a, b], \quad \text{e} \quad F_2(s)\ definida\ \ em \quad [c, d]$$

são equivalentes sempre que existir uma aplicação φ bijectiva e crescente de $[a, b]$ sobre $[c, d]$ tal que $F_1(t) = F_2[(\varphi(t))]$, fica assim definida uma relação de equivalência que permite dividir as funções contínuas em classes de funções equivalentes entre si.

É a cada uma destas classes de equivalência que se chama curva contínua. A mesma curva pode, portanto, ter uma infinidade de representações paramétricas. Na exposição que se segue considerar-se-á, porém, apenas uma dessas representações.

Considerando o vector $\overrightarrow{r}$ que une a origem das coordenadas ao ponto genérico (x, y, z) (vd. Figura 9.1) da curva definida parametricamente pela equações (9.1) tem-se $\overrightarrow{r} = x(t)\overrightarrow{i} + y(t)\overrightarrow{j} + z(t)\overrightarrow{k}$.

Quando t percorre $[a, b]$ as coordenadas x, y, z do ponto P, extremidade do vector $\overrightarrow{r}$, variam e o ponto P descreve a curva.

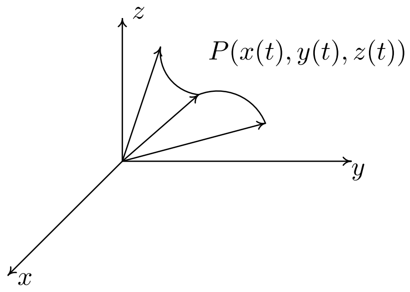

Figura 9.1:

A representação paramétrica (9.1) pode, portanto, ser substituída pela **representação vectorial**

$$\overrightarrow{r} = x(t)\,\overrightarrow{i} + y(t)\,\overrightarrow{j} + z(t)\,\overrightarrow{k}, \quad a \leq t \leq b$$

ou

$$\overrightarrow{r} = \overrightarrow{r}(t), \quad a \leq t \leq b.$$

O vector $\overrightarrow{r}(t)$ chama-se o **raio vector** dos pontos da curva e $\overrightarrow{r} = \overrightarrow{r}(t)$, $a \leq t \leq b$, a equação vectorial da curva. Uma curva pode, portanto, definir-se como sendo uma função vectorial contínua de uma variável real.

Por uma questão de maior simplicidade na linguagem, em vez de curva de equação vectorial $\overrightarrow{r} = \overrightarrow{r}(t)$, $a \leq t \leq b$, diz-se geralmente apenas curva $\overrightarrow{r} = \overrightarrow{r}(t)$, $a \leq t \leq b$.

Numa curva $\overrightarrow{r} = \overrightarrow{r}(t)$, $a \leq t \leq b$, o ponto $\overrightarrow{r}(a)$ denomina-se a origem ou ponto inicial da curva e $\overrightarrow{r}(b)$ a extremidade ou ponto final da curva. Se a origem coincide com a extremidade, isto é, se $\overrightarrow{r}(a) = \overrightarrow{r}(b)$ a curva diz-se **curva fechada**; se $\overrightarrow{r}(a) \neq \overrightarrow{r}(b)$, a curva diz-se **curva aberta** ou **arco da curva**.

Se numa curva $\overrightarrow{r} = \overrightarrow{r}(t)$, $a \leq t \leq b$, houver, entre a e b, valores do parâmetro $t_1 \neq t_2$ para os quais $\overrightarrow{r}(t_1) = \overrightarrow{r}(t_2)$ então diz-se que a curva se intersecta a si mesma ou que tem um ponto duplo.

Uma curva contínua que não se intersecta a si mesma denomina-se **curva simples** ou de **Jordan**; se a curva é aberta diz-se um **arco simples** ou **arco de Jordan**, enquanto que, se for fechada, se diz **curva fechada simples** ou **curva de Jordan**.

Uma curva

$$\overrightarrow{r} = \overrightarrow{r}(t) = x(t)\overrightarrow{i} + y(t)\overrightarrow{j} + z(t)\overrightarrow{k}, \quad a \leq t \leq b,$$

diz-se **curva lisa** (ou **regular**) se existem e são contínuas as derivadas $x'(t)$, $y'(t)$, $z'(t)$ em $[a, b]$ e não se anulam simultaneamente nesse intervalo. Se estas condições não se verificam apenas num número finito de pontos nos quais, todavia, existem as derivadas laterais, então a curva diz-se **seccionalmente lisa** (ou regular por secções). Os pontos em que as condições não se verificam denominam-se **pontos singulares** da curva, em contraposição aos demais, que se denominam **pontos ordinários**.

Se $C_1, C_2, \ldots, C_n$ for uma colecção finita de arcos de Jordan tais que, a partir do segundo, a origem de cada um coincide com a extremidade do anterior, a sua reunião $C = C_1 \cup C_2 \cup \ldots \cup C_n$, constitui um arco de Jordan seccionalmente liso que se denomina **caminho** ou **contorno**. O contorno diz-se **contorno fechado** se a extremidade de C_n coincide com a origem de C_1. Se os arcos componentes de um contorno forem segmentos de recta (que são arcos lisos de Jordan), então o contorno denomina-se **poligonal**.

As definições anteriores aplicam-se a curvas de $\mathbb{R}^n$, qualquer que seja $n \geq 2$.

9.2 Curvas em $\mathbb{R}^2$

De acordo com a definição de Jordan, curva em $\mathbb{R}^2$ é toda a imagem contínua em $\mathbb{R}^2$ de um segmento da recta real (ou da recta), isto é, o conjunto C dos pontos (x, y) do plano cartesiano cujas coordenadas são funções contínuas de um parâmetro real t definidas num intervalo $[a, b] \subset \mathbb{R}$, tais que

$$x = x(t), \quad y = y(t). \tag{9.2}$$

As equações (9.2) chamam-se as **equações paramétricas** da curva e definem a **representação paramétrica** da curva.

Designando por $\overrightarrow{r}$ o vector que une a origem das coordenadas ao ponto genérico (x, y) da curva, a representação paramétrica (9.2) pode ser substituída pela **representação vectorial**

$$\overrightarrow{r} = x(t)\,\overrightarrow{i} + y(t)\,\overrightarrow{j}, \quad a \leq t \leq b$$

ou abreviadamente

$$\overrightarrow{r} = \overrightarrow{r}(t), \quad a \leq t \leq b.$$

A equação $\overrightarrow{r} = \overrightarrow{r}(t)$ denomina-se a **equação vectorial** da curva. Muitas vezes uma curva de $\mathbb{R}^2$ é representada por uma equação da forma

$$y = f(x), \quad a \leq x \leq b$$

que se denomina a **equação cartesiana** da curva. Esta representação corresponde à representação paramétrica

$$\begin{aligned} x &= t \\ y &= f(t), \quad a \leq t \leq b. \end{aligned}$$

Inversamente, se uma curva for dada pela representação paramétrica

$$x = x(t), \quad y = y(t), \quad a \leq t \leq b$$

para se obter a sua representação cartesiana $y = f(x)$, basta eliminar t nas equações paramétricas.

9.2.1 Tangente em $\mathbb{R}^2$

Seja

$$\overrightarrow{r}(t) = x(t)\overrightarrow{i} + y(t)\overrightarrow{j}, \quad a \leq t \leq b$$

uma curva lisa simples e P o seu ponto de coordenadas

$$x_0 = x(t_0), \ y_0 = y(t_0), \ t_0 \in [a, b].$$

Como se sabe, a equação da recta do plano cartesiano que passa pelo ponto $P(x_0, y_0)$ é

$$\frac{x - x_0}{m} = \frac{y - y_0}{n}$$

em que m e n são quantidades proporcionais aos co-senos directores da recta. Ora, como a derivada

$$\frac{d\overrightarrow{r}}{dt}(t_0) = x'(t_0)\overrightarrow{i} + y'(t_0)\overrightarrow{j}$$

é um vector orientado segundo a tangente à curva dada em $P(x_0, y_0)$, então a equação cartesiana da tangente à curva nesse ponto é

$$\frac{x - x_0}{x'(t_0)} = \frac{y - y_0}{y'(t_0)} \tag{9.3}$$

ou

$$y - y_0 = \frac{y'(t_0)}{x'(t_0)}(x - x_0).$$

As equações paramétricas da tangente são, por seu turno

$$\begin{aligned} x &= x(t_0) + \lambda x'(t_0) \\ y &= y(t_0) + \lambda y'(t_0), \quad -\infty < \lambda < \infty \end{aligned} \tag{9.4}$$

e a equação vectorial

$$\overrightarrow{r} = x(t_0)\overrightarrow{i} + y(t_0)\overrightarrow{j} + \lambda[x'(t_0)\overrightarrow{i} + y'(t_0)\overrightarrow{j}] \tag{9.5}$$

ou

$$\overrightarrow{r} = \overrightarrow{r}(t_0) + \lambda\overrightarrow{r}'(t_0), \quad -\infty < \lambda < \infty.$$

9.2.2 Normal em $\mathbb{R}^2$

Chama-se **normal** a uma curva lisa num ponto $P(x_0, y_0)$ à recta que passa por esse ponto e é perpendicular à tangente à curva nesse ponto. Se a curva é dada na forma paramétrica

$$x = x(t), \ y = y(t), \ a \leq t \leq b,$$

de (9.3) resulta imediatamente que a equação cartesiana da normal à curva no ponto $P(x_0, y_0)$ é

$$y - y_0 = -\frac{x'(t_0)}{y'(t_0)}(x - x_0).$$

Como os vectores $x'(t_0)\overrightarrow{i} + y'(t_0)\overrightarrow{j}$ e $y'(t_0)\overrightarrow{i} - x'(t_0)\overrightarrow{j}$ são perpendiculares, uma equação vectorial da normal à curva no ponto $P(x_0, y_0)$ é

$$\overrightarrow{r} = x(t_0)\overrightarrow{i} + y(t_0)\overrightarrow{j} + \mu[y'(t_0)\overrightarrow{i} - x'(t_0)\overrightarrow{j}], \quad -\infty < \mu < \infty. \tag{9.6}$$

Exemplo 9.1. Determinar uma equação vectorial da tangente e da normal à curva $\overrightarrow{r}(t) = 4t^2\overrightarrow{i} + 4t\overrightarrow{j}$ no ponto correspondente a $t = 1$.

Resolução. De (9.5) resulta para equação da tangente

$$\overrightarrow{r} = 4\overrightarrow{i} + 4\overrightarrow{j} + \lambda(8\overrightarrow{i} + 4\overrightarrow{j})$$

e de (9.6) para equação da normal

$$\overrightarrow{r} = 4\overrightarrow{i} + 4\overrightarrow{j} + \mu(4\overrightarrow{i} - 8\overrightarrow{j}). \quad \blacktriangleleft$$

9.2.3 Rectificação de uma curva em $\mathbb{R}^2$

Seja C uma curva lisa de $\mathbb{R}^2$ definida parametricamente pelas equações

$$\begin{aligned} x &= x(t) \\ y &= y(t), \ a \le t \le b. \end{aligned}$$

A cada partição P do intervalo $[a, b]$ por meio de pontos

$$t_0 = a < t_1 < \ldots < t_i < \\ < \ldots < t_n = b$$

corresponde uma linha poligonal inscrita na curva, com vértices

$$A_i = (x(t_i), y(t_i)), \ i = 0, 1, \ldots, n.$$

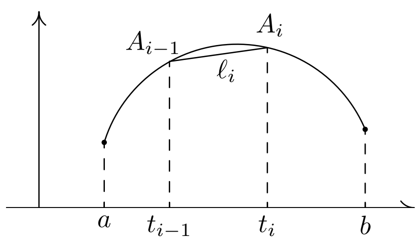

Figura 9.2:

Designando por l_i o comprimento do lado da poligonal definido pelos pontos A_{i-1} e A_i logo se vê que

$$l_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$$

em que, para maior simplicidade na escrita, se fez $x_i = x(t_i)$ e $y_i = y(t_i)$. O perímetro da poligonal será pois

$$S_n = \sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}.$$

Variando a partição do intervalo $[a, b]$ considerada, variará também em geral o perímetro S_n da correspondente poligonal. Em particular, quando a norma da partição tender para zero (o que implica que n tende para infinito) a poligonal tende a ajustar-se à curva e o seu perímetro S_n cresce.

O supremo dos perímetros de todas as poligonais que se podem inscrever na curva é, por definição, o **comprimento** da curva. A curva diz-se **rectificável** se esse supremo for finito, isto é, se existir, finito e determinado, o limite

$$S = \lim_{||P|| \to 0} \sum_{i=1}^{n} l_i.$$

Se a curva for lisa, então, de acordo com o **teorema dos acréscimos finitos de Lagrange** , existem entre t_{i-1} e t_i pontos intermediários θ_i e ϵ_i tais que

$$\begin{aligned} l_i &= \sqrt{[x'(\theta_i)]^2(t_i - t_{i-1})^2 + [y'(\epsilon_i)]^2(t_i - t_{i-1})^2} \\ &= \sqrt{[x'(\theta_i)]^2 + [y'(\epsilon_i)]^2} \cdot \Delta t_i \end{aligned}$$

pondo $\Delta t_i = t_i - t_{i-1}$.

Como $x'(t)$ e $y'(t)$ são contínuas, visto a curva ser lisa, o radical anterior é também uma função contínua e portanto o limite

$$S = \lim_{||P|| \to 0} \sum_{i=1}^{n} l_i$$

existe e é o integral definido

$$S = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2}\ dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\ dt.$$

Se na expressão anterior se deixar o limite superior de integração variável no intervalo $[a, b]$ obtém-se a função integral

$$S = f(t) = \int_a^t \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\ dt \qquad (9.7)$$

que, como se sabe, é derivável e tem por derivada a função integranda nos pontos em que esta é contínua. Ora, como no caso presente a função integranda é contínua em todo o intervalo $[a, b]$ tem-se então

$$\frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$

de onde resulta

$$ds^2 = dx^2 + dy^2,$$

expressão que se denomina o **diferencial do arco** de curva.

Se a curva for dada na forma cartesiana $y = f(x)$ a fórmula para a sua rectificação passa a ser

$$s = \int_a^b \sqrt{1 + (y')^2}\ dx.$$

Pode demonstrar-se que o comprimento de uma curva é independente da representação paramétrica utilizada, isto é, todas as representações equivalentes conduzem ao mesmo valor para o comprimento.

Fixando a origem A de uma curva, a posição de qualquer outro ponto P da curva pode ser determinado em função do comprimento do arco AP, isto é, através da fórmula (9.7) e por isso se chama **abcissa curvilínea** de P à variável s definida pela integral

$$S = f(t) = \int_a^t \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

Assim, os pontos da curva passam a ser determinados pelo parâmetro s, chamado **parâmetro intrínseco** dos pontos da curva. Utilizando esta mudança de parâmetro obtém-se para a curva a representação paramétrica

$$\overrightarrow{r} = \overrightarrow{r}(s)$$

que é equivalente à representação $\overrightarrow{r} = \overrightarrow{r}(t)$ (já que sendo a função $s = f(t)$ invertível se pode obter $t = f^{-1}(s)$).

Da regra de derivação da função composta resulta

$$\frac{d\overrightarrow{r}}{ds} = \frac{d\overrightarrow{r}}{dt}\frac{dt}{ds} = \frac{\dfrac{d\overrightarrow{r}}{dt}}{\dfrac{ds}{dt}}.$$

Ora, como

$$\frac{d\overrightarrow{r}}{dt} = \frac{dx}{dt}\overrightarrow{i} + \frac{dy}{dt}\overrightarrow{j}$$

é um vector tangente à curva e, por outro lado,

$$\frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$

é igual à norma desse vector, segue-se que

$$\frac{d\overrightarrow{r}}{ds} = \frac{\dfrac{d\overrightarrow{r}}{dt}}{\left\|\dfrac{d\overrightarrow{r}}{dt}\right\|}$$

é um **vector unitário** (ou versor) tangente à curva no ponto considerado. Tal vector notar-se-á $\overrightarrow{t}$, escrevendo-se

$$\overrightarrow{t} = \frac{d\overrightarrow{r}}{ds} \quad \text{ou} \quad \overrightarrow{t} = \frac{\dfrac{d\overrightarrow{r}}{dt}}{\left\|\dfrac{d\overrightarrow{r}}{dt}\right\|}.$$

9.2.4 Curvatura duma curva. Raio de curvatura. Evoluta

Dado um arco da curva lisa $\widehat{AB}$, considerem-se as tangentes à curva nos ponto A e B (vd. Figura 9.3). Designando por Δs o comprimento do arco e por $\Delta\alpha$ o ângulo de rotação da tangente quando se passa do ponto A ao ponto B, chama-se **curvatura média** K_m do arco $\widehat{AB}$ ao quociente

$$K_m = \frac{\Delta\alpha}{\Delta s}.$$

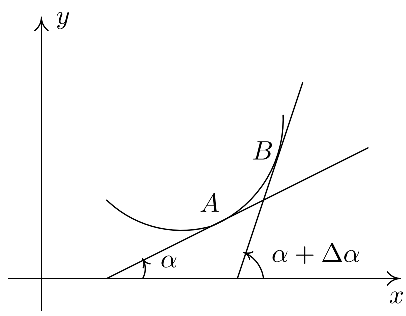

Figura 9.3:

A curvatura média permite caracterizar o grau de encurvamento dum arco mas já não permite caracterizar o encurvamento na vizinhança de um ponto. Para isso introduz-se a noção de curvatura num ponto.

Chama-se **curvatura** duma curva no ponto A, e nota-se K_A, ao limite da curvatura média do arco $\widehat{AB}$, quando B tende para A (ou seja, quando o comprimento Δs do arco tende para zero):

$$K_A = \lim_{\Delta s \to 0} K_m = \lim_{\Delta s \to 0} \frac{\Delta \alpha}{\Delta s}.$$

Como num sistema de coordenadas cartesianas α e s são ambos funções de x, pode-se considerar α como uma função de s expressa por equações paramétricas em x. Então

$$K_A = \lim_{\Delta s \to 0} \frac{\Delta \alpha}{\Delta s} = \frac{d\alpha}{ds} = \frac{\frac{d\alpha}{dx}}{\frac{ds}{dx}}.$$

Suponhamos que a curva é dada em coordenadas cartesianas por uma equação da forma $y = f(x)$ e que $f(x)$ é de classe C^2. Como, atendendo ao significado geométrico da noção de derivada, é $\alpha = \operatorname{arctg} y'$, tem-se

$$\frac{d\alpha}{dx} = \frac{y''}{1 + y'^2}.$$

Por outro lado, como já se viu, é $\dfrac{ds}{dx} = \sqrt{1 + y'^2}$ e por consequência resulta

$$K = \frac{d\alpha}{dx} = \frac{\frac{d\alpha}{dx}}{\frac{ds}{dx}} = \frac{y''}{\sqrt{(1 + y'^2)^3}} \tag{9.8}$$

para a expressão geral da curvatura num ponto.

Exemplo 9.2. Determinar a curvatura da recta $y = ax + b$ num ponto genérico $A(x, y)$.

Resolução. Como $y' = a$ e $y'' = 0$, substituindo na fórmula (9.8) tem-se $K = 0$. A recta é, portanto, uma curva de curvatura nula. ◀

Suponhamos agora que a curva é definida parametricamente pelas equações

$$x = \varphi(t), \quad y = \psi(t), \quad a \leq t \leq b$$

e que $\varphi(t)$ e $\psi(t)$ são funções de classe C^2. Neste caso é $\alpha = \operatorname{arctg}\dfrac{\psi'}{\varphi'}$ e portanto

$$\frac{d\alpha}{dt} = \frac{\varphi'\psi'' - \psi'\varphi''}{\varphi'^2 + \psi'^2}.$$

Como, por outro lado, é $\dfrac{ds}{dt} = \sqrt{\varphi'^2 + \psi'^2}$, resulta então para expressão da curvatura

$$K = \frac{\varphi'\psi'' - \psi'\varphi''}{\sqrt{(\varphi'^2 + \psi'^2)^3}}. \tag{9.9}$$

Exemplo 9.3. Determinar a curvatura num ponto genérico da ciclóide $x = a(t - \operatorname{sen} t)$, $y = a(1 - \cos t)$.

Resolução. Como $\dfrac{dx}{dt} = a(1 - \cos t)$; $\dfrac{d^2x}{dt^2} = a \operatorname{sen} t$; $\dfrac{dy}{dt} = a \operatorname{sen} t$ e $\dfrac{d^2y}{dt^2} = a \cos t$, substituindo na fórmula (9.9) tem-se

$$K = \frac{1}{4a \operatorname{sen} \dfrac{t}{2}}. \quad ◀$$

Donde, chama-se **raio de curvatura** de uma curva num ponto A ao módulo do inverso da curvatura nesse ponto, isto é

$$R = \frac{1}{|K|}.$$

Se sobre a normal à curva no ponto A se marcar, no sentido da concavidade, um comprimento igual ao raio de curvatura obtém-se um ponto que se denomina **centro de curvatura** da curva no ponto A.

O círculo com centro no centro de curvatura e raio igual ao raio de curvatura no ponto A chama-se o **círculo de curvatura da curva** no ponto A (a correspondente circunferência chama-se **circunferência osculatriz** no ponto A e é evidentemente tangente à curva nesse ponto).

Se a curva for definida cartesianamente pela equação $y = f(x)$, as coordenadas (α, β) do centro de curvatura são dadas pelas expressões

$$\begin{aligned} \alpha &= x - \frac{y'(1+y'^2)}{y''} \\ \beta &= y + \frac{1+y'^2}{y''}. \end{aligned} \tag{9.10}$$

O lugar geométrico dos centros de curvatura de uma curva constitui uma nova curva que se denomina **evoluta** da curva considerara.

Se a curva for dada pela equação $y = f(x)$, as equações paramétricas da curva são precisamente as equações (9.10) tomando x como parâmetro.

Exemplo 9.4. Determinar a equação da evoluta da parábola

$$y = ax^2.$$

Resolução. Como $y' = 2ax$ e $y'' = 2a$, as coordenadas do centro de curvatura num ponto genérico de abcissa x são de acordo com (9.10)

$$\alpha = -4a^2x^3, \quad \beta = \frac{1+6a^2x^2}{2a}.$$

Estas equações são precisamente as equações paramétricas da evoluta da parábola considerada. ◀

9.3 Curvas em $\mathbb{R}^3$

Como foi já referido anteriormente, chama-se curva contínua de $\mathbb{R}^3$ ao contradomínio de uma aplicação contínua de $\mathbb{R}$ em $\mathbb{R}^3$, isto é, ao conjunto C dos pontos (x, y, z) determinados por um sistema de funções contínuas

$$\begin{cases} x = x(t) \\ y = z(t) \\ z = z(t) \end{cases} \tag{9.11}$$

definidas num intervalo $a \leq t \leq b$.

As equações (9.11) chamam-se equações paramétricas da curva e constituem a sua representação paramétrica , à qual corresponde a representação vectorial

$$\overrightarrow{r}(t) = x(t)\,\overrightarrow{i} + y(t)\,\overrightarrow{j} + z(t)\,\overrightarrow{k}. \tag{9.12}$$

Uma curva de $\mathbb{R}^3$ pode também definir-se como sendo o lugar geométrico dos pontos de intersecção de duas superfícies

$$\begin{cases} F(x, y, z) = 0 \\ \\ G(x, y, z) = 0. \end{cases} \tag{9.13}$$

Por exemplo, as equações $x^2 + y^2 + z^2 = 9, \quad z = 2$ que representam uma esfera e um plano, definem, quando consideradas simultaneamente, uma circunferência.

9.3.1 Tangente em $\mathbb{R}^3$

Se $\overrightarrow{r}(t) = x(t)\overrightarrow{i} + y(t)\overrightarrow{j} + z(t)\overrightarrow{k}$, $a \le t \le b$ for uma representação vectorial de uma curva lisa de $\mathbb{R}^3$, então existe a derivada em $t_0 \in [a, b]$

$$\frac{d\overrightarrow{r}}{dt}(t_0) = \overrightarrow{r}(t_0) = x'(t_0)\overrightarrow{i} + y'(t_0)\overrightarrow{j} + z'(t_0)\overrightarrow{k}$$

que é, como se sabe, um vector orientado segundo a tangente à curva no ponto de coordenadas $x_0 = x(t_0)$, $y_0 = y(t_0)$, $z_0 = z(t_0)$.

A equação da tangente nesse ponto é portanto

$$\frac{x - x_0}{x'(t_0)} = \frac{y - y_0}{y'(t_0)} = \frac{z - z_0}{y'(t_0)}. \tag{9.14}$$

A correspondente representação paramétrica é então

$$\begin{cases} x = x(t_0) + \lambda x'(t_0) \\ y = y(t_0) + \lambda y'(t_0), \quad -\infty < \lambda < \infty \\ z = z(t_0) + \lambda z'(t_0) \end{cases} \tag{9.15}$$

e a representação vectorial

$$\overrightarrow{r} = \overrightarrow{r}(t_0) + \lambda \overrightarrow{r}'(t_0), \quad -\infty < \lambda < \infty. \tag{9.16}$$

Resta agora considerar o caso em que a curva é definida como a intersecção de duas superfícies

$$\begin{cases} F(x, y, z) = 0 \\ G(x, y, z) = 0. \end{cases} \tag{9.17}$$

Se nos pontos que verificam simultaneamente as duas equações as funções F e G admitem derivadas parciais de primeira ordem contínuas e se for $\dfrac{\partial(F,G)}{\partial(x,y)} \neq 0$ num ponto (x_0, y_0, z_0), então o teorema das funções implícitas permite concluir que numa vizinhança desse ponto as equações (9.17) definem x e y como funções de z de classe C^1. Nessa vizinhança fica pois definida uma curva lisa com uma representação paramétrica da forma

$$\begin{cases} x = x(z) \\ y = y(z) \\ z = z. \end{cases}$$

Para determinar a equação da tangente a esta curva no ponto (x_0, y_0, z_0), há que determinar as derivadas em ordem a z das funções $x(z)$ e $y(z)$ nesse ponto.

Ora, de acordo com o teorema das funções implícitas

$$\frac{dx}{dz} = -\frac{\dfrac{\partial(F,G)}{\partial(z,y)}}{\dfrac{\partial(F,G)}{\partial(x,y)}} \quad \text{e} \quad \frac{dy}{dz} = -\frac{\dfrac{\partial(F,G)}{\partial(x,z)}}{\dfrac{\partial(F,G)}{\partial(x,y)}}.$$

Como, por outro lado, $\dfrac{dz}{dz} = 1$, a equação cartesiana da tangente à curva em (x_0, y_0, z_0), é então

$$\frac{x - x_0}{\dfrac{\partial(F,G)}{\partial(y,z)}} = \frac{y - y_0}{\dfrac{\partial(F,G)}{\partial(z,x)}} = \frac{z - z_0}{\dfrac{\partial(F,G)}{\partial(x,y)}} \tag{9.18}$$

sendo os jacobianos calculadas em (x_0, y_0, z_0); sendo a fórmula válida quando pelo menos um dos jacobianos é diferente de zero. Se num ponto da curva os três jacobianos se anularem simultaneamente, o ponto chama-se ponto singular e a curva não tem tangente nesse ponto.

De (9.18) resulta imediatamente a representação paramétrica

$$\begin{aligned} x &= x_0 + \lambda \cdot \frac{\partial(F,G)}{\partial(y,z)} \\ y &= y_0 + \lambda \cdot \frac{\partial(F,G)}{\partial(z,x)}, \quad -\infty < \lambda < \infty \\ z &= z_0 + \lambda \cdot \frac{\partial(F,G)}{\partial(x,y)} \end{aligned} \tag{9.19}$$

e a representação vectorial

$$\overrightarrow{r} = \overrightarrow{r}(x_0, y_0, z_0) + \lambda \left(\frac{\partial(F,G)}{\partial(y,z)} \overrightarrow{i} + \frac{\partial(F,G)}{\partial(z,x)} \overrightarrow{j} + \frac{\partial(F,G)}{\partial(x,y)} \overrightarrow{k} \right), \tag{9.20}$$

com $-\infty < \lambda < \infty$.

Exemplo 9.5. Determinar a equação cartesiana da tangente à curva $x = t - \cos t$, $y = 3 + \operatorname{sen} 2t$, $z = 1 + \cos 3t$ no ponto em que $t = \dfrac{\pi}{2}$.

Resolução. Como $x'(\pi/2) = 2$, $y'(\pi/2) = -2$ e $z'(\pi/2) = 3$, a equação pretendida é

$$\frac{x - \pi/2}{2} = \frac{y-3}{-2} = \frac{z-1}{3}. \quad \blacktriangleleft$$

Exemplo 9.6. Determinar a equação da tangente à curva

$$x + y + z = 3, \; x^2 - y^2 + 2z^2 = 2$$

no ponto $P(1, 1, 1)$.

Resolução. Pondo $F(x,y,z) = x + y + z - 3 = 0$ e

$$G(x,y,z) = x^2 - y^2 + 2z^2 - 2 = 0$$

tem-se

$$\left(\frac{\partial(F,G)}{\partial(y,z)}\right)_P = -6, \quad \left(\frac{\partial(F,G)}{\partial(z,x)}\right)_P = 2, \quad \left(\frac{\partial(F,G)}{\partial(x,y)}\right)_P = 4$$

e portanto a equação da tangente à curva é

$$\frac{x-1}{-6} = \frac{y-1}{2} = \frac{z-1}{4}.$$

A correspondente representação vectorial é

$$\overrightarrow{r} = (\overrightarrow{i} + \overrightarrow{j} + \overrightarrow{k}) + \lambda(-6\overrightarrow{i} + 2\overrightarrow{j} + 4\overrightarrow{k})$$

e a representação paramétrica é

$$\begin{aligned} x &= 1 - 6\lambda, \\ y &= 1 + 2\lambda, \\ z &= 1 + 4\lambda. \quad \blacktriangleleft \end{aligned}$$

9.3.2 Normal e plano normal em $\mathbb{R}^3$

Se uma curva de $\mathbb{R}^3$ tem tangente num dado ponto, chama-se **normal** à curva nesse ponto a toda a recta que seja perpendicular à tangente no ponto de tangência. Como é evidente, há uma infinidade de normais à curva nesse ponto, as quais estão todas situadas num mesmo plano (perpendicular à tangente à curva) que se denomina **plano normal** à curva no ponto considerado.

Como se sabe, a equação do plano que passa por um ponto (x_0, y_0, z_0), e é perpendicular a uma recta de parâmetros directores A, B e C é

$$A(x - x_0) + B(y - x_0) + C(z - z_0) = 0.$$

Então, a equação do plano normal à curva $\overrightarrow{r} = \overrightarrow{r}(t)$ no ponto (x_0, y_0, z_0), é, atendendo à expressão (9.14) a tangente

$$x'(t_0)(x - x_0) + y'(t_0)(y - y_0) + z'(t_0)(z - z_0) = 0. \qquad (9.21)$$

Exemplo 9.7. Determinar as equações da tangente e do plano normal à curva

$$\overrightarrow{r}(t) = t\,\overrightarrow{i} + t^2\,\overrightarrow{j} + t^3\,\overrightarrow{k}$$

no ponto $(3, 9, 27)$.

Resolução. O ponto em causa é o que corresponde a $t = 3$. Como $x'(3) = 1$, $y'(3) = 6$, $z'(3) = 27$ a equação da tangente é

$$\frac{x-3}{1} = \frac{y-9}{6} = \frac{z-27}{27},$$

e a do plano normal

$$(x - 3) + 6(y - 9) + 27(z - 27) = 0$$

ou

$$x + 6y + 27z = 786. \quad \blacktriangleleft$$

9.3.3 Rectificação de curvas em $\mathbb{R}^3$

O que se referiu anteriormente (parágrafo 9.2.3) acerca da rectificação de curvas de $\mathbb{R}^2$ estende-se sem dificuldades às curvas de $\mathbb{R}^3$.

Assim, se C é uma curva de $\mathbb{R}^3$ definida por

$$x = x(t), \quad y = y(t), \quad z = z(t), \quad a \leq t \leq b$$

e $x(t)$, $y(t)$, $z(t)$ são funções de classe C^1, o seu comprimento é

$$s = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt. \tag{9.22}$$

Para o arco da curva compreendido entre a extremidade $P(a)$ e o ponto variável $P(t)$ o comprimento é dado pela função

$$s(t) = \int_a^t \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt \tag{9.23}$$

que é uma função derivável de t.

Derivando obtém-se:

$$\frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} \tag{9.24}$$

de onde resulta

$$ds = \sqrt{dx^2 + dy^2 + dz^2} \tag{9.25}$$

que se denomina o diferencial do arco.

Fixando a origem $P(a)$ do arco e imaginando sobre ele um ponto variável $P(t)$, a posição de P pode, portanto, determinar-se através da fórmula (9.23) e, por isso, se chama **abcissa curvilínea** de P à grandeza $s(t)$ definida por essa fórmula. Nessa altura a representação vectorial da curva passa a ser $\overrightarrow{r} = \overrightarrow{r}(s)$.

Tal como em $\mathbb{R}^2$, o vector

$$\frac{d\overrightarrow{r}}{ds} = \frac{dx}{ds}\overrightarrow{i} + \frac{dy}{ds}\overrightarrow{j} + \frac{dz}{ds}\overrightarrow{k}$$

é um vector unitário tangente à curva que se representa por $\overrightarrow{t}$

$$\overrightarrow{t} = \frac{d\overrightarrow{r}}{ds} = \frac{d\overrightarrow{r}}{dt}\frac{dt}{ds} = \frac{\frac{d\overrightarrow{r}}{dt}}{\frac{ds}{dt}} = \frac{\frac{d\overrightarrow{r}}{dt}}{\left\|\frac{d\overrightarrow{r}}{dt}\right\|}. \tag{9.26}$$

9.3.4 Fórmulas de Frenet

Como é evidente, o vector unitário $\overrightarrow{t}$ é variável em direcção de ponto para ponto da curva, pelo que pode considerar-se como função do arco $\widehat{AB}$, ou seja, do parâmetro intrínseco s. E, como a derivada dum vector de módulo constante é um vector perpendicular ao dado, derivando $\overrightarrow{t}$ em ordem a s obtém-se um vector de direcção perpendicular a $\overrightarrow{t}$. Designando por $\overrightarrow{n}$ o vector unitário que define a direcção e sentido desse vector e por k a sua grandeza, tem-se então

$$\frac{d\overrightarrow{t}}{ds} = k\overrightarrow{n} \tag{9.27}$$

que se denomina **primeira fórmula de Frenet**.

Ao valor absoluto de k chama-se a **curvatura de flexão** da curva no ponto considerado e a $\frac{1}{|K|}$ o **raio de curvatura de flexão** . O vector $\overrightarrow{n}$ é evidentemente normal à curva no ponto considerado; a recta que passa por esse ponto e tem a direcção de $\overrightarrow{n}$ chama-se a **normal principal** nesse ponto.

Ao plano definido por $\overrightarrow{t}$ e por $\overrightarrow{n}$, que é tangente à curva no ponto P de aplicação desses vectores visto conter $\overrightarrow{t}$, chama-se o **plano osculador** à curva em P. O plano osculador a uma curva num dado ponto é, pois, o plano definido pela tangente e pela **normal principal** à curva nesse ponto.

Considere-se no ponto P um terceiro vector unitário $\overrightarrow{b}$, de modo que o triedro definido por $\overrightarrow{t}$, $\overrightarrow{n}$ e $\overrightarrow{b}$ tenha a mesma orientação que o triedro definido pelos vectores $\overrightarrow{i}$, $\overrightarrow{j}$, $\overrightarrow{k}$. O vector unitário $\overrightarrow{b}$, que é também normal à curva em P, denomina-se versor da **binormal**.

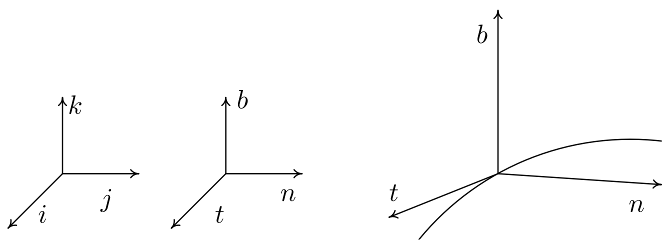

Figura 9.4:

O triedro assim definido é o **triedro de Frenet** (ou intrínseco) no ponto considerado. O plano definido por $\overrightarrow{b}$ e por $\overrightarrow{t}$ é o **plano rectificante** (Figura 9.5).

Como é evidente, é

$$\overrightarrow{t} = \overrightarrow{n} \times \overrightarrow{b}; \quad \overrightarrow{n} = \overrightarrow{b} \times \overrightarrow{t} \quad \text{e} \quad \overrightarrow{b} = \overrightarrow{t} \times \overrightarrow{n}.$$

Como $\overrightarrow{b}$ e $\overrightarrow{t}$ são perpendiculares entre si é $\overrightarrow{b} \cdot \overrightarrow{t} = 0$ e, portanto, derivando esta igualdade tem-se

$$\frac{d\overrightarrow{b}}{ds} \cdot \overrightarrow{t} + \overrightarrow{b} \cdot \frac{d\overrightarrow{t}}{ds} = 0$$

e atendendo ao valor de $\dfrac{d\overrightarrow{t}}{ds}$

$$\frac{d\overrightarrow{b}}{ds}\cdot\overrightarrow{t}=-\overrightarrow{b}\cdot(k\overrightarrow{n})=0$$

visto que $\overrightarrow{b}$ e $\overrightarrow{n}$ são perpendiculares.

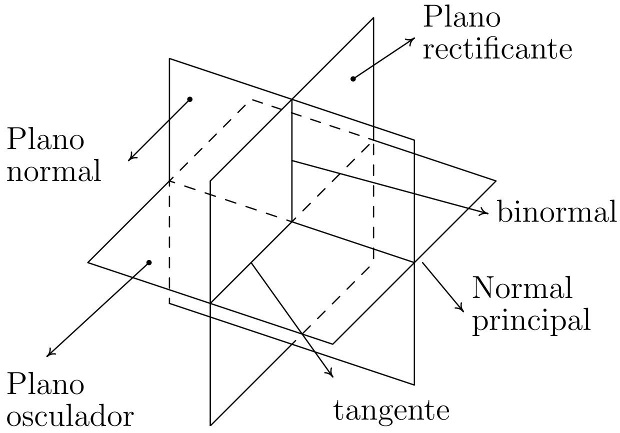

Figura 9.5:

A relação anterior mostra pois que $\dfrac{d\overrightarrow{b}}{ds}$ é perpendicular a $\overrightarrow{t}$; mas como é também perpendicular $\overrightarrow{b}$, há-de ter forçosamente a direcção de $\overrightarrow{n}$. Designado por T a sua grandeza tem-se portanto

$$\frac{d\overrightarrow{b}}{ds}=T\overrightarrow{n} \tag{9.28}$$

que é a **segunda fórmula de Frenet**.

O número T definido por esta fórmula chama-se a **curvatura de torção** (ou simplesmente torção) da curva no ponto considerado; ao inverso da torção, $\dfrac{1}{T}$, chama-se **raio de torção**.

Como é evidente, se uma curva for plana a sua torção é nula em todos os pontos. Se a torção não for nula a curva diz-se **curva empenada** ou **torça**.

Como $\overrightarrow{n} = \overrightarrow{b} \times \overrightarrow{t}$, derivando esta igualdade obtém-se

$$\frac{d\overrightarrow{n}}{ds} = \frac{d\overrightarrow{b}}{ds} \times \overrightarrow{t} + \overrightarrow{b} \times \frac{d\overrightarrow{t}}{ds} = T(\overrightarrow{n} \times \overrightarrow{t}) + k(\overrightarrow{b} \times \overrightarrow{n}) = -T\overrightarrow{b} - k\overrightarrow{t}$$

ou seja

$$\frac{d\overrightarrow{n}}{ds} = -k\overrightarrow{t} - T\overrightarrow{b} \tag{9.29}$$

que é a **terceira fórmula de Frenet.**

9.3.5 Equações da normal principal e da binormal

Como a **normal principal** à curva $\overrightarrow{r} = \overrightarrow{r}(s)$, no ponto $P_0(x_0, y_0, z_0)$ é a recta que passa por esse ponto e tem a direcção do vector unitário $\overrightarrow{n}$, a sua equação vectorial é

$$\overrightarrow{r} - \overrightarrow{r}_0 = \lambda\overrightarrow{n} \tag{9.30}$$

de onde resulta, tendo em conta que

$$\overrightarrow{n} = \frac{1}{k}\frac{d\overrightarrow{t}}{ds} = \frac{1}{k}\left(\frac{d^2x}{ds^2}\overrightarrow{i} + \frac{d^2y}{ds^2}\overrightarrow{j} + \frac{d^2z}{ds^2}\overrightarrow{k}\right),$$

a expressão cartesiana

$$\frac{x - x_0}{x''(s_0)} = \frac{y - y_0}{y''(s_0)} = \frac{z - z_0}{z''(s_0)} \tag{9.31}$$

em que as derivadas são calculadas em relação ao parâmetro intrínseco s.

Se a curva estiver definida, não em função do parâmetro intrínseco s, mas de outro parâmetro t, a aplicação da fórmula (9.26) permite passar das derivadas em ordem a s para as derivadas em ordem a t e obter assim uma expressão formalmente igual à (9.31), mas com derivadas em relação a t.

Finalmente para a **recta binormal** , cuja direcção é a de $\overrightarrow{b}$, tem-se

$$\overrightarrow{r} - \overrightarrow{r}_0 = \lambda \overrightarrow{b} \tag{9.32}$$

ou

$$\overrightarrow{r} - \overrightarrow{r}_0 = \lambda(\overrightarrow{t} \times \overrightarrow{n})$$

ou ainda

$$\overrightarrow{r} - \overrightarrow{r}_0 = \lambda \begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ x'_0 & y'_0 & z'_0 \\ x''_0 & y''_0 & z''_0 \end{vmatrix}$$

Desta expressão resulta a equação cartesiana da recta binormal

$$\frac{x - x_0}{y'_0 z''_0 - y''_0 z'_0} = \frac{y - y_0}{x''_0 z'_0 - x'_0 z''_0} = \frac{z - z_0}{x'_0 y''_0 - x''_0 y'_0}. \tag{9.33}$$

Exemplo 9.8. Determinar as equações cartesianas das rectas tangente, normal principal e binormal no ponto correspondente a $t = t_0$

da hélice cilíndrica $\begin{cases} x = a\cos t \\ y = a \text{ sen } t, \quad -\infty \leq t \leq \infty \\ z = bt \end{cases}$ com a e b constantes.

Resolução. Como em $t = t_0$ é

$$\begin{array}{lll} x_0 = a\cos t_0, & x_0' = -a \text{ sen } t_0, & x_0'' = -a\cos t_0 \\ y_0 = a \text{ sen } t_0, & y_0' = a\cos t_0, & y_0'' = -a \text{ sen } t_0 \\ z_0 = bt_0, & z_0' = b, & z_0'' = 0 \end{array}$$

e

$$\begin{array}{l} y_0'z_0'' - y_0''z_0' = ab\sin t_0, \\ x_0''z_0' - x_0'z_0'' = -ab\cos t_0, \\ x_0'y_0'' - x_0''y_0' = a^2 \end{array}$$

substituindo estes valores em (9.14), (9.31) e (9.33) obtêm-se, respectivamente, as equações cartesianas das rectas

Tangente: $\dfrac{x - x_0}{-a \text{ sen } t_0} = \dfrac{y - y_0}{a\cos t_0} = \dfrac{z - z_0}{b}$

Normal Principal: $\begin{cases} \dfrac{x - x_0}{\cos t_0} = \dfrac{y - y_0}{\text{sen } t_0} \\ z - z_0 = 0 \end{cases}$

Binormal: $\dfrac{x - x_0}{b \text{ sen } t_0} = \dfrac{y - y_0}{-b\cos t_0} = \dfrac{z - z_0}{a}$ ◀

9.3.6 Equações dos planos normal, rectificante e osculador

De acordo com a sua própria definição, o plano normal num ponto $P_0(x_0, y_0, z_0)$ à curva $\overrightarrow{r} = \overrightarrow{r}(s)$ é o lugar geométrico dos pontos com vector posicional $\overrightarrow{r}$ tais que o vector $\overrightarrow{r} - \overrightarrow{r}_0$ seja perpendicular ao vector $\overrightarrow{t}$. Portanto, a equação vectorial do **plano normal** em P_0 (com vector posicional $\overrightarrow{r}_0$) é

$$(\overrightarrow{r} - \overrightarrow{r}_0) \cdot \overrightarrow{t} = 0. \tag{9.34}$$

Como

$$\overrightarrow{r} = x(s)\,\overrightarrow{i} + y(s)\,\overrightarrow{j} + z(s)\,\overrightarrow{k}$$

e

$$\overrightarrow{t} = \left(\frac{dx}{ds}\right)_0 \overrightarrow{i} + \left(\frac{dy}{ds}\right)_0 \overrightarrow{j} + \left(\frac{dz}{ds}\right)_0 \overrightarrow{k},$$

em P_0 a expressão cartesiana do plano normal será

$$\left(\frac{dx}{ds}\right)_0 (x - x_0) + \left(\frac{dy}{ds}\right)_0 (y - y_0) + \left(\frac{dz}{ds}\right)_0 (z - z_0) = 0. \tag{9.35}$$

Analogamente, a equação do **plano rectificante** é

$$(\overrightarrow{r} - \overrightarrow{r}_0) \cdot \overrightarrow{n} = 0 \tag{9.36}$$

ou em forma cartesiana

$$\left(\frac{d^2x}{ds^2}\right)_0 (x - x_0) + \left(\frac{d^2y}{ds^2}\right)_0 (y - y_0) + \left(\frac{d^2z}{ds^2}\right)_0 (z - z_0) = 0. \tag{9.37}$$

Para o **plano osculador**, por seu turno tem-se

$$(\overrightarrow{r} - \overrightarrow{r}_0) \cdot \overrightarrow{b} = 0 \tag{9.38}$$

ou

$$(\overrightarrow{r} - \overrightarrow{r}) \cdot (\overrightarrow{t} \times \overrightarrow{n}) = 0$$

de onde resulta a equação cartesiana

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_0' & y_0' & z_0' \\ x_0'' & y_0'' & z_0'' \end{vmatrix} = 0. \tag{9.39}$$

Exemplo 9.9. Determinar as equações cartesianas dos planos normal, rectificante e osculador à curva considerada no Exemplo (9.8) no ponto correspondente a $t = t_0$.

Resolução. Substituindo os valores calculados no exemplo anterior (Exemplo 9.8) nas fórmulas (9.35), (9.37) e (9.39) resultam, respectivamente, as equações dos

Plano normal: $-(a \text{ sen } t_0)x + (a \cos t_0)y + bz = b^2 t_0$

Plano rectificante: $(\cos t_0)x + (\text{ sen } t_0)y = a$

Plano osculador: $(b \text{ sen } t_0)x - (b \cos t_0)y + az = abt_0$. ◀

9.4 Superfícies em $\mathbb{R}^3$

Definição 9.1.

Chama-se superfície no espaço $\mathbb{R}^3$ ao conjunto dos pontos (x, y, z) desse espaço descrito por três funções contínuas

$$\begin{aligned} x &= x(u, v) \\ y &= y(u, v) \\ z &= z(u, v) \end{aligned} \tag{9.40}$$

definidas num dado domínio D fechado de $\mathbb{R}^2$.

Uma superfície é, pois, o contradomínio da transformação contínua definida por (9.40). As equações (9.40) denominam-se **equações paramétricas** da superfície e definem a chamada **representação paramétrica** da superfície.

Se em $\mathbb{R}^3$ estiver fixado um referencial cartesiano ortogonal de (9.40) deduz-see imediatamente a chamada **representação vectorial** da superfície

$$\overrightarrow{r} = \overrightarrow{r}(u, v) = x(u, v)\,\overrightarrow{i} + y(u, v)\,\overrightarrow{j} + z(u, v)\,\overrightarrow{k}\,.$$

No caso particular em que em (9.40) se tomam como parâmetros $x = u$, $y = v$, obtém-se a representação cartesiana explícita

$$z = f(x, y).$$

Uma superfície pode, porém, ser dada também na forma implícita

$$F(x, y, z) = 0.$$

Definição 9.2.

Uma superfície definida parametricamente pelas equações 9.40 diz-se lisa se as funções $x(u,v)$, $y(u,v)$ e $z(u,v)$ são de classe C^1 e, além disso, os Jacobianos

$$\frac{\partial(x,y)}{\partial(u,v)}, \quad \frac{\partial(y,z)}{\partial(u,v)} \quad e \quad \frac{\partial(z,x)}{\partial(u,v)},$$

não se anulam simultaneamente em nenhum ponto de D.

Os pontos de D em que isto acontece denominam-se **pontos ordinários** da superfície; os pontos em que essas condições falham, são **pontos singulares** da superfície.

Se a superfície for definida cartesianamente por uma equação $F(x,y,z) = 0$, então será uma superfície lisa se F for de classe C^1 e as suas primeiras derivadas parciais não se anularem simultaneamente.

Se em (9.40) se fizer $u = f(t)$ e $v = g(t)$, com f e g de classe C^1, obtém-se uma curva sobre a superfície, de equações paramétricas

$$x = x(t), \quad y = y(t), \quad z = z(t).$$

A recta tangente a esta curva num dado ponto $P(x,y,z)$ é, por definição, uma recta tangente à superfície nesse ponto. Como por um mesmo ponto ordinário P da superfície passa uma infinidade de curvas sobre ela situadas, há uma infinidade de rectas tangentes à superfície nesse ponto. Essas rectas situam-se todas, porém, num mesmo plano, como é fácil de provar.

Proposição 9.1. *Todas as rectas tangentes a uma superfície $F(x, y, z) = 0$ num ponto ordinário P estão situadas sobre um mesmo plano.*

Demonstração. Consideremos sobre a superfície uma curva qualquer que passe pelo ponto P, de equações paramétricas

$$x = x(t), \quad y = y(t), \quad z = z(t).$$

A equação da curva pode, pois, escrever-se

$$F[x(t), y(t), z(t)] = 0.$$

Derivando em ordem a t vem

$$\frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} = 0,$$

o que corresponde a

$$\left(\frac{\partial F}{\partial x}\overrightarrow{i} + \frac{\partial F}{\partial y}\overrightarrow{j} + \frac{\partial F}{\partial z}\overrightarrow{k}\right) \cdot \left(\frac{dx}{dt}\overrightarrow{i} + \frac{dy}{dt}\overrightarrow{j} + \frac{dz}{dt}\overrightarrow{k}\right) = 0$$

ou seja

$$\text{grad } F \cdot \frac{d\overrightarrow{r}}{dt} = 0;$$

com grad $F \neq 0$ visto P ser ponto ordinário.

Como $\dfrac{d\overrightarrow{r}}{dt}$ é tangente à curva considerada no ponto P, segue-se que a tangente a essa curva é perpendicular ao vector $(\text{grad } F)_P$.

Mas como a curva considerada era arbitrária, então todas as rectas tangentes à superfície no ponto P são perpendiculares ao mesmo vector $(\text{grad } F)_P$ e, portanto, são complanares, como se pretendia demonstrar. ■

9.4.1 Plano tangente

Definição 9.3 (Plano tangente).

> *Chama-se* ***plano tangente*** *à superfície* $F(x, y, z) = 0$ *num ponto ordinário* $P(x_0, y_0, z_0)$ *ao plano formado por todas as tangentes em* P *às curvas traçadas sobre a superfície que passem por esse ponto.*

De acordo com a proposição anterior, o plano tangente a $F(x, y, z) = 0$ no ponto ordinário P é perpendicular ao vector $(\text{grad } F)_P$. A sua equação é portanto

$$\left(\frac{\partial F}{\partial x}\right)_P (x - x_0) + \left(\frac{\partial F}{\partial y}\right)_P (y - y_0) + \left(\frac{\partial F}{\partial z}\right)_P (z - z_0) = 0.$$

9.4.2 Normal

Definição 9.4 (Normal à superfície).

> *Chama-se* ***normal*** *à superfície* $F(x, y, z) = 0$ *no ponto ordinário* $P(x_0, y_0, z_0)$ *à recta perpendicular ao plano tangente à superfície nesse ponto.*

Como a normal tem a mesma orientação de (grad $F)_P$, a sua equação cartesiana é

$$\frac{x - x_0}{\left(\dfrac{\partial F}{\partial x}\right)_P} = \frac{y - y_0}{\left(\dfrac{\partial F}{\partial y}\right)_P} = \frac{z - z_0}{\left(\dfrac{\partial F}{\partial z}\right)_P}.$$

Nota. Caso a equação cartesiana da superfície seja dada na forma explícita $z = f(x, y)$, pondo $F(x, y, z) = f(x, y) - z = 0$ resultam as equações

$$z - z_0 = \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) \quad \text{(plano tangente)}$$

$$\frac{x - x_0}{\dfrac{\partial f}{\partial x}} = \frac{y - y_0}{\dfrac{\partial f}{\partial y}} = \frac{z - z_0}{-1} \quad \text{(plano normal)}$$

em que as derivadas se supõem calculadas no ponto (x_0, y_0).

Exemplo 9.10. Determinar as equações do plano tangente e da normal à superfície $x^3 + y^3 + z^3 + xyz - 6 = 0$ no ponto $P(1, 2 - 1)$.

Resolução. Pondo $F(x, y, z) = x^3 + y^3 + z^3 + xyz - 6$ tem-se:

$$\frac{\partial F}{\partial x} = 3x^2 + yz, \quad \frac{\partial F}{\partial y} = 3y^2 + xz, \quad \frac{\partial F}{\partial z} = 3z^2 + xy$$

e no ponto $P(1, 2, -1)$

$$\left(\frac{\partial F}{\partial x}\right)_P = 1, \quad \left(\frac{\partial F}{\partial y}\right)_P = 11, \quad \left(\frac{\partial F}{\partial z}\right)_P = 5.$$

A equação do plano tangente é pois

$$(x-1) + 11(y-2) + 5(z+1) = 0$$

ou seja

$$x + 11y + 5z - 18 = 0.$$

A equação da normal, por sua vez, é

$$\frac{x-1}{1} = \frac{y-2}{11} = \frac{z+1}{5}. \quad \blacktriangleleft$$

Exemplo 9.11. Determinar as equações do plano tangente e da normal à superfície de equação $z = xy$ no ponto $P(1,1,1)$.

Resolução. Pondo $z = f(x,y) = xy$ vem

$$\frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x$$

e no ponto $P(1,1)$: $\left(\dfrac{\partial f}{\partial x}\right)_{(1,1)} = 1 \quad \text{e} \quad \left(\dfrac{\partial f}{\partial y}\right)_{(1,1)} = 1.$

A equação do plano tangente é então

$$z - 1 = (x-1) + (y-1) \quad \text{ou seja} \quad x + y - z - 1 = 0.$$

A equação da normal é: $\dfrac{x-1}{1} = \dfrac{y-1}{1} = \dfrac{z-1}{-1}. \quad \blacktriangleleft$

As superfícies até agora estudadas foram superfícies de $\mathbb{R}^3$, isto é, transformações contínuas de $\mathbb{R}^2$ em $\mathbb{R}^3$. De uma forma mais geral, podem considerar-se superfícies em $\mathbb{R}^n$, as quais se definem como transformações contínuas de $\mathbb{R}^{n-1}$ em $\mathbb{R}^n$ e se denominam, frequentemente, **hipersuperfícies**.

9.5 Superfícies quádricas

As figuras tridimensionais que correspondem às cónicas são chamadas **superfícies quádricas**, que são definidas por equações do 2º grau a três variáveis

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + K = 0.$$

Por rotação ou translação dos eixos (ou ambas as coisas) a equação geral pode converter-se em um dos seguintes tipos:

$$Ax^2 + By^2 + Cz^2 = D \quad \text{ou} \quad Ax^2 + By^2 + Iz = 0,$$

que se chamam equações reduzidas da quádrica.

As quádricas mais importantes que consideraremos são as seguintes.

9.5.1 Esfera

Se a esfera tem centro na origem e raio a (Figura 9.6) sua equação é

$$x^2 + y^2 + z^2 = a^2.$$

Se o centro é ponto (h, k, j) a equação é

$$(x - h)^2 + (y - k)^2 + (z - j)^2 = a^2.$$

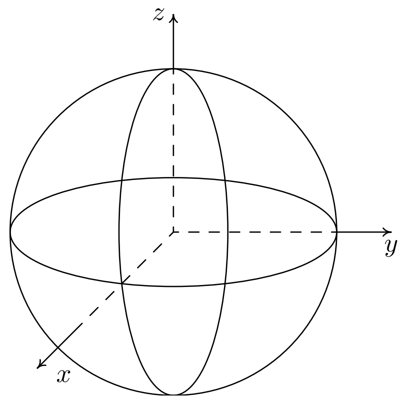

Figura 9.6:

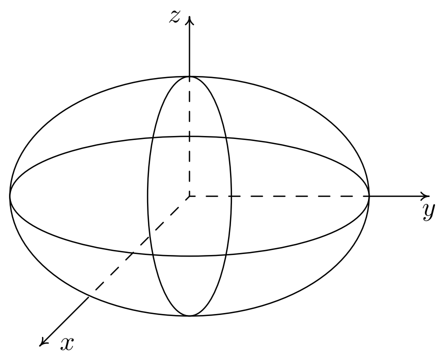

Figura 9.7:

9.5.2 Elipsóide

A equação de elipsóide é:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

A superfície assim definida tem centro na origem e a, b, c são os seus semi-eixos (Figura 9.7).

Se $a = b$ o elipsóide é um elipsóide de revolução em torno de OZ; se $a = c$ em torno de OY; se $b = c$ em torno de OX.

Se o centro for o ponto (h, k, j) a equação toma a forma

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} + \frac{(z-j)^2}{c^2} = 1.$$

9.5.3 Parabolóide elíptico

A equação do parabolóide elíptico é:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z.$$

As secções por planos $z = k$ são elipses; as secções por planos paralelos aos outros eixos coordenados são parábolas (Figura 9.8).

Se $a = b$ a superfície é uma superfície de revolução e as secções horizontais são circunferências e o parabolóide diz-se parabolóide de revolução.

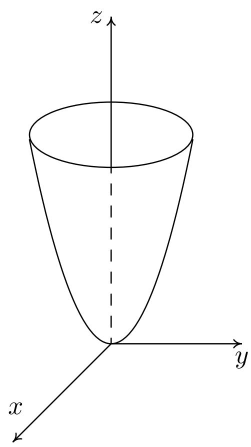

Figura 9.8:

9.5.4 Hiperbolóide de uma folha

A equação da hiperbolóide de uma folha (Figura 9.9) é:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

As secções por planos $z = k$ são elipses (excepto se $a = b$, caso em que o hiperbolóide é de revolução); as secções por planos $x = k$ ou $y = k$ são hipérboles.

9.5.5 Hiperbolóide de duas folhas

A equação de uma hiperbolóide de duas folhas (Figura 9.10) é:

$$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

As secções por planos $x = k$ ou $y = k$ são hipérboles; por planos $z = k$ são elipses.

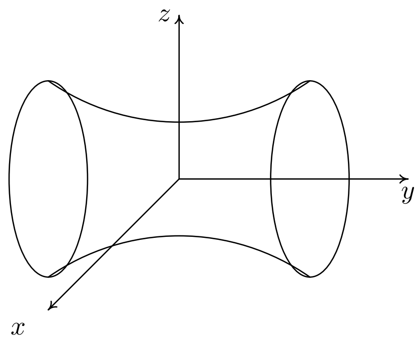

Figura 9.9:

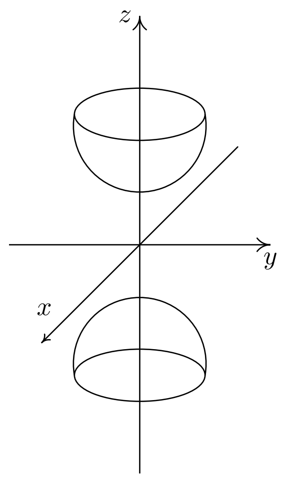

Figura 9.10:

9.5.6 Cone elíptico

A equação de um cone elíptico (Figura 9.11) é:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

As secções por planos $z = k > 0$ são elipses; as secções por planos $x = k$ ou $y = k$ são hipérboles. Se $a = b$ o cone é um cone circular.

9.5.7 Superfície cilíndrica

A equação de uma superfície cilíndrica (Figura 9.12) é:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Se $a = b$ é um cilindro circular; se $a \neq b$ elíptico.

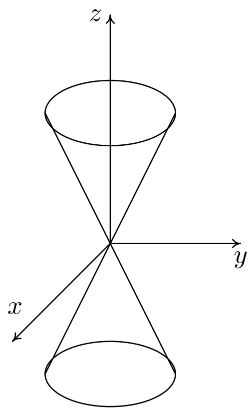

Figura 9.11:

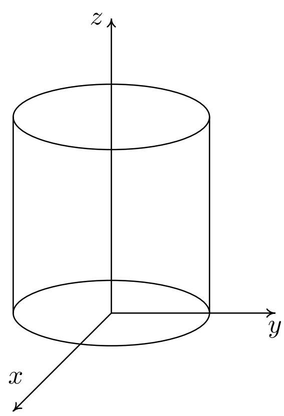

Figura 9.12:

9.6 Exercícios

1. Determine a equação cartesiana de cada uma das seguintes curvas planas:

   (a) $\overrightarrow{r}(t) = t^2\overrightarrow{i} + 2t\overrightarrow{j}, \quad 0 \leq t \leq 2;$

   (b) $\overrightarrow{r}(t) = \cos t\,\overrightarrow{i} + \operatorname{sen}^2 t\,\overrightarrow{j}, \quad 0 \leq t \leq 2\pi;$

   (c) $\overrightarrow{r}(t) = \operatorname{sen} t\,\overrightarrow{i} + (1 - 2\cos t)\overrightarrow{j}, \quad 0 \leq t \leq 3\pi.$

2. Determine as equações da tangente e do plano normal às curvas seguintes nos pontos indicados:

   (a) $x = 6 \operatorname{sen} t$, $y = 4\cos 3t$, $z = 2 \operatorname{sen} 5t$ para $t = \dfrac{\pi}{4}$;

   (b) $x + y + z = 3$, $x^2 - y^2 + 2z^2 = 2$ em $(1, 1, 1)$;

   (c) $3x^2y + y^2z = -2$, $2xz - x^2y = 3$ em $(1, -1, 1)$.

3. Determine as equações da tangente e do plano normal à curva

$$\overrightarrow{r}(t) = e^{2t}\overrightarrow{i} + e^{-2t}\overrightarrow{j} + 3t\overrightarrow{k}, \quad (t = 0).$$

4. Determine as equações do plano tangente e da normal a:

   (a) $3x^4 - 4y^3z + 4z^2xy - 4z^3x + 1 = 0$ em $(1, 1, 1)$;

   (b) $z = 2x^2 - 4y^2$ em $(2, 1, 4)$.

5. Determine o comprimento do arco da curva $x = t - \operatorname{sen} t$, $y = 1 - \cos t$, $z = 4 \operatorname{sen}\left(\dfrac{t}{2}\right)$ compreendido entre $t = 0$ e $t = 2\pi$.

6. Dada a curva $x = \cos t + \operatorname{sen}^2 t$, $y = \operatorname{sen} t - \operatorname{sen} t \cos t$, $z = -\cos t$, determine os vectores $\overrightarrow{t}$, $\overrightarrow{n}$ e $\overrightarrow{b}$ no ponto $t = \dfrac{\pi}{2}$.

7. Determine a curvatura e a torção da curva

$$\overrightarrow{r} = e^t \overrightarrow{i} + e^{-t} \overrightarrow{j} + t\sqrt{2}\,\overrightarrow{k}.$$

8. Determine o raio de curvatura da curva definida por

$$x^2 + y^2 + z^2 = 4, \quad x + y - z = 0.$$

9. Determine a equação do plano osculador à curva $x = y^2$, $z = x^2$ no ponto $(1, 1, 1)$.

10. Identifique as superfície de equações:

    (a) $x^2 + y^2 + z^2 - 2x + 4y + 4 = 0;$

    (b) $x^2 + y^2 - 2y - z + 1 = 0;$

    (c) $z^2 - x^2 - 3y^2 = 0;$

    (d) $2x^2 + 3y^2 - z^2 - 12x + 12y + 2z + 29 = 0;$

    (e) $x^2 + 2y^2 + 2z^2 + 4z + 1 = 0.$

Respostas dos exercícios

Exercícios 1.5 Não aplicável.

Exercícios 2.3 Não aplicável.

Exercícios 3.3 Não aplicável.

Exercícios 4.6

1. (a)$-1 \leq x \leq 1$ e $y \neq 0$; (b) $|x+y| \leq 1$; (c) $xy > 0$.

2. (a) 5; (b) 0; (c) Não existe.

3. (a) Pontos das rectas $x = 0$ e $y = 0$; (b) Pontos da circunferência $x^2 + y^2 = 1$; (c) Pontos dos planos coordenados $x = 0$, $y = 0$, $z = 0$.

Exercícios 5.11

1. (a) $\dfrac{\partial z}{\partial x} = yx^{y-1}$, $\dfrac{\partial z}{\partial y} = x^y \ln x$;
(b) $\dfrac{\partial z}{\partial x} = \operatorname{sen} xy + xy \cos xy$, $\dfrac{\partial z}{\partial y} = x^2 \cos xy$.

2. (a) $\dfrac{\partial^2 z}{\partial x^2} = \dfrac{2y^2}{(x-y)^3}$, $\dfrac{\partial^2 z}{\partial y^2} = \dfrac{2x^2}{(x-y)^3}$, $\dfrac{\partial^2 z}{\partial x \partial y} = \dfrac{\partial^2 z}{\partial y \partial x} =$
$= -\dfrac{2xy}{(x-y)^3}$;

(b) $\dfrac{\partial^2 z}{\partial x^2} = 0,\ \dfrac{\partial^2 z}{\partial y^2} = xe^y,\quad \dfrac{\partial^2 z}{\partial x \partial y} = \dfrac{\partial^2 z}{\partial y \partial x} = e^y.$

4. (a) $dz = \dfrac{1}{1+x^2}dx + \dfrac{1}{1+y^2}dy$; (b) $du = \dfrac{x}{u}dx + \dfrac{y}{u}dy + \dfrac{z}{u}dz.$

5. (a) $d^2z = 2dxdy$; (b) $d^2z = 2\ \text{sen}\ (2y)dxdy + 2x\cos(2y)dy^2.$

7. (a) $\dfrac{dz}{dx} = \cos(x - 6x^2)e^{\text{ sen } x - 2x^3}$; (b) $\dfrac{\partial z}{\partial u} = 0,\ \dfrac{\partial z}{\partial v} = 1$;
(c) $\dfrac{du}{dt} = \dfrac{t}{\sqrt{y}}\ \text{cotg}\ \dfrac{x}{\sqrt{y}} \cdot \left(6 - \dfrac{x}{2y^2}\right).$

6. $(adx + bdy)^n e^{ax+by}.$

9. (a) $2xyz^3\overrightarrow{i} + x^2z^3\overrightarrow{j} + 3x^2yz^2\overrightarrow{k}$; (b) $z - 2y$;
(c) $2x^2\overrightarrow{i} + (x - 4xy)\overrightarrow{j}$; (d) $3x^2yz^4 - 3x^2y^2z^3 + 6x^4y^2z^2.$

11. (a) $-\sqrt{22}$; (b) $\dfrac{1}{2}$; (c) $-\dfrac{4}{\sqrt{5}}$; (d) 1.

Exercícios 6.3

1. (a) $f(x,y) = 1 - \dfrac{1}{2}(x^2+y^2) + \dfrac{1}{24}(x^4 + 6x^2y^2 + y^4) + \ldots$;
(b) $f(x,y) = \displaystyle\sum_{n=0}^{14}(-1)^n(x+y)^n$;
(c) $f(x,y) = 1 + (x-1)y - \dfrac{1}{2}(x-1)^2y + R_3.$

2. (a) $f(x,y) = 1 - \dfrac{x^2}{2} - \dfrac{y^2}{2} + R_2$; (b) $f(x,y) = y + xy + R_2.$

3. $f(x,y) = xy\left(2 - \dfrac{x^2+y^2}{3!} + \dfrac{x^4+y^4}{5!} - \dfrac{x^6+y^6}{7!} + \ldots\right).$

4. $xy - \dfrac{1}{2}x^2y^2 + \dfrac{1}{3}x^3y^3 - \dfrac{1}{4}x^4y^4 + \ldots.$

5. $e^x\cos y + e^x(h\cos y - k\ \ \text{sen}\ y) + \dfrac{e^x}{2!}(h^2\cos y - 2hk\ \ \text{sen}\ y - k^2\ \ \text{sen}\ y) + \ldots.$

Exercícios 7.3

1. (a) $z_{\text{min}} = 0$ em $(1, \frac{1}{2})$; (b) $z_{\text{max}} = -3$ em $(-1, -1)$;
(c) Não tem extremos; (d) $z_{\text{min}} = -2e^{-1}$ em $(-2, 0)$;
(e) $z_{\text{max}} = a^2b^2$ em (a, b).

2. (a) $z_{\text{min}} = \dfrac{1}{2}$; (b) $z_{\text{max}} = \dfrac{e^{a^2}}{4}$; (c) $z_{\text{max}} = \dfrac{1}{2}$, $z_{\text{min}} = -\dfrac{1}{2}$;
(d) $z_{\text{min}} = -\dfrac{605}{32}$; (e) $z_{\text{max}} = \sqrt{3}$, $z_{\text{min}} = -\sqrt{3}$.

3. $r = 1$, $h = 2$.

4. Cubo.

5. $\sqrt{20}$.

6. Equilátero.

7. Equilátero.

8. $\dfrac{2\sqrt{3}}{3}$.

Exercícios 8.6

1. $-\dfrac{2xy}{(y^2 - x^2)^3}$.

2. $\dfrac{2xu^2 + 3y^3}{3x^2uv^2 + y}$.

3. $\dfrac{\partial u}{\partial x} = \dfrac{1 - 12v}{1 - 8uv}, \quad \dfrac{\partial v}{\partial y} = -\dfrac{1 + 4u}{1 - 8uv}$.

4. $h^2 - g - 2f = 0$.

5. $u = \frac{1}{2}(x + 2y)$, $v = \frac{1}{2}(y - 2x)$

6. (a) $2(v - u) + w - 8 = 0$; (b) $v = \cos e^u$.

7. (a) Máx para $x = -1$; (b) Não há.

8. Máximo em $(0, 0)$.

Exercícios 9.6

1. (a) $y^2 - 4x = 0$, $0 \leq y \leq 4$; (b) $x^2 + y - 1 = 0$, $-1 \leq x \leq 1$; (c) $2x^2 - y = 0$, $-1 \leq x \leq 1$.

2. (a) $\dfrac{x - 3\sqrt{2}}{3} = \dfrac{y + 2\sqrt{2}}{-6} = \dfrac{z + \sqrt{2}}{-5}$, $3x - 6y - 5z = 26\sqrt{2}$; (b) $\dfrac{x-1}{3} = \dfrac{y-1}{-1} = \dfrac{z-1}{-2}$, $3x - y - 2z = 0$; (c) $\dfrac{x-1}{3} = \dfrac{y+1}{16} = \dfrac{z-1}{2}$, $3x + 16y + 2z = -11$.

3. $\dfrac{x-1}{2} = \dfrac{y-1}{-2} = \dfrac{z}{3}$, $2x - 2y + 3z = 0$.

4. (a) $3x - 2y - 2z + 1 = 0$, $\dfrac{x-1}{3} = \dfrac{y-1}{-2} = \dfrac{z-1}{-2}$; (b) $8x - 8y - z - 4 = 0$, $\dfrac{x-2}{8} = \dfrac{y-1}{-8} = \dfrac{z-4}{-1}$.

5. 4π.

6. $\overrightarrow{t} = \dfrac{-\overrightarrow{i} + \overrightarrow{j} + \overrightarrow{k}}{\sqrt{3}}$, $\overrightarrow{n} = \dfrac{-5\overrightarrow{i} - 4\overrightarrow{j} - \overrightarrow{k}}{\sqrt{42}}$, $\overrightarrow{b} = \dfrac{\overrightarrow{i} - 2\overrightarrow{j} + 3\overrightarrow{k}}{\sqrt{14}}$.

7. Curvatura $= \dfrac{\sqrt{2}}{(x+y)^2}$, Torção $= \dfrac{-\sqrt{2}}{(x+y)^2}$.

8. 2.

9. $6x - 8y - z + 3 = 0$.

10. (a) Esfera com centro em $(1, -2, 0)$ e raio 1; (b) Parabolóide elíptico; (c) Cone elíptico; (d) Hiperbolóide de uma folha; (e) Elipsóide com centro em $(0, 0, 1)$.

Bibliografia

[1] AGUDO, J. DIAS, *Análise Real*, Livraria Escolar Editora, Lisboa, 1969.

[2] APOSTAL, TOM, *Mathematical Analusis*, Addison–Wesley Publishing Co, Reading, Mass., 1963.

[3] DIEUDONNÉ, J., *Foundations of Modern Analysis*, Academic Press, New York, 1960.

[4] FRIIEDMAN, AVNER, *Foundations of Modern Analysis*, Hlat, Rinehart and Winston, Inc., New York, 1970.

[5] FULKS, WATSON, *Advanced Calculus*, John Wiley and Sons, New York, 1969.

[6] KAPLAN, WILFRED, *Advanced Calculus*, Addison–Wesley Publishing Co, Reading, Mass., 1972.

[7] LIMA, F, *Análise Real*, IMPA, 1989.

[8] PASTOR, J. REY, *Análisis Matemático*, 8ª ed., Kapeluz, Buenos Aires, 1969.

[9] RANDOLPH, JOHN F., *Basic Real And Abstract Analysis*, Academic Press, New York, 1968.

[10] ROSENLICHT, MAXWELL, *Introduction to Analysis*, Scott, Fresman and Co, Glenview, Illinois, 1968.

[11] VALIRON, GEORGES, *Théorie des Fonctions*, Masson, Paris, 1966.

Índice Remissivo