Cálculo em $\mathbb{R}^m$

Pedro Lopes
Departamento de Matemática
Instituto Superior Técnico
1o. Semestre 2006/2007

Estas notas constituem um material de apoio ao curso de Análise Matemática II para as licenciaturas de Engenharia Aeroespacial, Engenharia Mecânica e Engenharia e Arquitectura Naval do Instituto Superior Técnico no 1o. semestre de 2006/2007 e não pretendem ser um substituto dos manuais escolares disponíveis.

1 Introdução. Estrutura Algébrica e Topológica em $\mathbb{R}^m$

Iniciamos aqui o estudo de funções de várias variáveis. Tipicamente as nossas funções serão dadas por expressões como:

$$f(x,y) = \frac{1}{x-y}, \qquad g(x,y,z) = x\log|y-z|, \qquad h(x_1,x_2,x_3,x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

Os domínios destas funções são:

$$\begin{aligned} D_f &= \{(x,y) \mid x,y \in \mathbb{R} \text{ e } x-y \neq 0\} \\ D_g &= \{(x,y,z) \mid x,y,z \in \mathbb{R} \text{ e } y-z \neq 0\} \\ D_h &= \{(x_1,x_2,x_3,x_4) \mid x_1,x_2,x_3,x_4 \in \mathbb{R}\} \end{aligned}$$

que são subconjuntos de $\mathbb{R}^2$ (ver figura 1), $\mathbb{R}^3$ (ver figura 2) e $\mathbb{R}^4$ (de facto, $D_h = \mathbb{R}^4$).

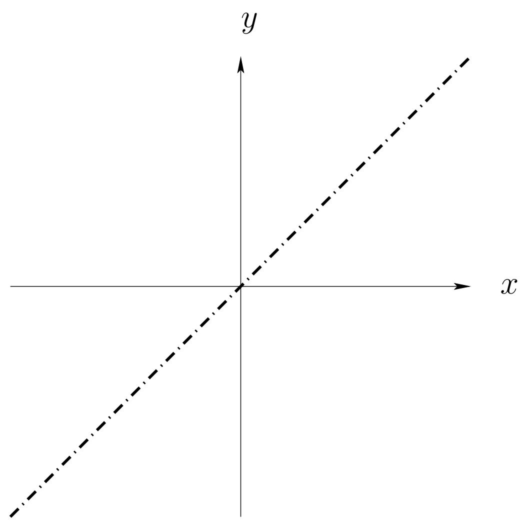

Figure 1: D_f é o plano XY excepto a linha representada na figura

Em $\mathbb{R}$ era comum usarmos operações algébricas como adição, subtracção, multiplicação e divisão para, entre outras coisas, definirmos as expressões analíticas das nossas funções. Quais destas operações fazem ainda sentido em $\mathbb{R}^m$ com $m > 1$?

Adição em $\mathbb{R}^m$

Sejam $(x_1,x_2,\ldots,x_m)$ e $(y_1,y_2,\ldots,y_m)$ dois elementos genéricos de $\mathbb{R}^m$. Definimos adição destes dois elementos:

$$(x_1,x_2,\ldots,x_m) + (y_1,y_2,\ldots,y_m) = (x_1+y_1, x_2+y_2, \ldots, x_m+y_m)$$

Esta operação é comutativa:

$$(x_1,x_2,\ldots,x_m) + (y_1,y_2,\ldots,y_m) = (x_1+y_1, x_2+y_2, \ldots, x_m+y_m) = (y_1+x_1, y_2+x_2, \ldots, y_m+x_m)$$

O elemento de $\mathbb{R}^m$ com todas as coordenadas nulas, $(0,0,\ldots,0)$, é tal que

$$(0,0,\ldots,0) + (x_1,x_2,\ldots,x_m) = (x_1,x_2,\ldots,x_m) = (x_1,x_2,\ldots,x_m) + (0,0,\ldots,0)$$

ou seja, $(0,0,\ldots,0)$ é o elemento neutro da adição em $\mathbb{R}^m$. Para cada $(x_1,x_2,\ldots,x_m)$, existe um único $(-x_1,-x_2,\ldots,-x_m)$ para os quais se tem:

$$(x_1,x_2,\ldots,x_m) + (-x_1,-x_2,\ldots,-x_m) = (0,0,\ldots,0) = (-x_1,-x_2,\ldots,-x_m) + (x_1,x_2,\ldots,x_m)$$

ou seja cada elemento de $\mathbb{R}^m$ tem um inverso em $\mathbb{R}^m$. Finalmente, esta operação é associativa tendo portanto, todas as propriedades que já conhecíamos da adição em $\mathbb{R}$.

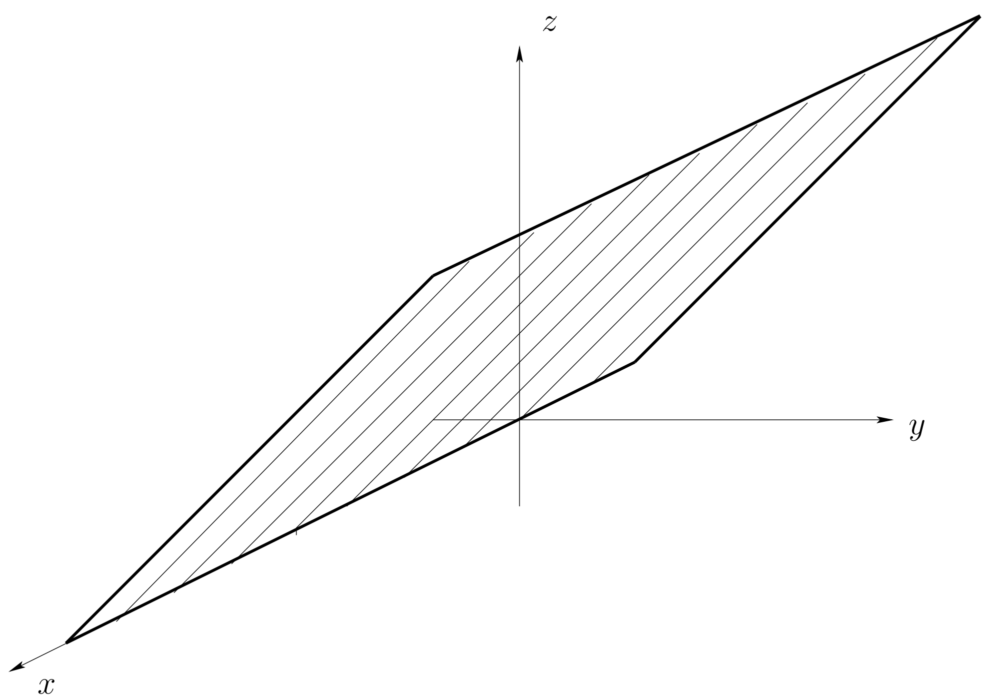

Figure 2: D_g é $\mathbb{R}^3$ excepto o plano representado na figura

Não existe noção de multiplicação em $\mathbb{R}^m$ tal como a conhecíamos em $\mathbb{R}$. Que outras maneiras de associar dois elementos para produzir novos elementos de $\mathbb{R}^m$ temos ainda?

Multiplicação por escalar

Dado um número real α (dito "escalar") e um elemento de $\mathbb{R}^m$, $(x_1, x_2, \ldots, x_m)$, definimos multiplicação por escalar da seguinte maneira:

$$\alpha(x_1, x_2, \ldots, x_m) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_m)$$

Distância entre dois elementos

Em $\mathbb{R}$, a distância entre dois elementos, x e y, é dada pelo módulo da diferença entre os dois, $|x - y|$. Em $\mathbb{R}^2$, usamos o Teorema de Pitágoras:

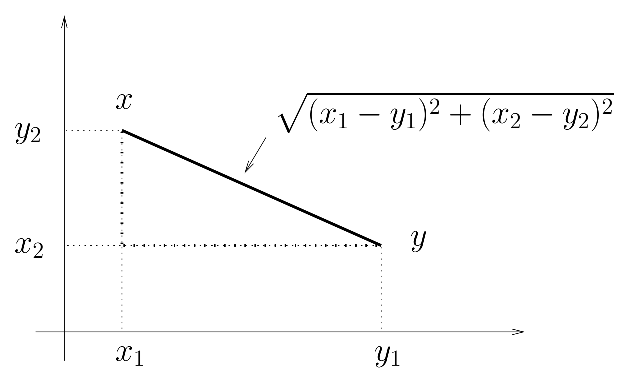

Figure 3: Distância entre dois pontos em $\mathbb{R}^2$, dadas as suas coordenadas

Em $\mathbb{R}^3$:

O que indica que, de um modo geral, $d(x, y)$, a distância entre os elementos $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$, em $\mathbb{R}^m$, é dada por:

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

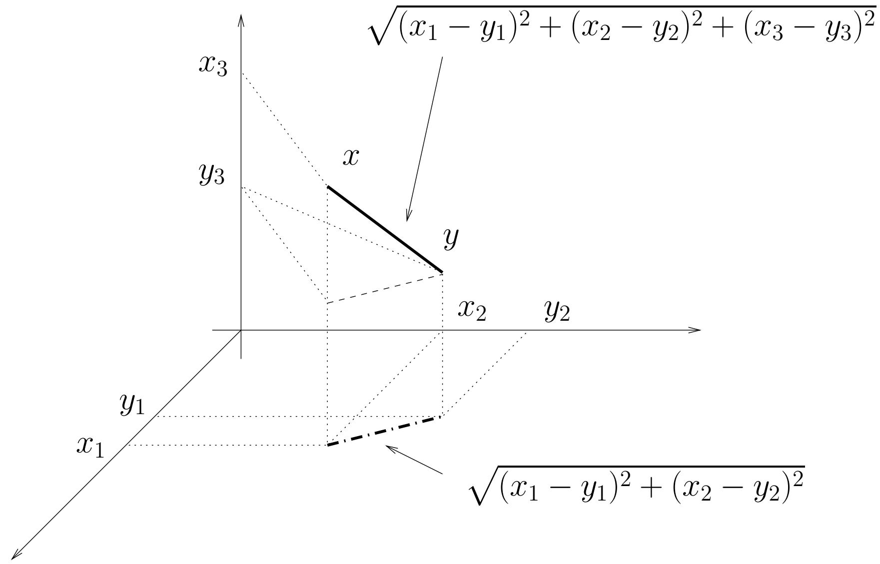

Figure 4: Distância entre dois pontos em $\mathbb{R}^3$, dadas as suas coordenadas

Se fizermos $(y_1, y_2, \ldots, y_m) = (0, 0, \ldots, 0)$, obtemos

$$\begin{aligned}\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_m-y_m)^2} &= \sqrt{(x_1-0)^2+(x_2-0)^2+\cdots+(x_m-0)^2} = \\ &= \sqrt{x_1^2+x_2^2+\cdots+x_m^2}\end{aligned}$$

que dá a distância do elemento $(0, 0, \ldots, 0)$ até ao elemento $(x_1, x_2, \ldots, x_m)$. A este valor chamamos norma de $(x_1, x_2, \ldots, x_m)$ e denotamos por $||(x_1, x_2, \ldots, x_m)||$:

$$||(x_1, x_2, \ldots, x_m)|| = \sqrt{x_1^2+x_2^2+\cdots+x_m^2}$$

Note-se que se fizermos $m = 1$ obtemos

$$||x_1|| = \sqrt{x_1^2} = |x_1|$$

Então a norma, $||\cdots||$, generaliza a noção de módulo. Por outro lado,

$$\begin{aligned}d(x,y) &= \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_m-y_m)^2} = ||(x_1-y_1, x_2-y_2, \cdots, x_m-y_m)|| = \\ &= ||(x_1, x_2, \cdots, x_m) - (y_1, -y_2, \cdots - y_m)|| = ||x-y||\end{aligned}$$

(com $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$) mais uma vez generalizando para $\mathbb{R}^m$ um facto nosso conhecido de $\mathbb{R}$: que a distância entre dois reais é o módulo da sua diferença.

Definimos tambem, para $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$ em $\mathbb{R}^m$, o seu produto interno:

$$x \cdot y = (x_1, x_2, \ldots, x_m) \cdot (y_1, y_2, \ldots, y_m) = x_1y_1 + x_2y_2 + \ldots x_my_m$$

Se, em particular, fizermos $y = x$, isto é, $(y_1, y_2, \ldots, y_m) = (x_1, x_2, \ldots, x_m)$, vem

$$x \cdot x = (x_1, x_2, \ldots, x_m) \cdot (x_1, x_2, \ldots, x_m) = x_1x_1 + x_2x_2 + \ldots x_mx_m = x_1^2 + x_2^2 + \ldots x_m^2 = ||x||^2$$

ou seja, o produto interno de um elemento por ele próprio é igual ao quadrado da sua norma. Então, dados dois elementos quaisquer de $\mathbb{R}^m$, $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$ e um número real (escalar) α, tem-se

$$\begin{aligned}
0 \leq ||x+\alpha y||^2 &= (x+\alpha y)\cdot(x+\alpha y) = \\
&= \big((x_1,x_2,\dots x_m)+\alpha(y_1,y_2,\dots y_m)\big)\cdot\big((x_1,x_2,\dots x_m)+\alpha(y_1,y_2,\dots y_m)\big) = \\
&= \big((x_1+\alpha y_1,x_2+\alpha y_2,\dots,x_m+\alpha y_m)\big)\cdot\big((x_1+\alpha y_1,x_2+\alpha y_2,\dots,x_m+\alpha y_m)\big) = \\
&= \big(x_1+\alpha y_1\big)\big(x_1+\alpha y_1\big)+\big(x_2+\alpha y_2\big)\big(x_2+\alpha y_2\big)+\dots+\big(x_m+\alpha y_m\big)\big(x_m+\alpha y_m\big) = \\
&= x_1\big(x_1+\alpha y_1\big)+\alpha y_1\big(x_1+\alpha y_1\big)+\dots+x_m\big(x_m+\alpha y_m\big)+\alpha y_m\big(x_m+\alpha y_m\big) = \\
&= \big(x_1^2+\alpha x_1y_1\big)+\big(\alpha y_1x_1+\alpha^2y_1^2\big)+\dots+\big(x_m^2+\alpha x_my_m\big)+\big(\alpha y_mx_m+\alpha^2y_m^2\big) = \\
&= \big(x_1^2+\dots+x_m^2\big)+2\alpha\big(x_1y_1+\dots+x_my_m\big)+\alpha^2\big(y_1^2+\dots+y_m^2\big) = \\
&= ||x||^2+2\alpha x\cdot y+\alpha^2||y||^2
\end{aligned}$$

ou seja, concluímos que o polinómio de grau dois em α:

$$p(\alpha) = \big(||y||^2\big)\alpha^2 + \big(2x\cdot y\big)\alpha + ||x||^2$$

é maior ou igual a zero. As raízes de um tal polinómio são dadas por:

$$\alpha = \frac{-2x\cdot y \pm \sqrt{(2x\cdot y)^2 - 4||y||^2||x||^2}}{2||y||^2}$$

Como o nosso polinómio tem, no máximo, uma raiz (porque é sempre maior ou igual a zero) então a expressão dentro da raiz tem de ser **menor ou igual a zero**, ou seja

$$(2x\cdot y)^2 - 4||y||^2||x||^2 \leq 0 \Leftrightarrow (x\cdot y)^2 \leq ||y||^2||x||^2 \Leftrightarrow |x\cdot y| \leq ||x||\,||y||$$

Então, para qualquer x e y em $\mathbb{R}^m$,

$$|x\cdot y| \leq ||x||\,||y||$$

que é conhecido por **desigualdade de Cauchy-Schwartz**.

Exercício 1.1

Para quaisquer x, y e z em $\mathbb{R}^m$ e α em $\mathbb{R}$, estabelecer as seguintes propriedades:

$$\begin{aligned}
&x\cdot y = y\cdot x \\
&(x+y)\cdot z = x\cdot z + y\cdot z \quad ; \qquad x\cdot(y+z) = x\cdot y + x\cdot z \\
&||\alpha x|| = |\alpha|\,||x|| \\
&(\alpha x)\cdot y = \alpha(x\cdot y) = x\cdot(\alpha y)
\end{aligned}$$

Da desigualdade de Cauchy-Schwartz resulta que

$$\begin{aligned}
||x+y||^2 &= (x+y)\cdot(x+y) = x\cdot(x+y) + y\cdot(x+y) = x\cdot x + x\cdot y + y\cdot x + y\cdot y = \\
&= ||x||^2 + 2x\cdot y + ||y||^2 \leq ||x||^2 + 2||x||\,||y|| + ||y||^2 = \big(||x|| + ||y||\big)^2
\end{aligned}$$

e portanto:

$$||x+y|| \leq ||x|| + ||y||$$

Desta desigualdade resulta que dados x, y e z em $\mathbb{R}^m$, se tem,

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

isto é

$$||x-y|| \leq ||x-z|| + ||z-y||$$

conhecida por **desigualdade triangular**. Verifica-se ainda trivialmente que, para qualquer x e y em $\mathbb{R}^m$ se tem

$$||x-y|| = ||y-x||$$

e

$$||x-y|| \geq 0$$

Estas três propriedades,

$$(1) \quad ||x-y|| \leq ||x-z|| + ||z-y||, \qquad (2) \quad ||x-y|| = ||y-x||, \qquad \text{e} \quad (3) \quad ||x-y|| \geq 0$$

são as que se esperam de uma distância entre dois pontos: (1) que não seja negativa; (2) que a distância de um ponto a um segundo ponto seja a mesma que a distância do segundo ponto ao primeiro; e que (3) a distancia de um ponto a um segundo seja menor ou, quando muito, igual que a distância do primeiro ponto a um terceiro ponto mais a distância desse terceiro ponto ao segundo (ir de Lisboa directo ao Porto percorre-se menos distância do que ir primeiro de Lisboa a Elvas e só depois de Elvas ao Porto).

Exercício 1.2

Considere um conjunto qualquer (ao qual chamamos X). Nesse conjunto considere a função:

$$d(x,y) = \begin{cases} 1, & \text{se } x = y \\ 0, & \text{se } x \neq y \end{cases}$$

Verifique que esta é uma função distância em X (isto é que verifica as propriedades (1), (2) e (3) acima).

NOTA: De agora em diante, a função distância que consideramos é:

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

em $\mathbb{R}^m$.

Tendo equipado $\mathbb{R}^m$ com uma função distância podemos agora falar, dado um ponto arbitrário a em $\mathbb{R}^m$, do conjunto de pontos que estão próximos de a a menos de um dado valor ϵ, ou seja de vizinhanças - agora conhecidas por bolas.

Definição 1.1

Dado $a = (a_1, \ldots, a_m)$ em $\mathbb{R}^m$ e $\epsilon > 0$ chamamos bola de centro a e raio ϵ ao conjunto:

$$B_\epsilon(a) = \{x \in \mathbb{R}^m \mid ||x-a|| < \epsilon\}$$

Com estas noções de distâncias e de bolas (vizinhanças) faz agora sentido falar de sucessões limitadas, sucessões convergentes; pontos interiores a conjuntos, pontos aderentes a conjuntos; etc., tal como fazíamos em $\mathbb{R}$.

2 Sucessões

Uma sucessão em $\mathbb{R}^m$ é uma correspondência que a cada número natural n associa um elemento de $\mathbb{R}^m$.

Exemplo 2.1

$$x_n = \underbrace{(n, n, \ldots, n)}_{m \text{ coordenadas}}$$

$$y_n = \underbrace{\left(\frac{1}{n}, n, \frac{1}{n}, n, \ldots\right)}_{m \text{ coordenadas}}$$

$$z_n = \underbrace{\left(\frac{1}{n}, \frac{2}{n}, \ldots, \frac{m}{n}\right)}_{m \text{ coordenadas}}$$

$$w_n = \underbrace{\left((-1)^n, (-1)^n, \ldots, (-1)^n\right)}_{m \text{ coordenadas}}$$

Note-se que aqui o indíce n em x_n, y_n, z_n, w_n, representa o termo da sucessão.

Definição 2.1 (Sucessão Limitada)

A sucessão (x_n) em $\mathbb{R}^m$ diz-se limitada se existir um real positivo R tal que

$$||x_n|| < R, \qquad \text{qualquer que seja } n$$

Exemplo 2.2

(z_n), (w_n) acima são sucessões limitadas, já que

$$||z_n|| = \left|\left|\left(\frac{1}{n}, \frac{2}{n}, \ldots, \frac{m}{n}\right)\right|\right| = \sqrt{\left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \cdots + \left(\frac{m}{n}\right)^2} \leq \sqrt{\left(\frac{m}{n}\right)^2 + \left(\frac{m}{n}\right)^2 + \cdots + \left(\frac{m}{n}\right)^2} =$$
$$= \sqrt{m\left(\frac{m}{n}\right)^2} = \frac{m}{n}\sqrt{m} \leq m\sqrt{m}$$

e

$$||w_n|| = \left|\left|\Big((-1)^n, (-1)^n, \ldots, (-1)^n\Big)\right|\right| = \sqrt{\Big((-1)^n\Big)^2 + \Big((-1)^n\Big)^2 + \cdots + \Big((-1)^n\Big)^2} = \sqrt{m}$$

Definição 2.2 (Sucessão Convergente)

A sucessão (x_n) em $\mathbb{R}^m$ diz-se convergente para a em $\mathbb{R}^m$ se, qualquer que seja $\epsilon > 0$. existir um inteiro positivo N tal que

$$n > N \Longrightarrow x_n \in B_\epsilon(a)$$

Exemplo 2.3

A sucessão $\left(\frac{1}{n} + 1, \frac{1}{n}\right)$ converge para $(1, 0)$. De facto,

$$\left|\left|\left(\frac{1}{n} + 1, \frac{1}{n}\right) - (1,0)\right|\right| = \left|\left|\left(\frac{1}{n}, \frac{1}{n}\right)\right|\right| = \sqrt{\left(\frac{1}{n}\right)^2 + \left(\frac{1}{n}\right)^2} = \sqrt{2\left(\frac{1}{n}\right)^2} = \frac{1}{n}\sqrt{2} = \frac{\sqrt{2}}{n}$$

Então, dado $\epsilon > 0$, tome-se um inteiro positivo N tal que

$$N > \frac{\sqrt{2}}{\epsilon}$$

donde

$$\frac{\sqrt{2}}{N} < \epsilon$$

Então, para todo o $n > N$ tem-se:

$$\left|\left|\left(\frac{1}{n} + 1, \frac{1}{n}\right) - (1,0)\right|\right| = \frac{\sqrt{2}}{n} < \frac{\sqrt{2}}{N} < \epsilon$$

ou seja, dado $\epsilon > 0$, determinámos um inteiro positivo N tal que,

$$n > N \quad \Longrightarrow \quad \left|\left|\left(\frac{1}{n} + 1, \frac{1}{n}\right) - (1,0)\right|\right| < \epsilon$$

Portanto, a sucessão $\left(\frac{1}{n} + 1, \frac{1}{n}\right)$ converge, por definição de convergência de sucessões, para $(1, 0)$.

Às sucessões como $\left(\frac{1}{n} + 1\right)$ e $\left(\frac{1}{n}\right)$ em relação à sucessão $\left(\frac{1}{n} + 1, \frac{1}{n}\right)$, chamamos sucessões coordenadas.

Seguidamente apresentamos alguns resultados sobre sucessões que são análogos a resultados que já conhecemos de sucessões em $\mathbb{R}$.

Proposição 2.1 *Uma sucessão convergente tem um único limite*

Dem. Suponha que a sucessão convergente (x_n) em $\mathbb{R}^m$ tem dois limites distintos, a e b em $\mathbb{R}^m$, com $a \neq b$. Seja $\epsilon = ||a-b||$. Como (x_n) converge, então existe um inteiro positivo N tal que para $n > N$, $x_n \in B_{\frac{\epsilon}{2}}(a)$ e $x_n \in B_{\frac{\epsilon}{2}}(b)$, ou seja, para $n > N$, tem-se, simultâneamente,

$$||x_n - a|| < \frac{\epsilon}{2} \quad \text{e} \quad ||x_n - b|| < \frac{\epsilon}{2}$$

o que implica que

$$||a-b|| = ||\big(a - x_n\big) + \big(x_n - b\big)|| \leq ||\big(a - x_n\big)|| + ||\big(x_n - b\big)|| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

donde em particular,

$$||a-b|| < \epsilon$$

o que contradiz a nossa escolha inicial de

$$\epsilon = ||a-b||$$

o que é absurdo, demonstrando-se assim que o limite de uma sucessão convergente é único. ■

Proposição 2.2 *Uma sucessão convergente é limitada.*

Dem. Suponha que a sucessão (x_n) converge para a. Então, por definição de convergência de sucessões, para todo o $\epsilon > 0$ existe um inteiro positivo N tal que para todo o $n > N$, $||x_n - a|| < \epsilon$. Como vimos nas práticas

$$\Big| ||x|| - ||y|| \Big| < ||x - y||$$

donde dado $\epsilon > 0$ existe um inteiro positivo tal que N

$$\Big| ||x_n|| - ||a|| \Big| < ||x_n - a|| < \epsilon$$

e portanto, desembaraçando de módulos

$$||a|| - \epsilon < ||x_n|| < ||a|| + \epsilon$$

ou seja existe um N tal que o conjunto dos termos da sucessão x_n, **a partir desse índice** N, é limitada. Falta portanto saber se o conjunto dos termos da sucessão até ao índice N é, ou não, um conjunto limitado. Mas este conjunto é um conjunto finito (só tem N elementos) logo é limitado. Finalmente, a união de dois conjuntos limitados (no caso, $\{||x_n|| \mid n \leq N\}$ e $\{||x_n|| \mid n > N\}$) é outra vez um conjunto limitado. Portanto, se uma sucessão é convergente então é uma sucessão limitada. ■

Proposição 2.3 *Uma sucessão é convergente (é limitada, respect.) se e só se todas as suas sucessões coordenadas forem convergentes (limitadas, respect.)*

Dem. Omitida ■

Proposição 2.4 *Se uma sucessão é limitada então tem subsucessões convergentes.*

Dem. Omitida. ■

Os resultados reunidos na proposição 2.5, são, de um modo geral, generalizações para $\mathbb{R}^m$, de resultados já conhecidos em $\mathbb{R}$ e cujas demonstrações são manipulações simples das definições e que deixamos então como exercício, a cargo de quem ler estas notas.

Proposição 2.5 *Se as sucessões (u_n) e (v_n) em $\mathbb{R}^m$ são convergentes então:*

- $(u_n \pm v_n)$ *tambem é convergente e* $\lim_{n\mapsto\infty}(u_n \pm v_n) = \lim_{n\mapsto\infty}(u_n) \pm \lim_{n\mapsto\infty}(v_n)$
- $(u_n \cdot v_n)$ *tambem é convergente e* $\lim_{n\mapsto\infty}(u_n \cdot v_n) = \lim_{n\mapsto\infty}(u_n) \cdot \lim_{n\mapsto\infty}(v_n)$
- $\big(||u_n||\big)$ *tambem é convergente e* $\lim_{n\mapsto\infty} ||u_n|| = || \lim_{n\mapsto\infty} u_n||$

Se, além disso, (α_n) é uma sucessão convergente de números reais (escalares) então

- $(\alpha_n u_n)$ *tambem é convergente e* $\lim_{n\mapsto\infty}(\alpha_n u_n) = \lim_{n\mapsto\infty}(\alpha_n) \lim_{n\mapsto\infty}(u_n)$

3 Noções Topologicas em $\mathbb{R}^m$

A Topologia tem a ver com a maneira como elementos se relacionam com conjuntos (na profundidade com que a estudamos aqui). A importância da topologia já deve ter ficado ilustrada aquando do estudo de funções em $\mathbb{R}$. Por exemplo, sabemos já que uma função contínua num intervalo fechado e limitado é uma função limitada e é tambem uma função integrável. A importância de "fechado" na expressão anterior fica clara quando pensamos na função

$$f(x) = \frac{1}{x} \quad \text{com } x \in]0,1]$$

Esta função, apesar de contínua, **não** é limitada, **nem** é integrável no intervalo (limitado) indicado. Qual é a aparente inconsistência com os resultados que recordámos acima? É que o intervalo $]0,1]$ **não é fechado**.

Tendo recordado a importância da Topologia em $\mathbb{R}$ passemos às definições em $\mathbb{R}^m$. Sejam D um subconjunto de $\mathbb{R}^m$ e a um elemento de $\mathbb{R}^m$.

Definição 3.1 (Elemento interior a um conjunto)

a é elemento interior a D se existe uma bola centrada em a toda contida em D, isto é, se existir $\epsilon > 0$ tal que

$$B_\epsilon(a) \subset D$$

como no exemplo (em $\mathbb{R}^2$) na figura 5. O conjunto dos elementos interiores a um conjunto D chama-se

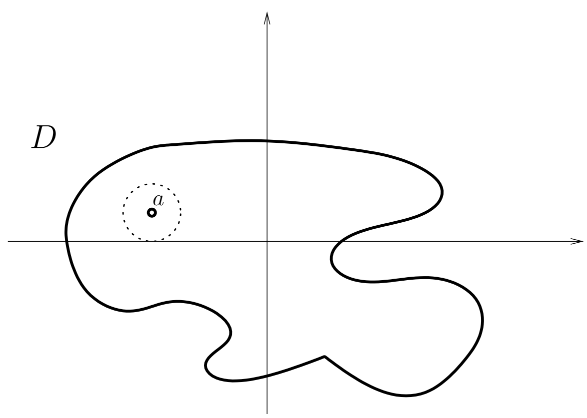

Figure 5: a é elemento interior a D (em $\mathbb{R}^2$)

"interior de D" e denota-se $int\ D$ ou $\overset{\circ}{D}$.

Definição 3.2 (Elemento aderente a um conjunto)

a é ponto aderente a D se a intersecção de D com qualquer bola centrada em a for não vazia isto é, se para qualquer $\epsilon > 0$,

$$B_\epsilon(a) \cap D \neq \emptyset$$

como no exemplo (em $\mathbb{R}^2$) na figura 6. O conjunto dos elementos aderentes a um conjunto D chama-se "aderência de D" ou "fecho de D"e denota-se $\overline{D}$.

Exercício 3.1

Seja X um subconjunto de $\mathbb{R}^m$. Mostre que

$$\overset{\circ}{X} \subset X \subset \overline{X}$$

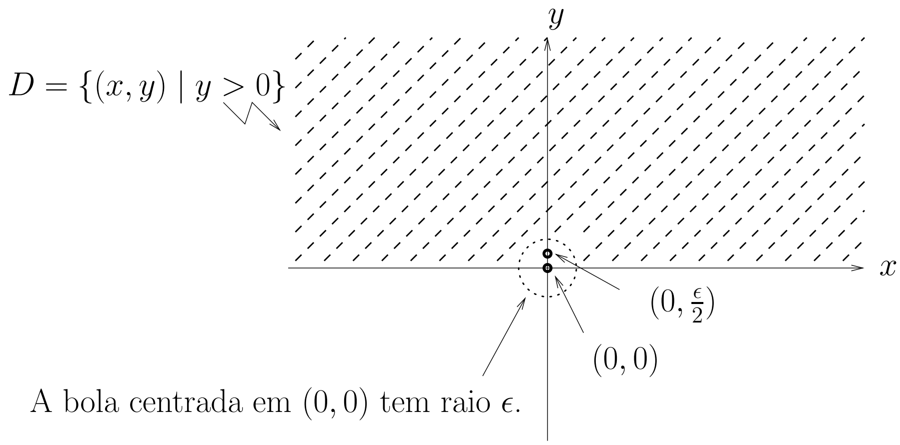

Figure 6: $(0, 0)$ é elemento aderente a D (em $\mathbb{R}^2$)

Definição 3.3 (Conjunto Aberto)

Um conjunto, D, diz-se aberto se cada um dos seu elementos for um elemento interior ao conjunto. Tendo em conta o exercício acima, um conjunto aberto, D, satisfaz,

$$int\ D = D$$

Definição 3.4 (Conjunto Fechado)

Um conjunto, D, diz-se fechado se cada um dos elementos aderentes a D for um elemento de D. Tendo em conta o exercício acima, um conjunto fechado, D, satisfaz,

$$\overline{D} = D$$

Proposição 3.1 *a é aderente a D se, e só se, existir uma sucessão de termos em D que convirja para a.*

Dem. Suponhamos que existe uma sucessão de termos em D, chamemos-lhe (x_n), que converge para a. Então, por definição, para todo o $\epsilon > 0$, existe um inteiro positivo N tal que

$$n > N \Rightarrow x_n \in B_\epsilon(a)$$

e portanto qualquer bola centrada em a contem elementos de D (já que $x_n \in D$ para qualquer n) ou seja, a é ponto aderente a D.

Reciprocamente, suponhamos que a é aderente a D. Então, por definição de elemento aderente, qualquer que seja o $\epsilon > 0$,

$$B_\epsilon(a) \cap D \neq \emptyset$$

Considere-se então a sucessão $\epsilon_n = \frac{1}{n}$. Para cada inteiro positivo n há-de existir um ponto $x_n \in D$ tal que $x_n \in B_{\frac{1}{n}}(a)$, já que a é aderente a D. Obtemos assim uma sucessão de pontos (x_n) em D. Vamos agora ver que essa sucesão tende para a. Como, para cada n, $x_n \in B_{\frac{1}{n}}(a)$ então, para cada n, $||x_n - a|| < \frac{1}{n}$. Então, dado $\delta > 0$, escollha-se um inteiro positivo N tal que $\frac{1}{\epsilon} < N$. Tem-se, então, para $n > N$,

$$||x_n - a|| < \frac{1}{n} < \frac{1}{N} < \epsilon$$

ou seja (x_n) converge para a, terminando a demonstração. ■

Corolário 3.1 *Um conjunto D é fechado se, e só se, toda a sucessão convergente de termos em D convergir para um elemento de D.*

Dem. Omitida. ■

Proposição 3.2 *Uma sucessão limitada tem pelo menos uma subsucessão convergente.*

Dem. Omitida. ■

Corolário 3.2 *Um conjunto D é limitado e fechado se, e só se, toda a sucessão limitada de termos em D tiver uma subsucessão convergente para um elemento de D.*

Dem. Omitida. ■

Definição 3.5 (Elemento exterior a um conjunto)

a diz-se exterior a D se for interior ao complementar de D.

Definição 3.6 (Elemento fronteiro a um conjunto)

a diz-se fronteiro a D se for, simultâneamente, aderente a D e ao complementar de D.

Voltando à importância da Topologia no estudo das funções, relembramos aqui o Teorema do Valor Intermédio em $\mathbb{R}$:

Teorema 3.1 *Seja f contínua num intervalo fechado e limitado $[a, b]$. Nestas condições f assume todos os valores entre $f(a)$ e $f(b)$.*

Em $\mathbb{R}^m$ (para $m > 1$) precisamos trocar "intervalo" por "conjunto conexo" para obter um tal resultado. Damos de seguida a definição de conjunto conexo depois de apresentar a definição de conjuntos separados.

Definição 3.7 (Conjuntos Separados)

Os subconjuntos não vazios de $\mathbb{R}^m$, A e B, dizem-se separados se nenhum deles contiver pontos aderentes ao outro isto é,

$$A \cap \overline{B} = \emptyset \qquad \text{e} \qquad B \cap \overline{A} = \emptyset$$

Exercício 3.2

(i) Se A e B são separados então A e B são disjuntos (isto é, a sua intersecção é vazia, isto é A e B não têm elementos em comum)

(ii) A recíproca não é verdadeira isto é, existem conjuntos disjuntos que não são separados. Encontrar um exemplo (... a figura 6 resolve *parcialmente* esta questão ...).

Dois conjuntos serem separados é então "mais forte" do que serem disjuntos. De facto, para serem separados, dois conjuntos têm que ser disjuntos mas de tal maneira que nenhum deles "toque" a aderência do outro. Um conjunto conexo é então um conjunto que não pode ser escrito como reunião de dois tais conjuntos:

Definição 3.8 (Conjunto não conexo)

Um subconjunto de $\mathbb{R}^m$, X diz-se **não** conexo se existirem dois conjuntos separados A e B tais que

$$X = A \cup B$$

4 Continuidade

A continuidade de uma função em $\mathbb{R}$ era a tradução em linguagem matemática da possibilidade de se poder desenhar o gráfico dessa função sem levantar o lápis do papel. Essa tradução era conseguida à custa das noções de distância/vizinhança:

$$f \text{ é contínua em } a \quad \overset{def.}{\Longleftrightarrow} \quad \text{para todo o } \epsilon > 0, \text{ existe } \delta > 0 \text{ tal que } \Big[\, x \in V_\delta(a) \Longrightarrow f(x) \in V_\epsilon\big(f(a)\big) \,\Big]$$

Como conseguimos "levar para" $\mathbb{R}^m$ esta noção de distância (e consequentemente de vizinhança, ou bola) então a definição de função contínua num ponto será:

Definição 4.1 (Função contínua num ponto)

f é contínua num ponto a do seu domínio se, por definição, para todo o $\epsilon > 0$, existir um $\delta > 0$ tal que

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

e note-se que estamos agora tambem a considerar funções cuja imagem está contida em $\mathbb{R}^p$ com $p > 1$.

Exemplo 4.1

Considere a função f de em $\mathbb{R}^2$ em $\mathbb{R}^3$ tal que, para qualquer $(x_1, x_2) \in \mathbb{R}^2$,

$$f(x_1, x_2) = (1, 3, 2)$$

f é então uma função constante cuja imagem é o elemento $(1, 3, 2)$ de $\mathbb{R}^3$. Seja então (a_1, a_2) um elemento qualquer de $\mathbb{R}^2$; vamos provar que f é contínua nesse ponto. Seja $\epsilon > 0$ e vejamos o que quer dizer para esta função f , $||f(x_1, x_2) - f(a_1, a_2)|| < \epsilon$. Como $f(x_1, x_2) - f(a_1, a_2) = (1, 3, 2) - (1, 3, 2) = 0(> \epsilon)$ - já que a função é constante - então qualquer que seja o $\delta > 0$, $||(x_1, x_2) - (a_1, a_2)|| < \delta$ implica $||f(x_1, x_2) - f(a_1, a_2)|| < \epsilon$. f é portanto contínua em (a_1, a_2)

Definição 4.2

Uma função diz-se contínua num domínio D se for contínua em qualquer ponto desse domínio D.

Voltando ao exemplo acima, como f é contínua num elemento genérico de $\mathbb{R}^2$, então f é contínua em $\mathbb{R}^2$. Note-se tambem que uma argumentação analoga permite mostrar a continuidade de qualquer função constante.

Exemplo 4.2

Considere a função g definida em $\mathbb{R}^m$ e com valores em $\mathbb{R}^m$ dada por

$$g(x_1, \dots, x_m) = (x_1, \dots, x_m) \qquad \text{(função identidade)}$$

Tome-se um elemento qualquer de $\mathbb{R}^m$, $(a_1, \dots, a_m)$; vamos provar que g é contínua nesse ponto (provando assim que é contínua em $\mathbb{R}^m$). Dado $\epsilon > 0$ vejamos o que quer dizer para esta função g , $||g(x_1, \dots, x_m) - g(a_1, \dots, a_m)|| < \epsilon$. Como g é a função identidade,

$$g(x_1, \dots, x_m) - g(a_1, \dots, a_m) = (x_1, \dots, x_m) - (a_1, \dots, a_m)$$

donde

$$||g(x_1, \dots, x_m) - g(a_1, \dots, a_m)|| = ||(x_1, \dots, x_m) - (a_1, \dots, a_m)||$$

Então basta tomar $\delta = \epsilon$:

$$\begin{aligned} &||(x_1, \dots, x_m) - (a_1, \dots, a_m)|| < \delta(= \epsilon) \Longrightarrow \\ &\qquad \Longrightarrow ||g(x_1, \dots, x_m) - g(a_1, \dots, a_m)|| \Big(= ||(x_1, \dots, x_m) - (a_1, \dots, a_m)||\Big) < \epsilon \end{aligned}$$

provando assim que g é contínua em $(a_1, \dots, a_m)$ e portanto em qualquer elemento de $\mathbb{R}^m$, já que $(a_1, \dots, a_m)$ é genérico em $\mathbb{R}^m$.

Exemplo 4.3

Considere a função h definida em $\mathbb{R}^m$ e com valores em $\mathbb{R}$ dada por

$$h(x_1, \dots, x_m) = ||(x_1, \dots, x_m)|| \left(= \sqrt{x_1^2 + \cdots + x_m^2}\right)$$

Tome-se um elemento qualquer de $\mathbb{R}^m$, $(a_1, \dots, a_m)$; vamos provar que g é contínua nesse ponto (provando assim que é contínua em $\mathbb{R}^m$). Dado $\epsilon > 0$,

$$\Big| h(x_1, \dots, x_m) - h(a_1, \dots, a_m) \Big| = \Big| ||(x_1, \dots, x_m)|| - ||(a_1, \dots, a_m)|| \Big|$$

Como vimos nas práticas, $\Big| ||(x_1,\ldots,x_m)|| - ||(a_1,\ldots,a_m)|| \Big| < ||(x_1,\ldots,x_m)-(a_1,\ldots,a_m)||$ donde, com $\delta = \epsilon$ vem,

$$\begin{aligned} &||(x_1,\ldots,x_m)-(a_1,\ldots,a_m)|| < \delta(=\epsilon) \quad \Longrightarrow \\ &\qquad \Longrightarrow \Big| h(x_1,\ldots,x_m) - h(a_1,\ldots,a_m) \Big| = \Big| ||(x_1,\ldots,x_m)|| - ||(a_1,\ldots,a_m)|| \Big| < \\ &\qquad\qquad < ||(x_1,\ldots,x_m)-(a_1,\ldots,a_m)|| < \epsilon \end{aligned}$$

e salientando só o que nos interessa:

$$||(x_1,\ldots,x_m)-(a_1,\ldots,a_m)|| < \delta \quad \Longrightarrow \quad \Big| h(x_1,\ldots,x_m) - h(a_1,\ldots,a_m) \Big| < \epsilon$$

provando assim que h é contínua em $(a_1,\ldots,a_m)$ e portanto em qualquer elemento de $\mathbb{R}^m$, já que $(a_1,\ldots,a_m)$ é genérico em $\mathbb{R}^m$.

Proposição 4.1 *Seja f uma função de domínio D contido em $\mathbb{R}^m$ e valores em $\mathbb{R}^p$, e $a \in D$. f é contínua em a se, e só se, para toda a sucessão (x_n) contida em D e convergente para a, a sucessão $\big(f(x_n)\big)$ convirja para $f(a)$.*

Dem. Seja f contínua em a e seja (x_n) uma sucessão convergente contida em D convergente para a. Dado $\epsilon > 0$ existe um $\delta > 0$ tal que, se $x \in D$ e $||x-a|| < \delta$ então $||f(x)-f(a)|| < \epsilon$. Como (x_n) converge para a então existe um inteiro positivo N tal que $||x_n - a|| < \delta$ sempre que $n > N$. E então como $x_n \in D$, qualquer que seja o inteiro positivo n, ter-se-á tambem para $n > N$, $||f(x_n)-f(a)|| < \epsilon$, o que prova que $f(x_n)$ converge para $f(a)$.

Suponhamos agora que f não é contínua em a. Então existe $\epsilon > 0$ tal que, quaquer que seja o $\delta > 0$ haverá pelo menos um ponto pertencente a D satisfazendo simultâneamente:

$$||x-a|| < \epsilon \qquad \text{e} \qquad ||f(x)-f(a)|| \geq \delta$$

Fazendo $\epsilon_n = \frac{1}{n}$ poderá portanto escolher-se uma sucessão de termos em D para a (como resulta da primeira das desigualdades acima) e tal que $\Big(f(x_n)\Big)$ não converge para $f(a)$ (resultando esta da segunda desigualdade acima) o que termina a demonstração. ∎

De forma sugestiva, embora um pouco imprecisa, pode dizer-se que a continuidade de f no ponto a equivale à possibilidade de permutar os símbolos "f" e "lim":

$$\lim f(x_n) = f(\lim x_n)$$

Através deste teorema e/ou de manipulações simples das definições obtêm-se os seguintes resultados (cujas demonstrações deixamos a cargo de quem ler estas notas):

Proposição 4.2

Sejam f e g funções definidas num domínio D contido em $\mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Seja α uma função definida em D e com valores em $\mathbb{R}$. Então, se f, g e α são contínuas em $a \in D$ então

- $f \pm g$ tambem são contínuas em a
- $f \cdot g$ tambem é contínua em a
- αf (e $\frac{1}{\alpha} f$ se $\alpha(a) \neq 0$)tambem é contínua em a

Se f e g assumem valores em $\mathbb{R}$, então

- fg tambem é contínua em a
- $\frac{f}{g}$ é contínua em a, se $g(a) \neq 0$.

(Continuidade da função composta) Seja f uma função definida num domínio D contido em $\mathbb{R}^m$ e com valores em $\mathbb{R}^p$, e g uma função definida num domínio D' (que contem $f(D)$) e com valores em $\mathbb{R}^q$.

- Se f é contínua em $a \in D$ e g é contínua em $f(a)$, então $g \circ f$ é contínua em a .

Com estes resultados, a análise da continuidade das funções dadas por expressões fica muito facilitado uma vez estabelecida a continuidade das seguintes funções:

Dado um inteiro positivo m e um i com $1 \leq i \leq m$, seja p_i a função de domínio $\mathbb{R}^m$ e valores em $\mathbb{R}$ dada por:

$$p_i(x_1, \ldots, x_i, \ldots, x_m) = x_i$$

Vamos mostrar que esta função é contínua em qualquer ponto $a = (a_1, \ldots, a_i, \ldots, a_m)$ de $\mathbb{R}^m$.Vejamos o que quer dizer aqui $|\, p_i(x_1, \ldots, x_i, \ldots, x_m) - p_i(a_1, \ldots, a_i, \ldots, a_m)|$. Como

$$|\, p_i(x_1, \ldots, x_i, \ldots, x_m) - p_i(a_1, \ldots, a_i, \ldots, a_m)| = |\, x_i - a_i \,| = \sqrt{(x_i - a_i)^2} \leq$$
$$\leq \sqrt{(x_1 - a_1)^2 + \cdots + (x_i - a_i)^2 + \cdots + (x_m - a_m)^2} = ||(x_1, \ldots, x_i, \ldots, x_m) - (a_1, \ldots, a_i, \ldots, a_m)||$$

Então, dado $\epsilon > 0$, para se ter $|\, p_i(x_1, \ldots, x_i, \ldots, x_m) - p_i(a_1, \ldots, a_i, \ldots, a_m)| < \epsilon$, basta tomar $\delta = \epsilon$ pois com $||(x_1, \ldots, x_i, \ldots, x_m) - (a_1, \ldots, a_i, \ldots, a_m)|| < \delta$ vem

$$|\, p_i(x_1, \ldots, x_i, \ldots, x_m) - p_i(a_1, \ldots, a_i, \ldots, a_m)| \leq ||(x_1, \ldots, x_i, \ldots, x_m) - (a_1, \ldots, a_i, \ldots, a_m)|| < \epsilon (= \delta)$$

provando assim que p_i é contínua em a e sendo este um ponto genérico de $\mathbb{R}^m$, então p_i é contínua em $\mathbb{R}^m$.

Exemplo 4.4

Considere a função

$$f(x_1, x_2) = \arctan\left(\frac{x_1^3 + x_2^2}{1 - x_1^2}\right)$$

que está definida no domínio D:

$$D = \{(x_1, x_2) \mid 1 - x_1 \neq 0\} = \mathbb{R}^2 \setminus \{(-1, 0), (1, 0)\}$$

Esta função é a composta das funções

$$g(x_1, x_2) = \frac{x_1^3 + x_2^2}{1 - x_1^2}$$

e

$$h(u) = \arctan(u)$$

isto é

$$f(x_1, x_2) = h \circ g(x_1, x_2)$$

$h(u) = \arctan(u)$ é uma função contínua de u. Quanto a g podemos reescrevê-la:

$$g(x_1, x_2) = \frac{\big(p_1(x_1, x_2)\big)^3 + \big(p_2(x_1, x_2)\big)^2}{1 - \big(p_1(x_1, x_2)\big)^2}$$

Como as funções p_i são contínuas e a soma, produto e divisão (sempre que a função no denominador não se anule, como é o caso) de funções contínuas é uma função contínua, então g é uma função contínua em $D = \mathbb{R}^2 \setminus \{(-1, 0), (1, 0)\}$. Assim a composta $f = h \circ g$ é uma função contínua em D.

E se a função tomar valores em $\mathbb{R}^m$ com $m > 1$? Por exemplo, a função

$$f(r, \theta) = \big(r\cos(\theta), r\sin(\theta)\big)$$

com as **funções coordenadas**

$$f_1(r, \theta) = r\cos(\theta), \qquad f_2(r, \theta) = r\sin(\theta)$$

definida para $r > 0$ e $\theta \in [0, 2\pi[$

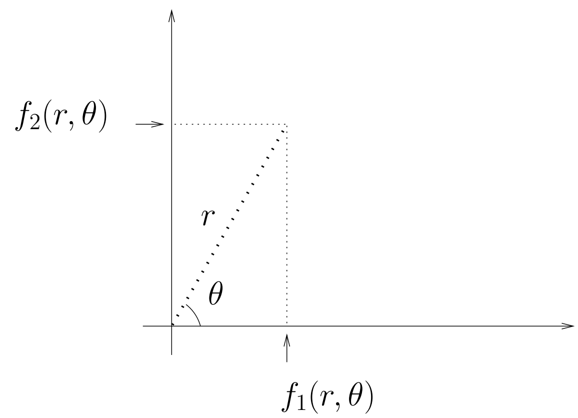

Figure 7: Função f

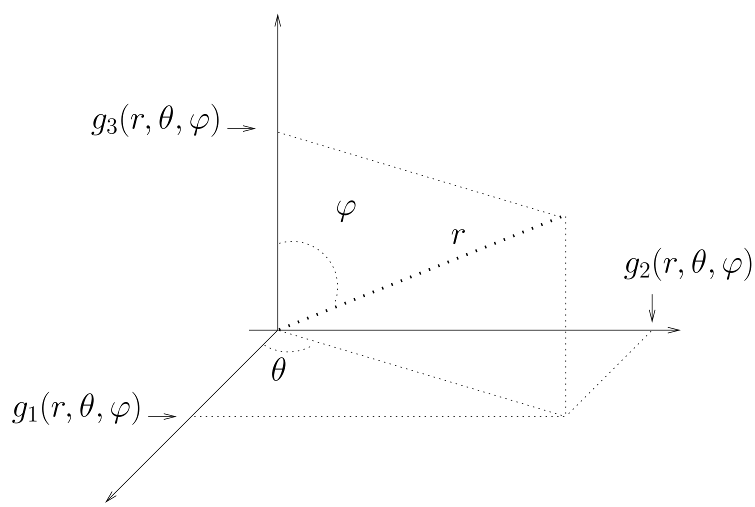

Figure 8: Função g

ou a função

$$g(r,\theta,\varphi) = \big(r\cos(\theta)\sin(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\varphi)\big)$$

com as funções coordenadas

$$g_1(r,\theta,\varphi) = r\cos(\theta)\sin(\varphi), \qquad g_2(r,\theta,\varphi) = r\sin(\theta)\sin(\varphi), \qquad g_3(r,\theta,\varphi) = r\cos(\varphi)$$

(qual é o domínio de g?) Para isso contamos com o seguinte resultado,

Proposição 4.3 *Seja f uma função definida num domínio D contido em $\mathbb{R}^m$ e com valores em $\mathbb{R}^p$. f é contínua em $a \in D$ se, e só se cada uma das funções coordenadas for contínua em a.*

Dem. Omitida. ■

Exercício 4.1

As funções f e g acima são contínuas?

Vamos agora ver como a topologia dos domínios afecta algumas características das funções contínuas definidas nesses domínios.

Proposição 4.4 *Seja f uma função contínua, definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Se D é fechado e limitado então $f(D)$ tambem é fechado e limitado.*

Dem. Vamos provar que qualquer sucessão em $f(D)$ tem pelo menos uma sucessão convergente para um ponto de $f(D)$. Seja então y_n uma subsucesão em $f(D)$. Para cada n escolha-se um ponto $x_n \in D$ tal que $f(x_n) = y_n$. Como D é limitado e fechado, (x_n) tem pelo menos uma subsucessão, chamemos-lhe $(x_{\alpha(n)})$, convergente para um ponto $a \in D$. Como f é contínua em D (e portanto em a) então $f(x_{\alpha(n)}) = y_{\alpha(n)}$ converge para o ponto $f(a)$ o que termina a demonstração. ■

Proposição 4.5 *Seja f uma função contínua, definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Se D é conexo então $f(D)$ tambem é conexo.*

Dem. Provaremos o contrarecíproco, ou seja se $f(D)$ não é conexo então D tambem não é conexo. Suponha-se então que $f(D)$ não é conexo. Existem então dois conjuntos separados, $A^\dagger$ e $B^\dagger$, isto é $A^\dagger \neq \emptyset \neq B^\dagger$ com $\overline{A^\dagger} \cap B^\dagger = \emptyset = \overline{B^\dagger} \cap A^\dagger$ tais que

$$f(D) = A^\dagger \cup B^\dagger$$

Seja A (B, respect.) o conjunto dos pontos de D que se transformam em pontos de $A^\dagger$ ($B^\dagger$, respect.) isto é:

$$A = \{x \in D \mid f(x) \in A^\dagger\} \qquad\qquad B = \{x \in D \mid f(x) \in B^\dagger\}$$

Então A e B são não vazios (porque $f(A) = A^\dagger$ e analogamente para B) e $A \cup B = D$. Vamos agora ver que

$$A \cap \bar{B} = \emptyset \qquad\qquad B \cap \bar{A} = \emptyset$$

Provaremos apenas a primeira das igualdades já que a segunda é analoga. Seja então a um ponto de A aderente a B. Então existe uma sucessão de pontos, (x_n), em B que converge para a. Pela continuidade de f, $\big(f(x_n)\big)$ converge para $f(a) \in A^\dagger$ mas como $x_n \in B$ para cada n então $f(x_n) \in B^\dagger$ para cada n donde o seu limite $f(a) \in \overline{B^\dagger}$ isto é $A^\dagger \cap \overline{B^\dagger} \neq \emptyset$ o que é absurdo pois assumimos que $A^\dagger$ e $B^\dagger$ são separados, o que termina a demonstração. ■

5 Limite

Antes de começarmos formalmente a falar de limites, recordemos a noção de continuidade.

Dada uma função f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$, f diz-se contínua em a pertencente a D se, por definição

$$\text{para todo o } \epsilon > 0, \text{ existe } \delta > 0 \text{ tal que } \|\, x - a \,\| < \delta \Rightarrow \|\, f(x) - f(a) \,\| < \epsilon$$

Como exemplo de tais funções falámos da função constante, da função identidade, da função "norma de", etc. Atentemos agora ao caso da função f definida em $\mathbb{R}^2 \setminus \{(0,0)\}$ e com valores em $\mathbb{R}$, dada por:

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

Apesar de esta função não estar definida em $(0,0)$, tem-se:

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

de onde obtemos:

$$|\, f(x,y) - 0 \,| \leq \sqrt{x^2 + y^2}$$

e assim, dado $\epsilon > 0$, tomando $\delta = \epsilon$ vem

$$\sqrt{(x-0)^2 + (y-0)^2} = \sqrt{x^2 + y^2} \leq \delta \Rightarrow |\, f(x,y) - 0 \,| \big(\leq \sqrt{x^2 + y^2}\big) \leq \epsilon$$

isto é,

$$\text{para todo o } \epsilon > 0, \text{ existe } \delta > 0 \text{ tal que } \|\, (x,y) - (0,0) \,\| \leq \delta \Rightarrow |\, f(x,y) - 0 \,| \leq \epsilon$$

ou seja apesar de a função não estar definida no elemento $(0,0)$, quando o argumento (x,y) se aproxima de $(0,0)$ os valores da função aproximam-se de 0 com o significado contido na expressão acima. Diz-se então que a função tem limite no ponto $(0,0)$. A definição de limite é então:

Definição 5.1

Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Diz-se que f tem limite $l \in \mathbb{R}^p$ num ponto $b \in \bar{D}$ se

$$\text{para todo o } \epsilon > 0, \text{ existe } \delta > 0 \text{ tal que } ||\, x - b \,|| < \delta \Rightarrow ||\, f(x) - l \,|| < \epsilon$$

e denota-se

$$\lim_{x \mapsto b} f(x) = l$$

Exemplo 5.1

Toda a função f, contínua num ponto a, tem limite $l = f(a)$ nesse ponto a.

Exemplo 5.2 (Prolongamento por continuidade)

Considere-se novamente o caso acima,

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

Apesar da função não estar definida no ponto $(0, 0)$ ela tem limite nesse ponto. Podemos assim definir uma nova função:

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

Esta nova função $\bar{f}$ é igual a f onde f já estava definida, sendo portanto aí contínua, e é igual a 0 em $(0, 0)$. Mas a sua definição em $(0, 0)$ foi feita de tal modo que $\bar{f}$ é tambem aí contínua. $\bar{f}$ diz-se então um prolongamento por continuidade de f ao ponto $(0, 0)$ que é um ponto aderente ao domínio de f. Em geral,

Definição 5.2

Seja f definida e contínua em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Diz-se que f é prolongável por continuidade a um ponto b pertencente a $\bar{D} \setminus D$ se existir o limite

$$\lim_{x \mapsto b} f(x)$$

A função

$$\bar{f}(x) = \begin{cases} f(x), & \text{se } x \in D \\ \lim_{x \mapsto b} f(x), & \text{se } x = b \end{cases}$$

chama-se prolongamento por continuidade de f a b.

Exemplo 5.3

Consideremos agora a função:

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

que está definida em $\mathbb{R}^2 \setminus \{(0, 0)\}$ e assume valores em $\mathbb{R}$. f é claramente uma função contínua no seu domínio (porquê?). Será que existe o limite de f quando (x, y) tende para $(0, 0)$? Comecemos por notar que ao longo do eixo dos XX, ou seja em pontos de $\mathbb{R}^2 \setminus \{(0, 0)\}$ da forma $(x, 0)$ se tem:

$$f(x, 0) = 1$$

ao passo que ao longo do eixo dos YY, isto é, em pontos de $\mathbb{R}^2 \setminus \{(0, 0)\}$ da forma $(0, y)$ se tem

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

Então num vizinhança de $(0,0)$ tão pequena quanto se queira, a função f assume valores tão distantes entre si como 1 e -1. Em particular, se existe o limite indicado (chamemos-lhe l), então, com $\epsilon = 1$ e para qualquer $\delta > 0$ tem-se $\| (\frac{\delta}{2}, 0) - (0,0) \| < \delta$ e $\| (0, \frac{\delta}{2}) - (0,0) \| < \delta$, donde

$$2 = | f(\frac{\delta}{2}, 0) - f(0, \frac{\delta}{2}) | = | \big(f(\frac{\delta}{2}, 0) - l\big) - \big(f(0, \frac{\delta}{2}) - l\big) | < | \big(f(\frac{\delta}{2}, 0) - l\big) | + | \big(f(0, \frac{\delta}{2}) - l\big) | < 1 + 1 = 2$$

ou seja $2 < 2$ o que é absurdo. Então o referido limite não existe.

Registemos aqui alguns resultados analogos a resultados que já vimos aquando do estudo da continuidade.

Proposição 5.1 *Seja f uma função de domínio D contido em $\mathbb{R}^m$ e valores em $\mathbb{R}^p$, e $b \in \bar{D}$. $\lim_{x \mapsto b} f(x) = l$ se, e só se, para toda a sucessão (x_n) contida em D e convergente para b, a sucessão $\big(f(x_n)\big)$ convirja para l.*

■

Proposição 5.2 *Seja f uma função de domínio D contido em $\mathbb{R}^m$ e valores em $\mathbb{R}^p$, e $b \in \bar{D}$. Designemos por f_j a j-ésima função coordenada de f e seja $l = (l_1, \ldots, l_p)$ um elemento de $\mathbb{R}^p$. $\lim_{x \mapsto b} f(x) = l$ se, e só se, $\lim_{x \mapsto b} f_j(x) = l_j$ (para todo o $j = 1, \ldots, m$).*

■

À semelhança do que acontecia com a continuidade, a existência de limite num ponto para duas funções comunica-se a certas funções obtidas dessas:

Proposição 5.3

Sejam f e g funções definidas num domínio D contido em $\mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Seja α uma função definida em D e com valores em $\mathbb{R}$. Então, se f, g e α têm limite em $b \in \bar{D}$ então tambem têm limite, no ponto b, as funções:

- $f \pm g$ com $\lim_{x \mapsto b}\big(f(x) \pm g\big) = \lim_{x \mapsto b}\big(f(x)\big) \pm \lim_{x \mapsto b}\big(g(x)\big)$
- $\| f \|$ com $\lim_{x \mapsto b} \| f(x) \| = \| \lim_{x \mapsto b} f(x) \|$
- $f \cdot g$ com $\lim_{x \mapsto b}\big(f(x) \cdot g\big) = \lim_{x \mapsto b}\big(f(x)\big) \cdot \lim_{x \mapsto b}\big(g(x)$
- αf (e $\frac{1}{\alpha} f$ se $\alpha(a) \neq 0$, respect.) com $\lim_{x \mapsto b}\big(\alpha f\big)(x) = \lim_{x \mapsto b} \alpha(x) \lim_{x \mapsto b} f(x)$ (com $\lim_{x \mapsto b} \frac{f}{\alpha}(x) = \frac{\lim_{x \mapsto b} f(x)}{\lim_{x \mapsto b} \alpha(x)}$, respect.)

(Limite da função composta) Seja f uma função definida num domíinio D contido em $\mathbb{R}^m$ e com valores em $\mathbb{R}^p$, g uma função definida num domínio D' (que contem $f(D)$) e com valores em $\mathbb{R}^q$, $l_1 \in \mathbb{R}^p$ e $l_2 \in \mathbb{R}^q$.

- Se $\lim_{x \mapsto b} f(x) = l_1$ e $\lim_{x \mapsto l_1} g(x) = l_2$ então $\lim_{x \mapsto b} g\big(f(x)\big) = l_2 = \lim_{x \mapsto l_1} g(x)$.

Exemplo 5.4

Atendendo a que

$$\lim_{(x,y) \mapsto (0,0)} \frac{x^2 - y^2}{\sqrt{x^2 + y^2}} = 0$$

e que

$$\lim_{u \mapsto 0} \sec(u) = 1$$

tem-se

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

5.1 Limites Direccionais

Os limites direccionais são limites que se tomam ao longo de uma recta passando pelo ponto em que se está a calcular o limite.

Definição 5.3 (Limite Direccional)

O limite direccional segundo a direcção v da função f no ponto a é:

$$\lim_{t \mapsto 0} f(a + tv)$$

Exemplo 5.5

Considere-se a função

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

Pretendemos calcular o limite desta função no ponto $(0, 0)$ ao longo de uma direcção v cujo declive é m. Por outras palavras pretendemos calcular o limite quando $(x, y) \mapsto (0, 0)$ ao longo do conjunto

$$A_m = \{(x, y) \mid y = mx \quad \text{e} \quad x \neq 0\}$$

Tem-se então:

$$\lim_{\substack{(x,y)\mapsto(0,0) \\ (x,y)\in A_m}} \frac{xy}{x^2 + y^2} = \lim_{x \mapsto 0^+} \frac{x \cdot mx}{x^2 + (mx)^2} = \lim_{x \mapsto 0^+} \frac{x^2 \cdot m}{x^2(1 + m^2)} = \frac{m}{1 + m^2}$$

Verificamos então que para cada inclinação m ter-se-à um limite distinto. Esta função não tem, portanto, limite no ponto $(0, 0)$. Este resultado permite-nos desde já realçar um aspecto importante dos limites direccionais: eles **podem ser** úteis na prova de que um certo limite não existe, como o exemplo atrás ilustra.

Será que por outro lado, sempre que **todos** os limites direccionais existam e sejam iguais, então a função tem limite no ponto em estudo?

Exercício 5.1

Estude os limites em $(0, 1)$ das seguintes funções de $\mathbb{R}^2$ em $\mathbb{R}$:

1.
$$f(x) = \begin{cases} 1, & \text{se } x \neq 0 \text{ e } y = x^2 \\ 0, & \text{se } x = 0 \text{ ou } y \neq x^2 \end{cases}$$

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

6 Diferenciabilidade

6.1 Introdução

A diferenciabilidade de uma função é uma questão local, isto é, tem a ver com o que se passa numa vizinhança de um ponto.

Recordemos alguns aspectos da diferenciabilidade em $\mathbb{R}$:

Quão depressa varia a função "junto" a um ponto a?

Medimos esta grandeza através do quociente

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

dito razão incremental, isto é o quociente de quanto a função varia sobre quanto o argumento varia no intervalo $[a, a + h]$. Obtemos assim uma grandeza cujas dimensões são as dimensões da função f sobre as dimensões do argumento. Exemplo: velocidade "=" espaço percorrido sobre tempo decorrido: metros por segundo ou algo equivalente, km por hora, milhas por hora, etc. Mas o que nos interessa é o que se passa sobre o ponto a. Consideramos então a razão incremental para intervalos $[a, a+h]$ com h cada vez mais pequeno, ou seja tomamos o limite quando h tende para zero. Se este limite existe, isto é, se existe

$$\lim_{h \mapsto 0} \frac{f(a+h) - f(a)}{h}$$

dizemos que a função f tem derivada no ponto a com o valor

$$f'(a) = \lim_{h \mapsto 0} \frac{f(a+h) - f(a)}{h}$$

e este valor é então a medida de quão depressa a função varia sobre o ponto a.

Graficamente o que é que isto quer dizer?

O facto de a função f ser diferenciável num ponto a quer dizer que podemos associar uma recta tangente ao gráfico de f no ponto $(a, f(a))$ - cujo declive vai ser precisamente $f'(a)$.

Aproximar a função f diferenciável no ponto a por outras funções mais simples:

Se, dada uma função f, se conhecer exactamente o seu valor num ponto a assim como o valor das suas derivadas, mas só nesse ponto, será possível aproximá-la por uma função mais simples, mais fácil de trabalhar?

Exemplo 6.1

Por exemplo, a função

$$f(x) = \sin(x)$$

cujo valor no ponto $x = \frac{\pi}{6}$ é bem conhecido assim como o das suas derivadas - embora numa pequena vizinhança de $x = \frac{\pi}{6}$ não saibamos o valor exacto desta função. Poderíamos aproximá-la junto a $x = \frac{\pi}{6}$ pela função

$$g(x) = \frac{1}{2}$$

mas a aproximação seria melhor se se usasse a função

$$h(x) = \frac{1}{2} + \frac{\sqrt{3}}{2}(x - \frac{\pi}{6})$$

De um modo geral, se uma função f é diferenciável num ponto a, então, como vimos no final do capítulo sobre fórmula de Taylor:

$$f'(a) = \lim_{h \mapsto 0} \frac{f(a+h) - f(a)}{h} \Leftrightarrow f(a+h) - f(a) = f'(a) \cdot h + o(h)$$

onde $o(h)$ é uma função que satisfaz:

$$\lim_{h \mapsto 0} \frac{o(h)}{h} = 0$$

coloquialmente dizemos que $o(h)$ é uma função que tende mais depressa para zero do que h quando $h \mapsto 0$. Então o facto de f ser diferenciável em a exprime tambem o facto de f ser "bem" aproximada por uma constante mais uma função linear no afastamento h em relação ao ponto aonde os valores da f e da sua derivada são conhecidos. Por outras palavras, a diferença

$$f(a+h) - f(a)$$

é bem aproximada por

$$f'(a) \cdot h$$

que é uma aplicação linear em h.

Recordemos aqui que uma aplicação linear, L, (tambem conhecida por transformação linear) é uma função entre espaços lineares, V_1, V_2 (tambem conhecidos por espaços vectoriais),

$$L : V_1 \longrightarrow V_2$$

satisfazendo

$$L(u+w) = L(u) + L(w) \qquad \text{e} \qquad L(\alpha u) = \alpha L(u) \qquad \text{para todos os } u, w \in V_1, \text{ e } \alpha \in \mathbb{R}$$

Como sabemos da Álgebra Linear, uma aplicação linear, L, fica univocamente representada por uma matriz desde que se fixem bases nos espacos lineares V_1 e V_2.Os espaços lineares que nos vão interessar de seguida são os $\mathbb{R}^m$ juntamente com as bases ditas **canónicas**, cujos elementos são:

$$e_1 = (1, 0, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \dots \quad , e_m = (0, 0, 0, \dots, 1)$$

Em particular, as transformações lineares de $\mathbb{R}$ em $\mathbb{R}$ têm a forma

$$T(x) = cx$$

onde c é uma constante.

O que é então a diferenciabilidade em $\mathbb{R}^m$?

Se a função em estudo, f, tem por domínio um subconjunto de $\mathbb{R}^m$ então NÃO faz sentido escrever $\frac{f(a+h)-f(a)}{h}$ porque $h \in \mathbb{R}^m$ e a divisão NÃO está definida em $\mathbb{R}^m$, logo tambem não fará sentido tomar o limite daquela razão incremental quando o h tende para zero. No entanto como vimos atrás, há uma condição equivalente à existência desse limite que é existir uma aplicação linear, $f'(a)$, e uma função que tende para zero mais depressa que h quando h tende para zero que tornam verdadeira a igualdade:

$$f(a+h) = f(a) + f'(a) \cdot h + o(|h|), \qquad \text{em que } \frac{o(|h|)}{|h|} \mapsto 0, \text{ quando } h \mapsto 0$$

Definimos, então, diferenciabilidade em $\mathbb{R}^m$

Definição 6.1 (Diferenciabilidade)

Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$. Diz-se que f é diferenciável em $a \in D$ se existir uma aplicação linear L_a e uma função $o(||h||)$ quando $h \mapsto 0$ para as quais se tem:

$$f(a+h) = f(a) + L_a(h) + o(||h||), \qquad \text{em que } \frac{o(||h||)}{||h||} \mapsto \underbrace{(0, 0, \dots, 0)}_{p\ zeros}, \text{ quando } ||h|| \mapsto 0$$

A aplicação linear L_a chama-se aplicação linear derivada de f em a que tambem se costuma denotar por

$$f'(a)$$

Exemplo 6.2

1. Função identidade:

$$f : \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

tal que $f(x) = x$, para todo o $x \in \mathbb{R}^m$. Vamos mostrar que esta função é diferenciável num ponto $a \in \mathbb{R}^m$, tentando escrever a diferença $f(a+h) - f(a)$ à custa de uma aplicação linear calculada em $h \in \mathbb{R}^m$ mais um o pequeno de $||h||$:

$$f(a+h) - f(a) = a + h - a = h = (h_1, \dots, h_m) = \cdots$$

onde os h_i's são as coordenadas de h na base canónica de $\mathbb{R}^m$,

$$\cdots = (h_1, \dots, h_m) + (0, \dots, 0) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} + (0, \dots, 0)$$

Então, L_a é representada, na base canónica, pela matriz

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

e a função $o(||h||)$ é a função constante $(0, 0, \dots, 0)$. De facto,

$$\frac{(0, 0, \dots, 0)}{||h||} = (\frac{0}{||h||}, \frac{0}{||h||}, \dots, \frac{0}{||h||}) = (0, 0, \dots, 0) \underset{||h|| \mapsto 0}{\mapsto} (0, 0, \dots, 0)$$

Assim, obtivémos

$$f(a + h) = f(a) + L_a(h) + o(||h||)$$

onde L_a é, fixadas a base canónica em $\mathbb{R}^m$, representada pela matriz

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

e

$$o(||h||) = (0, 0, \dots, 0)$$

Então, a função identidade $f(x) = x$ é diferenciável em qualquer ponto a de $\mathbb{R}^m$ sendo a aplicação linear derivada nesse ponto a representada pela matriz acima (matriz identidade) fixada a base canónica em $\mathbb{R}^m$.

2. Função constante. Seja $b = (b_1, \dots, b_p)$ um elemento de $\mathbb{R}^p$ e considere-se a função

$$g : \mathbb{R}^m \longrightarrow \mathbb{R}^p$$

dada por

$$g(x) = b$$

Vamos averiguar se g é diferenciável num ponto $a \in \mathbb{R}^m$. Seja então $h \in \mathbb{R}^m$. Tem-se:

$$g(a+h)-g(a) = b-b = \underbrace{(0, 0, \dots, 0)}_{p\ zeros} = \underbrace{(0, 0, \dots, 0)}_{p\ zeros} + \underbrace{(0, 0, \dots, 0)}_{p\ zeros} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} + \underbrace{(0, 0, \dots, 0)}_{p\ zeros}$$

(onde a matriz acima é a matriz cujas entradas são todas nulas) e portanto conseguimos obter

$$g(a + h) = g(a) + L_a(h) + o(||h||)$$

e portanto g é diferenciável em a. Como a é um ponto qualquer de $\mathbb{R}^m$, então g é diferenciável em qualquer ponto a em $\mathbb{R}^m$ sendo a aplicação linear derivada nesse ponto a representada pela matriz identicamente nula, fixadas as bases canónicas em $\mathbb{R}^m$ e $\mathbb{R}^p$.

3. Funções "projecção. Dado i tal que $1 \leq i \leq m$, considere-se a função:

$$p_i : \mathbb{R}^m \longrightarrow \mathbb{R}$$

dada por

$$p_i(x_1, \dots, x_m) = x_i$$

Dado um ponto a em $\mathbb{R}^m$ e h em $\mathbb{R}^m$, tem-se:

$$p_i(a + h) - p_i(a) = p_i(a_1 + h_1, \dots, a_m + h_m) - p_i(a_1) = a_i + h_i - a_i = h_i =$$

$$= \begin{pmatrix} 0 & 0 & \dots & 0 & \overset{i-ésima}{1} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} + 0$$

que, com considerações analogas às dos exercícios anteriores nos leva a concluir que p_i é uma função diferenciável em qualquer ponto a de $\mathbb{R}^m$, sendo a aplicação linear derivada em cada ponto a representada pela matriz

$$\begin{pmatrix} 0 & 0 & \dots & 0 & \overset{i-\acute{e}sima}{1} & 0 & \dots & 0 \end{pmatrix}$$

fixadas as bases canónicas em $\mathbb{R}^m$ e em $\mathbb{R}$.

4. Considere-se

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

dada por

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

Dado um ponto (a_1, a_2) em $\mathbb{R}^2$ e (h_1, h_2) em $\mathbb{R}^2$, tem-se:

$$\begin{aligned} & f\big((a_1,a_2)+(h_1,h_2)\big) - f\big((a_1,a_2)\big) = f\big((a_1+h_1, a_2+h_2)\big) - f\big((a_1,a_2)\big) = \\ & = (a_1+h_1)^2 + (a_2+h_2)^2 - \big(a_1^2+a_2^2\big) = a_1^2 + 2a_1h_1 + h_1^2 + a_2^2 + 2a_2h_2 + h_2^2 - a_1^2 - a_2^2 = \\ & = 2a_1h_1 + h_1^2 + 2a_2h_2 + h_2^2 = 2a_1h_1 + 2a_2h_2 + h_1^2 + h_2^2 = \begin{pmatrix} 2a_1 & 2a_2 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + h_1^2 + h_2^2 \end{aligned}$$

e argumentando como atrás só nos falta ver que $h_1^2 + h_2^2$ é um o-pequeno de h quando $||h||$ tende para zero:

$$\frac{h_1^2 + h_2^2}{||h||} = \frac{h_1^2 + h_2^2}{\sqrt{h_1^2 + h_2^2}} = \sqrt{h_1^2 + h_2^2} = ||h|| \underset{||h|| \mapsto 0}{\longmapsto} 0$$

e portanto f é diferenciável em qualquer (a_1, a_2) em $\mathbb{R}^2$, sendo a aplicação linear derivada num tal ponto (a_1, a_2) representada pela matriz

$$\begin{pmatrix} 2a_1 & 2a_2 \end{pmatrix}$$

fixadas as bases canónicas em $\mathbb{R}^2$ e $\mathbb{R}$. Note-se que este é o primeiro destes quatro exemplos em que a aplicação linear derivada varia de ponto para ponto; nos primeiros três exemplos a aplicação linear derivada era constante, era sempre a mesma independentemente do ponto onde estivesse a ser calculada.

Proposição 6.1 *Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$. f é diferenciável em $a \in D$ se, e só se, cada uma das suas funções coordenadas for diferenciável em a.*

Dem. Omitida. ■

6.2 Derivadas direccionais; derivadas parciais

Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}$, seja a um ponto de D e seja v um vector de $\mathbb{R}^m$.

Definição 6.2 (Derivada de f em a segundo v)

Se existir o limite:

$$\lim_{t \mapsto 0} \frac{f(a+tv) - f(a)}{t}$$

designa-se por **derivada de f em a segundo v** e denota-se:

$$D_v f(a) \quad \text{ou} \quad f'_v(a) \quad \text{ou} \quad \frac{\partial f}{\partial v}(a)$$

Se v na definição acima for um vector unitário, isto é $||v|| = 1$, então a derivada de f em a segundo v chama-se **derivada direccional de f em a, na direcção e sentido de v.** A derivada direccional pode interpretar-se como uma "taxa média de variação de f por unidade de comprimento.

Exemplo 6.3

Considere-se a função

$$f(x,y) = x^2 y$$

e calculemos a derivada de f em (a,b) segundo um vector $v = (\alpha, \beta)$. Tem-se:

$$D_{(\alpha,\beta)}f(a,b) = \lim_{t \mapsto 0} \frac{f\big((a,b)+t(\alpha,\beta)\big) - f\big((a,b)\big)}{t} = \lim_{t \mapsto 0} \frac{(a+t\alpha)^2(b+t\beta) - a^2 b}{t} =$$
$$= \lim_{t \mapsto 0} \frac{(a^2 + 2ta\alpha + t^2\alpha^2)(b+t\beta) - a^2 b}{t} = \lim_{t \mapsto 0} \frac{a^2b + 2tab\alpha + bt^2\alpha^2 + a^2t\beta + 2t^2a\beta\alpha + bt^3\beta\alpha^2 - a^2b}{t} =$$
$$= \lim_{t \mapsto 0} \frac{2tab\alpha + bt^2\alpha^2 + a^2t\beta + 2t^2a\beta\alpha + bt^3\beta\alpha^2}{t} = 2ab\alpha + a^2\beta + \lim_{t \mapsto 0}\Big(bt\alpha^2 + 2ta\beta\alpha + bt^2\beta\alpha^2\Big) =$$
$$= 2ab\alpha + a^2\beta$$

Em particular, fazendo $v = (1,0)$ e depois $v = (0,1)$, obtemos as derivadas direccionais segundo os eixos dos XX e dos YY:

$$D_{(1,0)}f(a,b) = 2ab \cdot 1 + a^2 \cdot 0 = 2ab \qquad D_{(0,1)}f(a,b) = 2ab \cdot 0 + a^2 \cdot 1 = a^2$$

As derivadas direcionais segundo os vectores unitários na direcção e sentido dos eixos, têm um nome especial; chamam-se derivadas parciais segundo o eixo em questão:

Definição 6.3 (Derivadas parciais de f em a) *A derivada direccional de f em a segundo o vector unitário na direcção e sentido do i-ésimo eixo coordenado chama-se "***derivada parcial de f em a em ordem a x_i***. A notação é:*

$$\frac{\partial f}{\partial x_i}(a)$$

e, por definição:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \mapsto 0} \frac{f(a+te_i) - f(a)}{t}$$

onde $e_i = (0,0,\ldots,0,\overset{i-ésima}{1},0,\ldots,0)$

Exemplo 6.4

1. Consideremos a função

$$f(x,y) = x^2 + \sin(xy)$$

Calcular

$$\frac{\partial f}{\partial x}(a,b)$$

corresponde a considerar a função

$$\varphi(x) = f(x,b) = x^2 + \sin(bx)$$

averiguar se esta função de x é diferenciável, em caso afirmativo, derivá-la em ordem ao x e calcular para $x = a$:

$$\varphi'(x) = 2x + b\cos(bx)$$

donde

$$\frac{\partial f}{\partial x}(a,b) = 2a + b\cos(ba)$$

Analogamente,

$$\frac{\partial f}{\partial y}(a,b) = a\cos(ba)$$

2. Considere-se a função

$$f(r,\theta) = (r\cos(\theta), r\sin(\theta))$$

portanto com as funções coordenadas

$$f_1(r,\theta) = r\cos(\theta), \qquad \text{e} \qquad f_2(r,\theta) = r\sin(\theta)$$

Tem-se:

$$\begin{array}{cc} \dfrac{\partial f_1}{\partial r} = \cos(\theta) & \dfrac{\partial f_1}{\partial \theta} = -r\sin(\theta) \\ \dfrac{\partial f_2}{\partial r} = \sin(\theta) & \dfrac{\partial f_2}{\partial \theta} = r\cos(\theta) \end{array}$$

6.3 Implicações da diferenciabilidade; conexão com as derivadas direccionais

Recordemos que se f é diferenciável em a isso quer dizer que existe uma aplicação linear L_a tal que

$$f(a+h) = f(a) + L_a(h) + o(||h||)$$

Proposição 6.2 *Se f é diferenciável em a, então*

- *f é contínua em a;*
- *para todo o v em $\mathbb{R}^m \setminus \{(0,\ldots,0)\}$, existe $D_v f(a)$.*

Dem. Sendo f diferenciável em a, então existe uma aplicação linear L_a tal que

$$f(a+h) = f(a) + L_a(h) + o(||h||) \qquad \text{quando } ||h|| \text{ tende para zero}$$

e tomando limites quando $||h|| \mapsto 0$ obtem-se:

$$\lim_{||h||\mapsto 0} f(a+h) = \lim_{||h||\mapsto 0}\Big(f(a) + L_a(h) + o(||h||)\Big) = f(a) + \lim_{||h||\mapsto 0}\Big(L_a(h) + o(||h||)\Big) = \ldots$$

... porque uma aplicação linear é contínua em qualquer ponto e aplica o vector de coordenadas todas nulas no vector de coordenadas todas nulas ...

$$\cdots = f(a) + \lim_{||h||\mapsto 0}\Big(o(||h||)\Big) = f(a)$$

... em consequência da definição de $o(||h||)$, e portanto, f é contínua em a.

Seja agora v em $\mathbb{R}^m \setminus \{(0,\ldots,0)\}$; tem-se:

$$f(a+tv) - f(a) = L_a(tv) + o(||tv||) = \ldots$$

trocando h por tv na expressão de diferenciabilidade, e

$$\cdots = tL_a(v) + t||v||\frac{o(||tv||)}{t||v||}$$

porque L_a é uma aplicação linear, donde:

$$\lim_{t\mapsto 0}\frac{f(a+tv)-f(a)}{t} = \lim_{t\mapsto 0}\left(L_a(v) + ||v||\frac{o(||tv||)}{t||v||}\right) = L_a(v) + \lim_{t\mapsto 0}||v||\frac{o(||tv||)}{t||v||} = L_a(v)$$

por definição de o-pequeno. Portanto

$$D_v f(a) = L_a(v)$$

e, portanto, existe. ■

Corolário 6.1 *Dada f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$ se f é diferenciável em $a \in D$ então a aplicação linear L_a é única. Fixadas as bases canónicas em $\mathbb{R}^m$ e $\mathbb{R}^p$, L_a é representada pela matriz:*

$$M_a^f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \ldots & \frac{\partial f_p}{\partial x_m}(a) \end{pmatrix}$$

Dem. Se f é diferenciável em a então $D_v f(a)$ existe e

$$L_a(v) = D_v f(a)$$

logo L_a é única. Fixadas as bases canónicas $\{e_1, \ldots, e_m\}$ e $\{e_1, \ldots, e_p\}$, a matriz que representa L_a tem por colunas

$$L_a(e_i) = \begin{pmatrix} D_{e_i} f_1(a) \\ D_{e_i} f_2(a) \\ \vdots \\ D_{e_i} f_p(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(a) \\ \frac{\partial f_2}{\partial x_i}(a) \\ \vdots \\ \frac{\partial f_p}{\partial x_i}(a) \end{pmatrix}$$

■

Observação 6.1

1. À matriz

$$M_a^f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \ldots & \frac{\partial f_p}{\partial x_m}(a) \end{pmatrix}$$

chama-se a **matriz jacobiana** ou simplesmente **jacobiana** de f no ponto a.

2. No caso $p = 1$, em que a matriz jacobiana se reduz a uma linha, esta é conhecida por **gradiente** com a seguinte notação:

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_m}(a) \right)$$

Observação 6.2

A matriz jacobiana de f em a é a matriz cujas entradas são as derivadas parciais de f em a na maneira acima explicitada. Se f é diferenciável em a então o cálculo das derivadas de f em a segundo uma direcção v fica muito simplificado pois, pela Proposição 6.2, esse valor é dado pelo cálculo da aplicação linear derivada em a no vector v, ou seja, fixadas as bases canónicas, pelo produto matricial da jacobiana de f em a pelo vector coluna v. Escusa-se assim de ter que usar a definição de derivada segundo uma direcção v que implicaria ter de calcular um limite que, como já é da nossa experiência, pode ser bastante complicado.

6.4 Outras implicações da diferenciabilidade

Proposição 6.3 *Seja f e g funções definidas em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$.*

- *Se f e g são diferenciáveis em $a \in D$ então $f+g$ tambem é diferenciável em a e a matriz jacobiana $f+g$ em a é a soma das jacobianas de f e de g no ponto a*
- *Se α é um número real, então αf é diferenciável no ponto a e a jacobiana de αf no ponto a obtem-se da jacobiana de f no ponto a multiplicando todas as entradas desta por α.*

Dem. Sejam f e g nas condições da proposição. Tem-se:

$$f(a+h) = f(a) + f'(a)(h) + o(||h||)$$
$$g(a+h) = g(a) + g'(a)(h) + o(||h||)$$

e somando ordenadamente obtemos,

$$(f+g)(a+h) = (f+g)(a) + f'(a)(h) + g'(a)(h) + o(||h||) + o(||h||) = (f+g)(a) + L_a(h) + o(||h||)$$

já que a soma de o-pequenos é outra vez um o-pequeno e a soma de aplicações lineares é outra vez uma aplicação linear. Assim

$$f'(a)(h) + g'(a)(h) = L_a(h) = (f+g)'(a)(h)$$

e portanto $f+g$ é diferenciável em a e a sua jacobiana é a soma das jacobianas de f e de g em a.

Seja agora α um número real. Se f é diferenciável em a,

$$f(a+h) = f(a) + f'(a)(h) + o(||h||)$$

e multiplicando ambos os lados por α,

$$\alpha f(a+h) = \alpha f(a) + \alpha f'(a)(h) + \alpha o(||h||)$$

ou seja

$$(\alpha f)(a+h) = (\alpha f)(a) + \alpha f'(a)(h) + o(||h||)$$

já que o-pequeno multiplicado por constante é outra vez um o-pequeno. Tambem porque aplicação linear multiplicada por constante é outra vez uma aplicação linear então (αf) é diferenciável em a e a sua jacobiana obtem-se da jacobiana de f em a multiplicando todos os elementos de matriz desta por α. ■

Proposição 6.4 *Seja g uma função definida em $D \subset \mathbb{R}^m$, com valores em $\mathbb{R}^p$, diferenciável em $a \in D$ e seja f uma função definidas em $g(D) \subset E \subset \mathbb{R}^p$, com valores em $\mathbb{R}^q$ e diferenciável em $g(a) = b \in E$. Então $f \circ g$ é diferenciável em a e*

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

Dem. Omitida ■

Decorre deste resultado que as jacobianas destas funções se relacionam da seguinte forma:

$$M_a^{f \circ g} = M_{g(a)}^f M_a^g$$

e explicitando algumas entradas destas matrizes:

$$M_a^f = \begin{pmatrix} \frac{\partial (f\circ g)_1}{\partial x_1}(a) & \frac{\partial (f\circ g)_1}{\partial x_2}(a) & \dots & \frac{\partial (f\circ g)_1}{\partial x_m}(a) \\ \frac{\partial (f\circ g)_2}{\partial x_1}(a) & \frac{\partial (f\circ g)_2}{\partial x_2}(a) & \dots & \frac{\partial (f\circ g)_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial (f\circ g)_p}{\partial x_1}(a) & \frac{\partial (f\circ g)_p}{\partial x_2}(a) & \dots & \frac{\partial (f\circ g)_p}{\partial x_m}(a) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(g(a)) & \frac{\partial f_1}{\partial y_2}(g(a)) & \dots & \frac{\partial f_1}{\partial y_p}(g(a)) \\ \frac{\partial f_2}{\partial y_1}(g(a)) & \frac{\partial f_2}{\partial y_2}(g(a)) & \dots & \frac{\partial f_2}{\partial y_p}(g(a)) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial y_1}(g(a)) & \frac{\partial f_p}{\partial y_2}(g(a)) & \dots & \frac{\partial f_p}{\partial y_p}(g(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \frac{\partial g_1}{\partial x_2}(a) & \dots & \frac{\partial g_1}{\partial x_m}(a) \\ \frac{\partial g_2}{\partial x_1}(a) & \frac{\partial g_2}{\partial x_2}(a) & \dots & \frac{\partial g_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1}(a) & \frac{\partial g_p}{\partial x_2}(a) & \dots & \frac{\partial g_p}{\partial x_m}(a) \end{pmatrix}$$

Em particular, para cada i em $\{1, 2, \dots, q\}$ e para cada j em $\{1, 2, \dots, m\}$, tem-se

$$\frac{\partial (f \circ g)_i}{\partial x_j}(a) = \sum_{k=1}^{p} \frac{\partial f_i}{\partial y_k}(g(a)) \frac{\partial g_k}{\partial x_j}(a)$$

Proposição 6.5 *Sejam f e g funções reais, isto é, definidas num domínio $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}$. Se f e g são diferenciáveis em $a \in D$ então o produto fg tambem é diferenciável em a e*

$$\big(fg\big)'(a) = g(a)f'(a) + f(a)g'(a)$$

Dem. Omitida ■

Corolário 6.2 *Se, além disso, $g(a) \neq 0$, então $\frac{f}{g}$ tambem é diferenciável em a e*

$$\left(\frac{f}{g}\right)' = \frac{g(a)f'(a) - f(a)g'(a)}{\big(g(a)\big)^2}$$

Dem. Omitida ■

Exemplo 6.5

Seja f definida em $\mathbb{R}^2 \setminus \{(1,0)\}$ dada por

$$f(x,y) = \Big(e^{2x+y^2}, \frac{(x-1)^2}{(x-1)^2+y^2}\Big)$$

f é diferenciável se e só se as suas funções coordenadas forem diferenciáveis.

$$f_1(x,y) = e^{2x+y^2} = \big(h_2 \circ h_1\big)(x,y)$$

Consideremos:

$$h_1(x,y) = 2x + y^2 = 2p_1(x,y) + \big(p_2(x,y)\big)^2$$

Como as funções projecção são funções diferenciáveis e somas e produtos de funções reais diferenciáveis são funções diferenciáveis, então h_1 é uma função diferenciável no seu domínio. Consideremos agora

$$h_2(z) = e^z$$

que é uma função diferenciável (cf. Análise Matemática I). Como a composta de funções diferenciáveis é uma função diferenciável, então $f_1 = h_2 \circ h_1$ é uma função diferenciável. Consideremos agora a segunda função coordenada:

$$f_2(x,y) = \frac{(x-1)^2}{(x-1)^2+y^2} = \frac{\Big(p_1(x,y)-1\Big)^2}{\Big(p_1(x,y)-1\Big)^2 + p_2(x,y)^2}$$

Novamente, as funções projecção são diferenciáveis. Produtos de funções reais diferenciáveis são funções diferenciáveis; somas de funções diferenciáveis são funções diferenciáveis. Logo o numerador e o denominador da fracção em questão são funções diferenciáveis. Finalmente, o quociente de funções reais diferenciáveis é diferenciável em todos os pontos onde o denominador não se anula. Então f_2 é uma função diferenciável no seu domínio. Finalmente, f é diferenciável no seu domínio porque as suas funções coordenadas são funções diferenciáveis.

6.5 Um critério de diferenciabilidade

O facto de existirem as derivadas parciais de uma função num ponto não significa que ela seja diferenciável nesse ponto:

Exemplo 6.6

Considere-se

$$f(x,y) = \begin{cases} 1, & \text{se } x \neq 0 \text{ e } y = x^2 \\ 0, & \text{se } x = 0 \text{ ou } y \neq x^2 \end{cases}$$

As derivadas parciais em $(0,0)$ são:

$$\frac{\partial f}{\partial x}(0,0) = D_{(1,0)}f(0,0) = \lim_{t\mapsto 0} \frac{f((0,0)+t(1,0)) - f(0,0)}{t} = \lim_{t\mapsto 0} \frac{f((t,0)) - 0}{t} = \lim_{t\mapsto 0} \frac{0}{t} = 0$$

e

$$\frac{\partial f}{\partial y}(0,0) = D_{(0,1)}f(0,0) = \lim_{t\mapsto 0} \frac{f((0,0)+t(0,1)) - f(0,0)}{t} = \lim_{t\mapsto 0} \frac{f((0,t)) - 0}{t} = \lim_{t\mapsto 0} \frac{0}{t} = 0$$

mas f não é contínua em $(0,0)$ pois

$$\lim_{\substack{(x,y)\mapsto(0,0) \\ y=0}} f(x,y) = 0 \qquad \text{enquanto que} \qquad \lim_{\substack{(x,y)\mapsto(0,0) \\ y=x^2}} f(x,y) = 1$$

e se uma função não é contínua num ponto então tambem não é diferenciável nesse ponto (6.2).

No entanto:

Proposição 6.6 *Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$ e $a \in D$. Se para todo o i em $\{1, \ldots, m\}$ e para todo o j em $\{1, \ldots, p\}$, as funções*

$$\frac{\partial f_j}{\partial x_i}$$

forem contínuas em a então f é diferenciável em a.

Dem. Omitida. ■

Assim uma maneira alternativa de averiguarmos se uma função é diferenciável num ponto a é verificarmos se as derivadas parciais são funções contínuas nesse ponto a.

Exemplo 6.7

Considere-se a função:

$$f(x, y) = \big(x^2 \cos(y), y^2 \sin(x)\big)$$

que tem por funções coordenadas as funções:

$$f_1(x, y) = x^2 \cos(y), \qquad \text{e} \qquad f_2(x, y) = y^2 \sin(x)$$

As derivadas parciais são:

$$\frac{\partial f_1}{\partial x} = 2x \cos(y); \qquad \frac{\partial f_1}{\partial y} = -x^2 \sin(y)$$

$$\frac{\partial f_2}{\partial x} = y^2 \cos(x); \qquad \frac{\partial f_2}{\partial y} = 2y \sin(x)$$

As derivadas parciais são então funções contínuas (porquê?) de (x, y), logo f é uma função diferenciável em qualquer (x, y) de $\mathbb{R}^2$.

Como estas funções derivadas parciais são diferenciáveis (porquê?) podemos calcular as suas derivadas parciais. Calculemos, então, as derivadas parciais de $\frac{\partial f_1}{\partial x}$ e de $\frac{\partial f_1}{\partial y}$:

$$\frac{\partial^2 f_1}{\partial x^2}(x, y) := \frac{\partial}{\partial x}\Big(\frac{\partial f_1}{\partial x}\Big)(x, y) = \frac{\partial}{\partial x}\big(2x \cos(y)\big) = 2 \cos(y)$$

$$\frac{\partial^2 f_1}{\partial x \partial y}(x, y) := \frac{\partial}{\partial x}\Big(\frac{\partial f_1}{\partial y}\Big)(x, y) = \frac{\partial}{\partial x}\big(-x^2 \sin(y)\big) = -2x \sin(y)$$

$$\frac{\partial^2 f_1}{\partial y \partial x}(x, y) := \frac{\partial}{\partial y}\Big(\frac{\partial f_1}{\partial x}\Big)(x, y) = \frac{\partial}{\partial y}\big(2x \cos(y)\big) = -2x \sin(y)$$

$$\frac{\partial^2 f_1}{\partial y^2}(x, y) := \frac{\partial}{\partial y}\Big(\frac{\partial f_1}{\partial y}\Big)(x, y) = \frac{\partial}{\partial y}\big(-x^2 \sin(y)\big) = -x^2 \cos(y)$$

Note-se que $\frac{\partial^2 f_1}{\partial x \partial y}(x, y) = \frac{\partial^2 f_1}{\partial y \partial x}(x, y)$. Será que sempre acontece?

Proposição 6.7 (Schwartz) *Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}^p$ e $a \in D$. Se as derivadas parciais de 2a. ordem de f forem funções contínuas em a então*

$$\frac{\partial^2 f_i}{\partial x_j \partial x_k}(a) = \frac{\partial^2 f_i}{\partial x_k \partial x_j}(a)$$

Dem. Omitida ■

7 Aspectos geométricos da diferenciabilidade

Nesta secção abordaremos alguma geometria associada ao facto de uma função ser diferenciável num ponto. Começamos por estabelecer alguns resultados auxiliares.

7.1 Recta perpendicular a uma direcção e passando por um ponto; plano perpendicular a uma direcção e passando por um ponto.

7.1.1 Recta perpendicular a uma direcção e passando por um ponto.

Em $\mathbb{R}^2$, a recta perpendicular ao vector de coordenadas (v_1, v_2) e passando pelo ponto (x_1^0, x_2^0) é o conjunto de pontos de coordenadas (x_1, x_2) tais que os vectores

$$(x_1 - x_1^0, x_2 - x_2^0) \qquad \text{e} \qquad (v_1, v_2)$$

são ortogonais (ver Figura 9: Isto equivale a exigir que o produto interno destes dois vectores seja nulo:

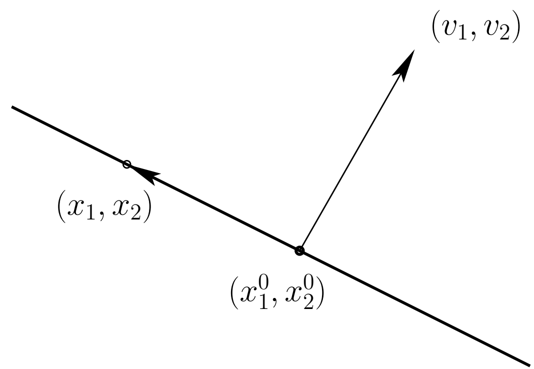

Figure 9: Recta perpendicular ao vector (v_1, v_2) e passando pelo ponto (x_1^0, x_2^0)

$$(x_1 - x_1^0, x_2 - x_2^0) \cdot (v_1, v_2) = 0$$

que podemos tambem escrever na forma:

$$(x_1 - x_1^0)v_1 + (x_2 - x_2^0)v_2 = 0$$

ou

$$x_2 = -\frac{v_1}{v_2}x_1 + x_2^0 + \frac{v_1}{v_2}x_1^0$$

7.1.2 Plano perpendicular a uma direcção e passando por um ponto.

Em $\mathbb{R}^3$, o plano perpendicular ao vector de coordenadas (v_1, v_2, v_3) e passando pelo ponto (x_1^0, x_2^0, x_3^0) é o conjunto de pontos de coordenadas (x_1, x_2, x_3) tais que os vectores

$$(x_1 - x_1^0, x_2 - x_2^0, x_3 - x_3^0) \qquad \text{e} \qquad (v_1, v_2, v_3)$$

são ortogonais (ver Figura 10. Isto equivale a exigir que o produto interno destes dois vectores seja nulo:

$$(x_1 - x_1^0, x_2 - x_2^0, x_3 - x_3^0) \cdot (v_1, v_2, v_3) = 0$$

que podemos tambem escrever na forma:

$$(x_1 - x_1^0)v_1 + (x_2 - x_2^0)v_2 + (x_3 - x_3^0)v_3 = 0$$

ou

$$x_3 = -\frac{v_1}{v_3}x_1 - \frac{v_2}{v_3}x_2 + x_3^0 + \frac{v_1}{v_3}x_1^0 + \frac{v_2}{v_3}x_2^0$$

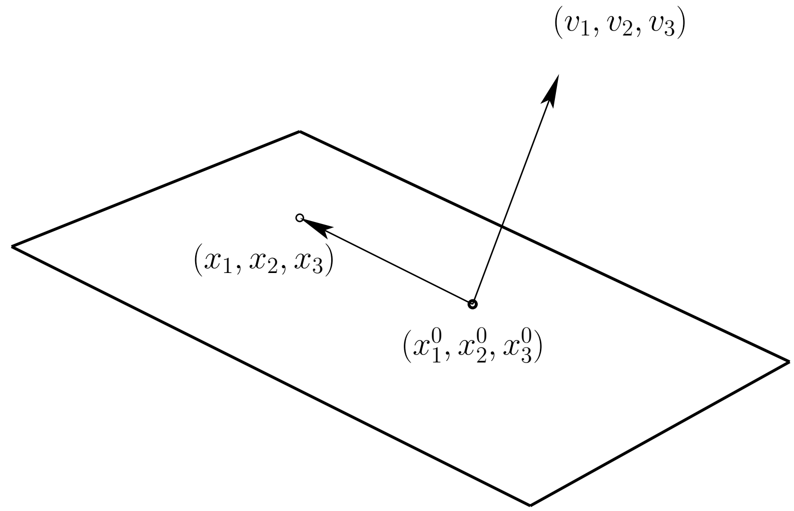

Figure 10: Plano perpendicular ao vector (v_1, v_2, v_3) e passando pelo ponto (x_1^0, x_2^0, x_3^0)

7.1.3 Hiperplano perpendicular a uma direcção e passando por um ponto.

Repetindo a abordagem atrás feita, fica claro que, em $\mathbb{R}^m$, se define o *hiperplano* ortogonal a uma direcção $v_1, v_2, \ldots, v_m$ e passando por um ponto $(x_1^0, x_2^0, \ldots, x_m^0)$ como o conjunto dos pontos $(x_1, x_2, \ldots x_m)$ tal que

$$(x_1 - x_1^0, x_2 - x_2^0, \ldots, x_m - x_m^0) \cdot (v_1, v_2, \ldots, v_m) = 0$$

que podemos tambem escrever na forma:

$$(x_1 - x_1^0)v_1 + (x_2 - x_2^0)v_2 + \cdots + (x_m - x_m^0)v_m = 0$$

ou

$$x_m = -\frac{v_1}{v_m}x_1 - \frac{v_2}{v_m}x_2 - \cdots - \frac{v_{m-1}}{v_m}x_{m-1} + x_m^0 + \frac{v_1}{v_m}x_1^0 + \frac{v_2}{v_m}x_2^0 + \cdots + \frac{v_{m-1}}{v_m}x_{m-1}^0$$

7.2 Funções diferenciáveis, gradiente e hipersuperfícies de nível

Seja f uma função real diferenciável num ponto a interior ao seu domínio. Qual a direcção em que nos devemos afastar de a de modo a, localmente, sentirmos a maior variação de f? Ou seja qual é o vector unitário v que maximiza o módulo da derivada direccional de f em a? Já que f é difereniciável em a tem-se:

$$|D_v f(a)| = |\nabla f(a) \cdot v| = ||\nabla f(a)||\,||v|||\cos\left(\widehat{\nabla f(a), v}\right)| = ||\nabla f(a)|||\cos\left(\widehat{\nabla f(a), v}\right)|$$

onde $\left(\widehat{\nabla f(a), v}\right)$ representa o menor ângulo entre $\nabla f(a)$ e v. Como o valor máximo de $|\cos\alpha|$ ocorre para $\alpha = 0 + 2k\pi$, então, maximizar esta derivada direccional equivale a tomar v unitário na direcção de gradiente de f em a

$$v = \frac{\nabla f(a)}{||\nabla f(a)||}$$

Por outro lado, em que direcção nos deveríamos afastar de a de modo a, localmente, não sentirmos variação de f? Desta vez, pretendemos anular a derivada direccional, donde,

$$0 = |D_v f(a)| = \cdots = ||\nabla f(a)|||\cos\left(\widehat{\nabla f(a), v}\right)|$$

donde, devemos tomar aqui, v ortogonal a $\nabla f(a)$.

Suponhamos então que a nossa função f é diferenciável num certo domínio, D, e que nesse domínio nós pretendemos encontrar os pontos a juntamente com a direcção e sentido v_a ao longo do qual nos devemos afastar de a de forma a, localmente, não sentirmos variação de f. A equação que pretendemos resolver aqui é

$$\cos\left(\widehat{\nabla f(a), v}\right) = 0$$

A resolução desta equação dá-nos, mais uma vez, os pontos a onde a função é diferenciável juntamente com o vector unitário v_a ao longo do qual nos devemos afastar de a de modo a não sentirmos, localmente, variação de f. Se o domínio de f está contido em $\mathbb{R}^2$, as linhas que unem esses pontos a e que, em cada um desses pontos a, são tangentes aos v_a chamam-se linhas de nível de a. A restrição de f a uma dessas linhas origina uma função constante. Se o domínio de f está contido em $\mathbb{R}^3$ obtemos as chamadas superfícies de nível. Se o domínio de f está contido em $\mathbb{R}^m$ obtemos as chamadas hipersuperfícies de nível.

Note-se que em cada uma destas hipersuperfícies de nível, o gradiente de f em a é ortogonal a esta hipersuperfície em a.

Exercício 7.1

Quais são as linhas de nível da função $f(x, y) = x^2 + y^2$?
E da função $g(x, y) = \frac{1}{1+x^2+y^2}$?

7.3 Hiperplano tangente e direcção ortogonal ao gráfico de uma função f num ponto $a, f(a)$

Comecemos por considerar uma função f definida num domínio D contido em $\mathbb{R}^2$ e com valores em $\mathbb{R}$. Suponhamos ainda que f é diferenciável num ponto (x_0, y_0) do interior de D. Na figura 11 esboçamos uma tal situação introduzindo tambem uma porção do plano tangente ao gráfico de f assim como um vector ortogonal ao gráfico de f no ponto $(x_0, y_0, f(x_0, y_0))$. O gráfico de f é formado pelos pontos (x, y, z) onde $z = f(x, y)$. Assim podemos considerar a função $g(x, y, z) = f(x, y) - z$ e desta maneira o gráfico de f passa a ser a superfície de nível da função g onde esta assume o valor constatnte zero. Então, a direcção ortogonal ao gráfico de f no ponto de $(x_0, y_0, f(x_0, y_0))$ é o gradiente da função g no ponto $(x_0, y_0, f(x_0, y_0))$:

$$\begin{aligned}\nabla g(x_0, y_0, f(x_0, y_0)) &= \left(\frac{\partial g}{\partial x}(x_0, y_0, f(x_0, y_0)), \frac{\partial g}{\partial y}(x_0, y_0, f(x_0, y_0)), \frac{\partial g}{\partial z}(x_0, y_0, f(x_0, y_0)),\right) = \\ &= \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1,\right)\end{aligned}$$

Assim, o plano tangente ao gráfico de f no ponto $(x_0, y_0, f(x_0, y_0))$ é,de acordo com os resultados de uma subsecção anterior,

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

Analogamente, se f está definida num domínio D contido em $\mathbb{R}^m$ e é diferenciável em $(x_1^0, x_2^0, \ldots, x_m^0)$ pertencente ao interior de D, então, a direcção ortogonal ao gráfico de f no ponto $(x_1^0, x_2^0, \ldots, x_m^0, f(x_1^0, x_2^0, \ldots, x_m^0))$ é dada por:

$$\left(\frac{\partial f}{\partial x_1}(x_1^0, x_2^0, \ldots, x_m^0), \frac{\partial f}{\partial x_2}(x_1^0, x_2^0, \ldots, x_m^0), \ldots, \frac{\partial f}{\partial x_2}(x_1^0, x_m^0, \ldots, x_m^0), -1,\right)$$

e o hiperplano tangente ao gráfico de f no ponto $(x_1^0, x_2^0, \ldots, x_m^0), f(x_1^0, x_2^0, \ldots, x_m^0))$ é dado pela expressão:

$$\begin{aligned}x_{m+1} = f(x_1^0, x_2^0, \ldots, x_m^0) + (x_1 - x_1^0)\frac{\partial f}{\partial x}(x_1^0, x_2^0, \ldots, x_m^0) + (x_2 - x_2^0)\frac{\partial f}{\partial y}(x_1^0, x_2^0, \ldots, x_m^0) + \ldots \\ \cdots + (x_m - x_m^0)\frac{\partial f}{\partial y}(x_1^0, x_2^0, \ldots, x_m^0)\end{aligned}$$

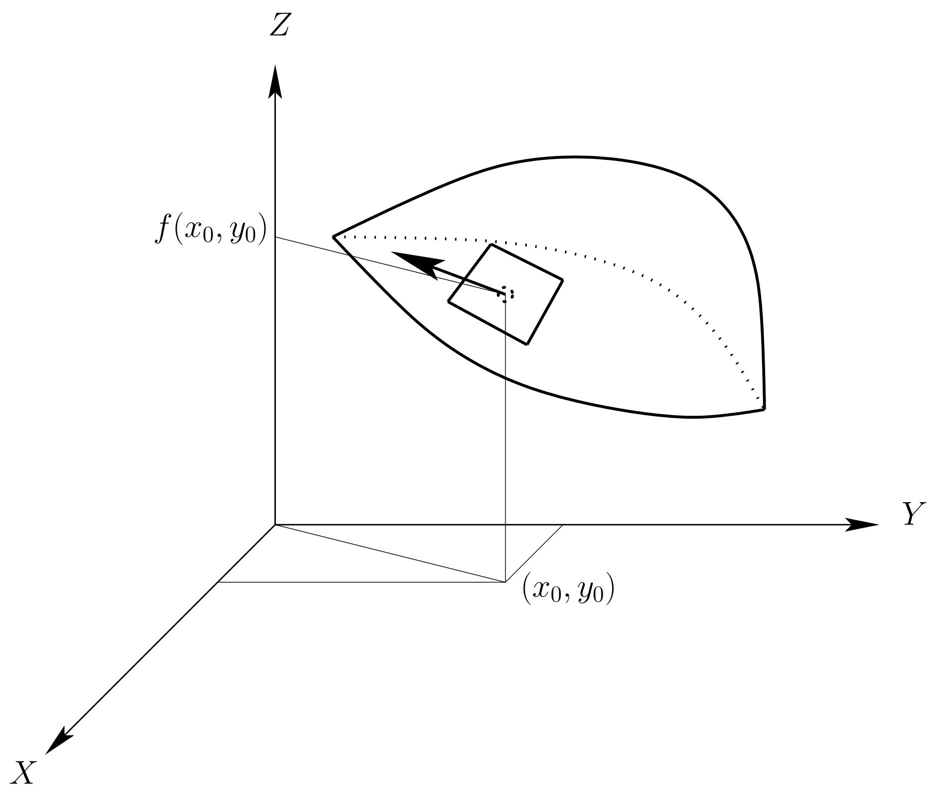

Figure 11: Plano tangente ao gráfico de f e vector ortogonal ao gráfico de f no ponto $(x_0, y_0, f(x_0, y_0))$

8 Fórmula de Taylor

Seja f uma função real, isto é, f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}$ e a um ponto interior a D onde existem todas as derivadas parciais de f até uma certa ordem l. Se $l = 2$ podemos escrever (porquê?):

$$f(a+h) = f(a) + \big(h \cdot \nabla\big) f(a) + o(||h||)$$

em que encaramos

$$\big(h \cdot \nabla\big) f(a)$$

como a função que resulta da operação $\big(h \cdot \nabla\big)$ em f, posteriormente calculada em a. Vejamos então mais explicitamente o que é essa operação de $\big(h \cdot \nabla\big)$ em f quando $m = 2$:

$$\big(h \cdot \nabla\big) f = \begin{pmatrix} h_1 & h_2 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{pmatrix} f = \Big(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y}\Big) = \Big(h_1 \frac{\partial f}{\partial x} + h_2 \frac{\partial f}{\partial y}\Big) =$$

e portanto

$$\big(h \cdot \nabla\big) f(a) = \Big(h_1 \frac{\partial f}{\partial x} + h_2 \frac{\partial f}{\partial y}\Big)(a) = h_1 \frac{\partial f}{\partial x}(a) + h_2 \frac{\partial f}{\partial y}(a)$$

Se $l = 3$ então obtemos

$$f(a+h) = f(a) + \big(h \cdot \nabla\big) f(a) + \frac{1}{2}\big(h \cdot \nabla\big)^2 f(a) + o(||h||^2)$$

onde

$$\big(h \cdot \nabla\big)^2 f = \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y}\right)^2 f = \Big(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y}\Big)\Big(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y}\Big) f =$$

$$= \Big(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y}\Big)\Big(h_1 \frac{\partial f}{\partial x} + h_2 \frac{\partial f}{\partial y}\Big) = h_1 \frac{\partial}{\partial x}\Big(h_1 \frac{\partial f}{\partial x} + h_2 \frac{\partial f}{\partial y}\Big) + h_2 \frac{\partial}{\partial y}\Big(h_1 \frac{\partial f}{\partial x} + h_2 \frac{\partial f}{\partial y}\Big) =$$

$$= h_1^2 \frac{\partial^2 f}{\partial x^2} + h_1 h_2 \frac{\partial^2 f}{\partial x \partial y} + h_2 h_1 \frac{\partial^2 f}{\partial y \partial x} + h_2^2 \frac{\partial^2 f}{\partial y^2} = \ldots$$

... já que h_1 e h_2 são constantes ...

$$\cdots = h_1^2 \frac{\partial^2 f}{\partial x^2} + 2h_1 h_2 \frac{\partial^2 f}{\partial x \partial y} + h_2^2 \frac{\partial^2 f}{\partial y^2}$$

usando Schwartz na última igualdade. Então:

$$\frac{1}{2}\big(h \cdot \nabla\big)^2 f(a) = \frac{1}{2} h_1^2 \frac{\partial^2 f}{\partial x^2}(a) + h_1 h_2 \frac{\partial^2 f}{\partial x \partial y}(a) + \frac{1}{2} h_2^2 \frac{\partial^2 f}{\partial y^2}(a)$$

Pode-se mostrar que se existem as derivadas parciais de f até à ordem $l+1$, então

$$f(a+h) = f(a) + \big(h \cdot \nabla\big) f(a) + \frac{1}{2}\big(h \cdot \nabla\big)^2 f(a) + \cdots + \frac{1}{l!}\big(h \cdot \nabla\big)^l f(a) + o(||h||^l)$$

8.1 Aplicação ao estudo de extremos

Definição 8.1

Seja f uma função real, isto é, f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}$. f tem um mínimo no ponto $a \in D$ se existir uma vizinhança de a, $B_\epsilon(a)$ tal que para todo o $x \in B_\epsilon(a)$, $f(x) \geq f(a)$.

Analogamente se diz que f tem máximo em $a \in D$ se existir uma vizinhança de a, $B_\epsilon(a)$ tal que para todo o $x \in B_\epsilon(a)$, $f(x) \leq f(a)$.

Mínimos e máximos de uma função designam-se por extremos dessa função.

Proposição 8.1 *Seja f definida em $D \subset \mathbb{R}^m$ e com valores em $\mathbb{R}$ e diferenciável num ponto $a \in D$. Se ocorre um extremo de f em a então, para todo o $v \neq (0, 0, \ldots, 0)$ a derivada de f segundo v em a é nula:*

$$D_v f(a) = 0$$

Dem. Dado $v \neq (0, 0, \ldots, 0)$, considere-se a função real de variável real dada por

$$g(t) = f(a + tv)$$

g é diferenciável (porquê?) em particular para $t = 0$. Porque f tem extremo em a, g tem extremo em $t = 0$. Então (cf. Análise Matemática I) a derivada de g em $t = 0$ é nula:

$$0 = g'(0) = \lim_{t \mapsto 0} \frac{g(t) - g(0)}{t - 0} = \lim_{t \mapsto 0} \frac{f(a + tv) - f(a)}{t - 0} = D_v f(a)$$

ou seja, para todo o $v \neq (0, 0, \ldots, 0)$

$$D_v f(a) = 0$$

■

Corolário 8.1 *Em particular,*

$$\nabla f(a) = (0, 0, \ldots, 0)$$

Dem. Omitida. ■

Definição 8.2

Seja f definida em $D \subset \mathbb{R}^m$, com valores em $\mathbb{R}$ e diferenciável num ponto $a \in D$. Se

$$\nabla f(a) = (0, 0, \ldots, 0)$$

então a diz-se um ponto de estacionariedade de f

Definição 8.3

Seja f definida em $D \subset \mathbb{R}^m$, com valores em $\mathbb{R}$ e diferenciável num ponto $a \in D$. Um ponto de estacionariedade de f, a, diz-se um ponto de sela se existir uma vizinhança de a, $B_\epsilon(a)$, e pelo menos dois pontos $x, y \in B_\epsilon(a)$ tais que $f(x) > f(a) > f(y)$.

Exemplo 8.1

1.
$$f(x,y) = x^2 + y^2$$

Esta função é claramente não negativa e assume o valor zero em $(0,0)$. Então f tem mínimo em $(0,0)$. Esse mínimo é ponto de estacionariedade já que f é diferenciável (porquê?) em $(0,0)$ e

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial x^2 + y^2}{\partial x}\Big|_{(0,0)} = 2x\Big|_{(0,0)} = 0$$

e

$$\frac{\partial f}{\partial y}(0,0) = \frac{\partial x^2 + y^2}{\partial y}\Big|_{(0,0)} = 2y\Big|_{(0,0)} = 0$$

2.
$$g(x,y) = xy$$

Esta função assume o valor zero em $(0,0)$, é positiva nos quadrantes ímpares e negativa nos quadrantes pares.

Para além disso

$$\frac{\partial g}{\partial x}(0,0) = \frac{\partial xy}{\partial x}\Big|_{(0,0)} = y\Big|_{(0,0)} = 0$$

e

$$\frac{\partial g}{\partial y}(0,0) = \frac{\partial xy}{\partial y}\Big|_{(0,0)} = x\Big|_{(0,0)} = 0$$

portanto g tem um ponto de sela em $(0,0)$.

Suponhamos agora que considerando uma função diferenciável, pretendemos localizar os seus extremos e pontos de sela. Os candidatos serão os pontos de estacionariedade. Haverá depois que distinguir quais desses pontos de estacionariedade são pontos de sela e quais são extremos e, de entre os extremos, decidir quais são os máximos e quais são os mínimos. Para isso fazemos uso da fórmula de Taylor (de 2a. ordem) aplicada ao ponto de estacionariedade a:

$$f(a+h) = f(a) + \big(h \cdot \nabla\big) f(a) + \frac{1}{2}\big(h \cdot \nabla\big)^2 f(a) + o(||h||^2)$$

Como o que nos interessa é perceber o sinal de $f(a+h) - f(a)$ reescrevemos:

$$f(a+h) - f(a) = \big(h \cdot \nabla\big) f(a) + \frac{1}{2}\big(h \cdot \nabla\big)^2 f(a) + o(||h||^2)$$

e como estamos a admitir que a é ponto de estacionariedade de f

$$f(a+h) - f(a) = \frac{1}{2}\big(h \cdot \nabla\big)^2 f(a) + o(||h||^2)$$

e como o $o(||h||^2)$ tende para zero mais depressa que $||h||^2$ quando $||h||$ tende para zero, o que interessa saber é como varia o sinal de

$$\big(h \cdot \nabla\big)^2 f(a) = h_1^2 \frac{\partial^2 f}{\partial x^2}(a) + 2h_1h_2 \frac{\partial^2 f}{\partial x \partial y}(a) + h_2^2 \frac{\partial^2 f}{\partial y^2}(a)$$

em função do h. Se este sinal for positivo, então $f(a+h) - f(a) > 0$ e portanto temos um mínimo em a; se for negativo, temos um máximo. Se para certos h o sinal é positivo enquanto que para outros o sinal é negativo, então temos um ponto de sela.

Se

$$h_1^2 \frac{\partial^2 f}{\partial x^2}(a) + 2h_1h_2 \frac{\partial^2 f}{\partial x \partial y}(a) + h_2^2 \frac{\partial^2 f}{\partial y^2}(a) = 0$$

então haverá que considerar mais termos na fórmula de Taylor para se chegar a alguma conclusão quanto à natureza do que ocorre no ponto a - não vamos considerar tais situações neste curso.

Consideremos novamente

$$h_1^2\frac{\partial^2 f}{\partial x^2}(a) + 2h_1h_2\frac{\partial^2 f}{\partial x\partial y}(a) + h_2^2\frac{\partial^2 f}{\partial y^2}(a)$$

e ponha-se h_2^2 em evidência:

$$h_2^2\Big[\frac{\partial^2 f}{\partial x^2}(a)\Big(\frac{h_1}{h_2}\Big)^2 + 2\frac{\partial^2 f}{\partial x\partial y}(a)\frac{h_1}{h_2} + \frac{\partial^2 f}{\partial y^2}(a)\Big]$$

donde obtemos uma função quadrática em $\frac{h_1}{h_2}$:

$$A\Big(\frac{h_1}{h_2}\Big)^2 + B\frac{h_1}{h_2} + C$$

com

$$A = \frac{\partial^2 f}{\partial x^2}(a), \qquad B = 2\frac{\partial^2 f}{\partial x\partial y}(a), \qquad C = \frac{\partial^2 f}{\partial y^2}(a)$$

e é o sinal desta função quadrática que importa agora conhecer. Para isso, queremos localizar os pontos onde esta função se anula. Estes são dados pela fórmula resolvente:

$$\frac{-B \pm \sqrt{B^2 - 4AC}}{2}$$

Então o que realmente queremos saber é o sinal de

$$B^2 - 4AC$$

ou ainda de

$$\frac{1}{4}\Big(B^2 - 4AC\Big) = \Big(\frac{\partial^2 f}{\partial x\partial y}(a)\Big)^2 - \frac{\partial^2 f}{\partial x^2}(a)\frac{\partial^2 f}{\partial y^2}(a)$$

Se o sinal de $\frac{1}{4}\Big(B^2 - 4AC\Big)$ for positivo então a função tanto assume valores positivos como negativos e portanto isso equival a dizer que, quanto à função f, ocorre um ponto de sela em a. Se este sinal for negativo quer dizer que **NÃO** há pontos onde a função se anula: esta é sempre positiva ou sempre negativa, o que corresponde a um extremo em a da função f; esse extremo será um máximo se $\frac{\partial^2 f}{\partial x^2}(a) < 0$ e mínimo se $\frac{\partial^2 f}{\partial x^2}(a) > 0$. $\frac{1}{4}\Big(B^2 - 4AC\Big)$ é o caso inconclusivo que nos referimos acima e que portanto não consideraremos.

Finalmente, se considerarmos a matriz

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x\partial y}(a) \\ \frac{\partial^2 f}{\partial x\partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}$$

obtem-se:

$$\det\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x\partial y}(a) \\ \frac{\partial^2 f}{\partial x\partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix} = -\Big(B^2 - 4AC\Big)$$

Então, em termos desta matriz, tem-se

$$\det\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x\partial y}(a) \\ \frac{\partial^2 f}{\partial x\partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix} < 0 \qquad \text{Ponto de sela em } a$$

$$\det\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x\partial y}(a) \\ \frac{\partial^2 f}{\partial x\partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix} > 0 \qquad \text{Extremo em } a\text{; máximo se } \frac{\partial^2 f}{\partial x^2}(a) < 0 \text{ e mínimo se } \frac{\partial^2 f}{\partial x^2}(a) > 0$$