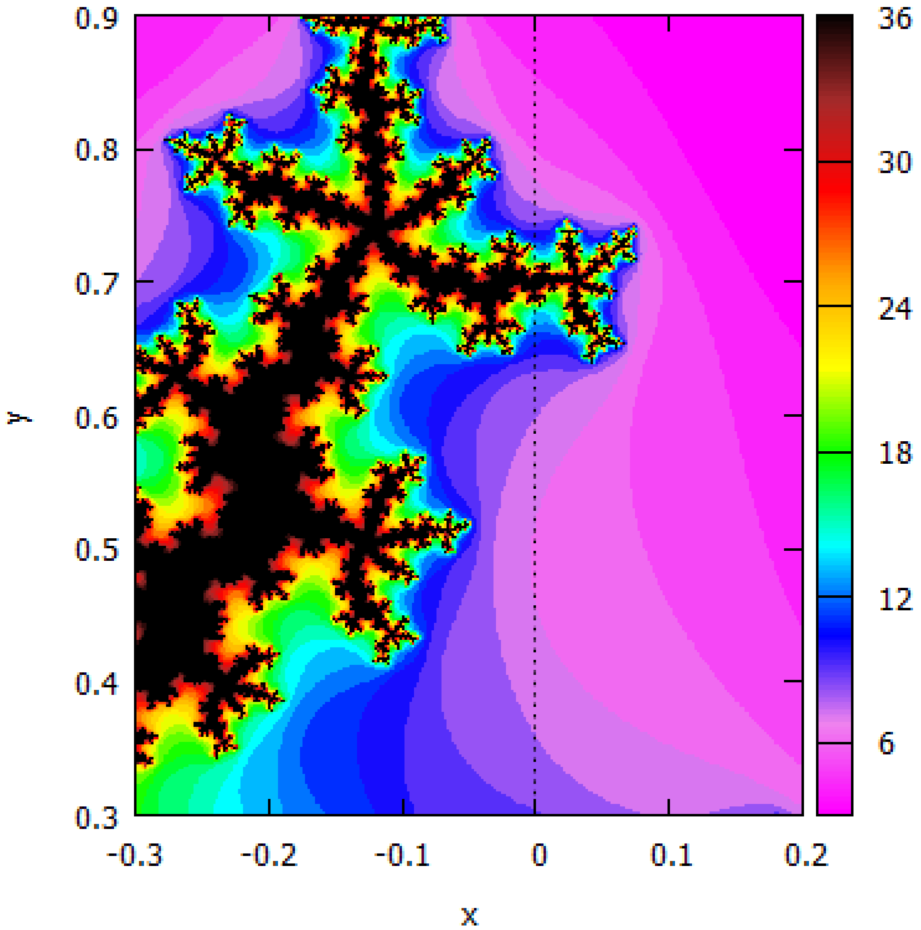
Apontamentos de

Análise Matemática II

2022/2023
Parte 1. Cálculo Diferencial em $\mathbb{R}^n$



Sara Fernandes, Clara Grácio, Luís Silva, Fátima Pereira, Telma Santos

Apontamentos de Análise Matemática II

2022/2023

Parte 1. Cálculo Diferencial em $\mathbb{R}^n$

Sara Fernandes, Clara Grácio, Luís Silva,

Fátima Pereira, Telma Santos

Departamento de Matemática

Escola de Ciências e Tecnologia

Universidade de Évora

.

Capa: Conjunto de Julia Software Maxima

Conteúdo

Capítulo 1

Introdução

Estas folhas constituem a base das aulas teóricas de Análise Matemática II da docente Sara Fernandes, no ano lectivo 2022/2023 e resultam de um trabalho colectivo acumulado de anos de experiência na leccionação desta unidade curricular por diversos docentes.

Não sendo um texto completo, mas apenas uma base de trabalho, aconselham-se os alunos a ter em conta a bibliografia recomendada, bem como a frequência das aulas teóricas e práticas.

O conhecimento teórico constante nestes apontamentos é um conhecimento básico e fundamental para as aplicações matemáticas que se consideram necessárias aos cursos a que a disciplina é ministrada, por isso, a tentativa usual de "fazer a cadeira"sem olhar para a sua parte teórica é, do nosso ponto de vista, desaconselhável por levar à ideia de que "em Matemática é só empinar fórmulas"ou "decorar exercícios resolvidos". Toda a matéria das aplicações/exercícios práticos tem fundamento teórico contido nestas folhas.

A par dos exercícios propostos no texto e no final de cada secção, será disponibilizada uma compilação de exercícios recomendados, que o aluno deverá ter em conta.

1.1 Preliminares

Neste capítulo introduziremos as noções básicas e notações que necessitaremos nos restantes capítulos. Alguns conceitos são já familiares ao aluno, no entanto, a linguagem utilizada procurará ser mais formal e rigorosa do que a utilizada durante o ensino secundário. Neste nível de ensino da Matemática procura-se um maior rigor de linguagem, necessidade nem sempre entendida pelo aluno. As definições, enunciados de teoremas e as suas demonstrações fazem parte da matéria avaliável, a par dos exercícios de aplicação. Não queremos com isto dizer que o aluno deverá decorar sem perceber, pelo contrário, fazemos questão que a matéria seja compreendida, mas entendemos que não é possível compreender plenamente resultados sem demonstração. Procuraremos ilustrar com exemplos toda a matéria, mas queremos deixar claro que exemplos são apenas exemplos, não substituem as provas rigorosas dos resultados. Procuraremos levar o aluno a raciocinar sobre objectos matemáticos, tentando sempre encontrar exemplos práticos que vão ao encontro dos interesses dos cursos em que se estão a formar. No entanto, quando as aplicações de determinados resultados sejam importantes para a formação dos alunos, mesmo que a demonstração destes saia do âmbito do programa, eles serão apresentados sem demontração e exigir-se-á apenas que o aluno os saiba aplicar.

Usaremos as seguintes notações habituais:

$\mathbb{N}$ Conjunto dos números naturais

$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$

$\mathbb{Z}$ Conjunto dos números inteiros relativos

$\mathbb{Q}$ Conjunto dos números racionais

$\mathbb{R}$ Conjunto dos números reais

$\mathbb{R}^+$ Conjunto dos números reais positivos

$\mathbb{R}_0^+$ Conjunto dos números reais não negativos

$\mathbb{C}$ Conjunto dos números complexos

$\wedge$ Operador lógico da conjunção, lê-se "e"

$\vee$ Operador lógico da disjunção, lê-se "ou"

$\backslash$ Operador diferença de conjuntos: $A \backslash B = \{x \in A : x \notin B\}$

A^C Complementar do conjunto A. Se $A \subset \mathbb{R}$, $A^C = \mathbb{R}\backslash A$

(a, b) ou $]a, b[$ Intervalo aberto dos números reais maiores que a e menores que b

$[a, b]$ Intervalo fechado dos números reais maiores ou iguais a a e menores ou iguais a b

Vamos começar com a introdução de alguns elementos da teoria de conjuntos que o aluno certamente reconhecerá.

Definição 1.1 *Sejam A e B dois conjuntos. Definimos o* ***produto cartesiano*** *de A por B da seguinte forma*

$$A \times B = \{(a, b) : a \in A \wedge b \in B\}.$$

Ou seja são os pares ordenados em que o primeiro termo pertence ao primeiro conjunto e o segundo termo pertence ao segundo conjunto.

Exemplo:

a) $A = \{1, 2, 3\}$, $B = \{4, 5\}$, $A \times B = \{(1, 4), (1, 5), (2, 4), (2, 5), (3, 4), (3, 5)\}$;

b) $A = \{\boxplus, \oslash\}$, $B = \{\heartsuit, \clubsuit\}$, $A \times B = \{(\boxplus, \clubsuit), (\boxplus, \heartsuit), (\oslash, \heartsuit), (\oslash, \clubsuit)\}$;

c) $A = \mathbb{R}$, $B = \mathbb{R}$, $A \times B = \{(a, b) : a \in \mathbb{R} \wedge b \in \mathbb{R}\}$ é o conjunto dos pares ordenados de números reais. Este conjunto especial, $\mathbb{R} \times \mathbb{R}$, denota-se $\mathbb{R}^2$.

Usando o sistema de eixos cartesiano, bem conhecido do aluno, obtemos uma visão geométrica do conjunto $\mathbb{R}^2$ como um plano, cujos pontos se identificam com as coordenadas das projecções nos eixos dos xx e dos yy.

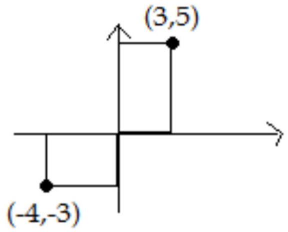


Figura 1.1: Sistema cartesiano em dimensão 2.

Podemos alargar a noção de produto cartesiano a mais do que dois conjuntos, assim, por exemplo, para A, B, C conjuntos, define-se

$$A \times B \times C = \{(a, b, c) : a \in A \wedge b \in B \wedge c \in C\}.$$

Aqui também assinalamos o caso especial em que $A = B = C = \mathbb{R}$, o conjunto dos trios ordenados de números reais, que denotaremos $\mathbb{R}^3$. Neste caso, fixado um sistema de eixos cartesiano os pontos do espaço são identificados pelas coordenadas das suas projecções nos eixos dos xx, dos yy e dos zz.

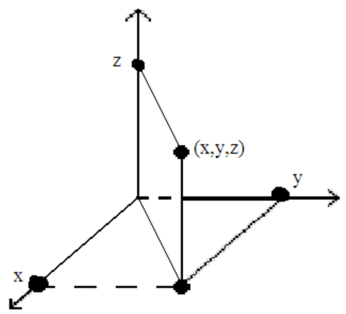


Figura 1.2: Sistema cartesiano em dimensão 3.

Quem define para 2 ou 3 conjuntos também pode definir o produto cartesiano para um número $n \in \mathbb{N}$ de conjuntos

$$A_1 \times A_2 \times ...A_n = \{(a_1, a_2, ..., a_n) : a_1 \in A_1 \wedge a_2 \in A_2 \wedge ... \wedge a_n \in A_n\}.$$

Igualmente aqui $\mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R}$ será denotado por $\mathbb{R}^n$. Apesar de não conseguirmos visualizar objectos em espaços de dimensão superior a 3 o aluno está familiarizado com a noção subjacente a esta visualização, que são as bases ortonormadas dos espaços vectoriais reais, $\mathbb{R}^n$, estudadas em Álgebra Linear no primeiro semestre.

Definição 1.2 *Seja V um conjunto não vazio e seja $+$ uma operação binária em V e $\cdot$ uma operação de $\mathbb{R} \times V$ em V. Dizemos que $(V, \mathbb{R}, +, \cdot)$ é um* ***espaço vectorial*** *sobre o corpo $\mathbb{R}$ se e só se se verificarem as seguintes propriedades:*

P1. Associatividade da operação $+$;

P2. Comutatividade da operação $+$;

P3. Existência de elemento neutro para a operação $+$;

P4. Existência de simétrico.

Além disso, para a e b números reais, $\mathbf{v}_1$ e $\mathbf{v}_2$ elementos de V

Q1. $a(b\mathbf{v}_1) = (ab)\,\mathbf{v}_1$;

Q2. $(a+b)\,\mathbf{v}_1 = a\mathbf{v}_1 + b\mathbf{v}_1$;

Q3. $a\,(\mathbf{v}_1 + \mathbf{v}_2) = a\mathbf{v}_1 + a\mathbf{v}_2$.

Q4. $1{\cdot}\mathbf{v}_1 = \mathbf{v}_1$.

Usaremos letras minúsculas a negrito para denotar vectores (elementos de V) e letras minúsculas sem negrito para denotar os escalares reais.

Evidentemente, o conjunto $\mathbb{R}^n$ munido das operações

$$\begin{array}{lll} + & : & (x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)\,, \\ \cdot & : & a \cdot (x_1, x_2, ..., x_n) = (ax_1, ax_2, ..., ax_n)\,, \end{array}$$

é um espaço vectorial real, onde o elemento neutro é o vector $\mathbf{0} = (0, 0, ..., 0)$ ao qual chamaremos origem ou vector nulo e o simétrico do vector $\mathbf{v} = (x_1, x_2, ..., x_n)$ é o vector $-\mathbf{v} = (-x_1, -x_2, ..., -x_n)$.

Vamos supor que o aluno está familiarizado com estes conceitos e terminologia.

Exercícios 1.1.

a) Esboce no sistema de eixos cartesiano os seguintes vectores de $\mathbb{R}^2$

$$\begin{aligned} \mathbf{x} &= (-1, 2)\,; \\ \mathbf{y} &= (0, 5); \\ \mathbf{z} &= (-5, 0); \\ \mathbf{w} &= (0, 0). \end{aligned}$$

b) Indique o simétrico de

$$\begin{aligned} (2, -2, 4, -4) &\in \mathbb{R}^4; \\ (1, 0, 2) &\in \mathbb{R}^3; \\ (0, 0) &\in \mathbb{R}^2. \end{aligned}$$

c) Resolva as seguintes equações em $\mathbb{R}^n$:

$$\begin{aligned} 5\,(x_1, x_2, x_3, x_4) + (7, 2, 3, -2) &= (0, -2, 6, 1)\,; \\ -2\,(x_1, x_2, x_3) + 2\,(1, 2, 3) &= 3\,(7, -2, -4)\,; \\ -\,(x_1, x_2) + 2\,(\pi, 2\pi) &= 3\pi\,(7, -2)\,. \end{aligned}$$

Não vamos repetir o estudo dos espaços vectoriais $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$ que foi feito em Álgebra Linear, o nosso ponto de vista é outro, estes espaços serão vistos do ponto de vista métrico e topológico e serão os conjuntos de partida e de chegada das funções que vamos estudar.

O primeiro conceito "novo"de que vamos necessitar é a noção de distância. Pusemos a palavra novo entre aspas, porque evidentemente o aluno sabe muito bem o que é uma distância entre pontos do espaço em que vivemos. Mas saberá o que é a distância entre elementos de $\mathbb{R}^5$? Qual a distânca entre $(1,2,3,4,5)$ e $(5,4,3,2,1)$? Recordemos o que sabemos. Quando se trata de números reais, falamos de distância entre eles, querendo dizer "diferença", ou melhor, "módulo da diferença", $d(x,y) = |x-y|$. A distância entre 1 e 5 é, claro está, $|1-5| = |-4| = 4$. A distância entre -2 e 4 é $|-2-4| = |-6| = 6$. Já a distância entre dois pontos/vectores do plano, identificados pelas coordenadas cartesianas, é a distância euclidiana, $d(\mathbf{v}_1, \mathbf{v}_2)$ ou $\|\mathbf{v}_1 - \mathbf{v}_2\|_2$, definida da seguinte forma:

$$\begin{aligned} \text{Para } \mathbf{v}_1 &= (x_1, x_2) \text{ e } \mathbf{v}_2 = (y_1, y_2) \\ d(\mathbf{v}_1, \mathbf{v}_2) &= \|\mathbf{v}_1 - \mathbf{v}_2\|_2 = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}. \end{aligned}$$

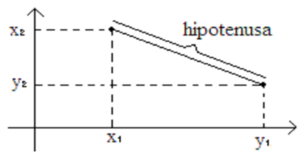


Figura 1.3: Distância

E também para pontos/vectores do espaço:

$$\begin{aligned} \text{Para } \mathbf{v}_1 &= (x_1, x_2, x_3) \text{ e } \mathbf{v}_2 = (y_1, y_2, y_3) \\ d(\mathbf{v}_1, \mathbf{v}_2) &= \|\mathbf{v}_1 - \mathbf{v}_2\|_3 = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2}. \end{aligned}$$

De uma forma geral, definimos a **distância euclidiana** entre dois vectores de $\mathbb{R}^n$ como se segue:

$$\begin{aligned} \text{Para } \mathbf{v}_1 &= (x_1, x_2, ..., x_n) \text{ e } \mathbf{v}_2 = (y_1, y_2, ..., y_n) \\ d(\mathbf{v}_1, \mathbf{v}_2) &= \|\mathbf{v}_1 - \mathbf{v}_2\|_n = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + ... + (x_n-y_n)^2}. \end{aligned}$$

Chamamos norma de um vector à sua distância à origem

$$\|\mathbf{v}\|_n = d(\mathbf{v}, \mathbf{0}) = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}.$$

Esta norma pode ser obtida através do produto interno (também estudado em Álgebra Linear no semestre passado), cuja definição recordamos brevemente.

Definição 1.3 *Sejam* $\mathbf{v}_1 = (x_1, x_2, ..., x_n)$ *e* $\mathbf{v}_2 = (y_1, y_2, ..., y_n)$ *dois vectores de* $\mathbb{R}^n$. *Definimos* ***produto interno*** *ou* ***produto escalar*** *de* $\mathbf{v}_1$ *e* $\mathbf{v}_2$ *por*

$$(\mathbf{v}_1, \mathbf{v}_2)_n = x_1 y_1 + x_2 y_2 + ... + x_n y_n.$$

Usam-se também outras notações para o produto interno, como $\mathbf{v}_1 \cdot \mathbf{v}_2$, $\mathbf{v}_1 | \mathbf{v}_2$ ou ainda $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle$.

Exercício: Mostre que

1. $(\mathbf{v}_1, \mathbf{v}_2) = (\mathbf{v}_2, \mathbf{v}_1)$;

2. $((\alpha \mathbf{v}_1 + \beta \mathbf{v}_2), \mathbf{v}_3) = \alpha(\mathbf{v}_1, \mathbf{v}_3) + \beta(\mathbf{v}_2, \mathbf{v}_3)$

Definido o produto interno, facilmente se prova que a norma $\|.\|_n$ verifica

$$\|\mathbf{v}_1\|_n^2 = (\mathbf{v}_1, \mathbf{v}_1)_n \,.$$

Nota: Poderíamos estudar outro tipo de distâncias/normas ou mesmo a noção generalizada de distância/norma e fazer um estudo mais geral sobre este conceito teórico. Infelizmente o número de horas semanais limitam esta possibilidade, estreitando-nos o ponto de vista. Fiquemo-nos então apenas com a distância/norma euclidianas.

Facilmente se verificam as seguintes propriedades:

Teorema 1.1 *Sejam* $\mathbf{x}$ *e* $\mathbf{y}$ *elementos de* $\mathbb{R}^n$ *e seja* $c \in \mathbb{R}$ *um escalar. Então*

1. $\|\mathbf{x}\|_n = 0$ *sse* $\mathbf{x} = \mathbf{0}$.

2. $\|\mathbf{x}\|_n > 0$ *sse* $\mathbf{x} \neq \mathbf{0}$ *(positividade).*

3. $\|c\, \mathbf{x}\|_n = |c| \;\; \|\mathbf{x}\|_n$ *(homogeneidade).*

4. $|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|_n \|\mathbf{y}\|_n$ *(desigualdade de Cauchy-Schwarz).*

5. $\|\mathbf{x} + \mathbf{y}\|_n \leq \|\mathbf{x}\|_n + \|\mathbf{y}\|_n$ *(desigualdade triangular).*

Demonstração:

1 e 2. são evidentes da definição de norma.

3. Resulta das propriedades das funções $\sqrt{x}$ e x^2.

4. Se $\mathbf{x} = \mathbf{0}$ a desigualdade verifica-se. Supomos que $\mathbf{x} \neq \mathbf{0}$. Então $(\mathbf{x} \cdot \mathbf{x}) \neq 0$ e podemos definir $c = \frac{(\mathbf{x}, \mathbf{y})}{(\mathbf{x}, \mathbf{x})}$. Da definição de produto interno obtemos

$$0 \leq (\mathbf{y} - c\,\mathbf{x}, \mathbf{y} - c\,\mathbf{x}) = (\mathbf{y}, \mathbf{y} - c\,\mathbf{x}) - c\,(\mathbf{x}, \mathbf{y} - c\,\mathbf{x}) = (\mathbf{y}\ ,\ \mathbf{y} - c\,\mathbf{x})$$

porque $c = \frac{(\mathbf{x}, \mathbf{y})}{(\mathbf{x}, \mathbf{x})}$ e logo $(\mathbf{x},\ \mathbf{y} - c\mathbf{x}) = (\mathbf{x}, \mathbf{y}) - c\,(\mathbf{x}, \mathbf{x}) = 0$.

Obtemos então

$$0 \leq (\mathbf{y}\ , \mathbf{y} - c\,\mathbf{x}) = (\mathbf{y}\ ,\ \mathbf{y}) - c\,(\mathbf{x}, \mathbf{y})$$

Somando $c\,(\mathbf{x}, \mathbf{y})$ a ambos os membros obtemos

$$c\,(\mathbf{x}, \mathbf{y}) \leq (\mathbf{y}, \mathbf{y})$$

Finalmente, multiplicando ambos os membros da desigualdade por $(\mathbf{x}, \mathbf{x})$ obtemos

$$(\mathbf{x}, \mathbf{y})^2 \leq (\mathbf{y}, \mathbf{y})\ \ (\mathbf{x}, \mathbf{x}) = \|\mathbf{y}\|_n^2 \|\mathbf{x}\|_n^2$$

donde se pode concluir a afirmação.

5.

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

Tomando a raiz quadrada obtemos a desigualdade pretendida.□

Conhecendo a noção de distância, estamos em condições de definir vizinhanças de vectores como generalização de intervalos reais centrados em dado número real x e de raio δ: $B\,(x_0, \delta) = \{x \in \mathbb{R} : |x - x_0| < \delta\} =]x_0 - \delta, x_0 + \delta[$.

Definição 1.4 *Chamamos* ***vizinhança*** *de* x_0 *ou* ***bola*** *centrada em* $\mathbf{x}_0 \in \mathbb{R}^n$ *e raio* $\delta \in \mathbb{R}^+$ *ao conjunto*

$$\begin{aligned} B(x_0, \delta) &= \{\mathbf{x} \in \mathbb{R}^n : \|x - x_0\|_n < \delta\} = \\ &= \{\mathbf{x} \in \mathbb{R}^n : \sqrt{(x_1 - x_{01})^2 + (x_2 - x_{02})^2 + ... + (x_n - x_{0n})^2} < \delta\}. \end{aligned}$$

Exercícios 1.2.

a) Esboce o conjunto dos pontos de $\mathbb{R}^2$ que pertence à bola centrada em $(2, 3)$ e raio 2.

b) Esboce os seguintes subconjuntos de $\mathbb{R}^2$

$$\begin{aligned} A &= \{\mathbf{v} \in \mathbb{R}^2 : 1 < \|\mathbf{v}\| < 3\}; \\ B &= \{\mathbf{v} \in \mathbb{R}^2 : \|\mathbf{v}\| \leq 5\}; \\ C &= \{\mathbf{v} \in \mathbb{R}^2 : \|\mathbf{v}\| \geq 1\}; \\ D &= \{(x, y) \in \mathbb{R}^2 : 1 < x < 3 \wedge -2 < y < 4\}; \\ E &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x + y < 5\}; \\ F &= \{(x, y) \in \mathbb{R}^2 : \operatorname{sen} x < y < 1\}. \end{aligned}$$

c) Que lugar geométrico representa uma vizinhança de um ponto em $\mathbb{R}^3$?

d) Que lugar geométrico representa o conjunto seguinte?

$$X = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 < 4 \wedge 0 < z < 5\}$$

e) Tente descrever analiticamente uma lua em quarto crescente (no plano).

Ao contrário do que possa parecer, a noção de distância é extremamente importante na Análise Matemática que, nunca é demais realçar, está assente fundamentalmente na noção de limite. Assim, para definir limite em Análise I o aluno necessitou de alguns conceitos topológicos que vieram da noção de distância em $\mathbb{R}$. Também em Análise II, antes da definição de limite vamos necessitar da introdução desses conceitos.

Em todas as definições que se seguem consideramos A um subconjunto de $\mathbb{R}^n$, $\mathbf{x}$ um elemento de $\mathbb{R}^n$.

Definição 1.5 *Dizemos que x é um **ponto interior** a A sse existe uma bola centrada em x e raio δ, $B(x,\delta)$, tal que $B(x,\delta) \subset A$.*

*Dizemos que x é um **ponto exterior** a A sse existe uma bola centrada em x e raio δ, $B(x,\delta)$, tal que $B(x,\delta) \cap A = \varnothing$.*

*Dizemos que x é um **ponto fronteira** de A sse qualquer que seja a bola centrada em x e raio δ, $B(x,\delta)$, se tem $B(x,\delta) \cap A \neq \varnothing$ e $B(x,\delta) \cap A^C \neq \varnothing$.*

*Chamamos **interior** de A ao conjunto dos seus pontos interiores, denotamo-lo $Int(A)$.*

*Chamamos **exterior** de A ao conjunto dos seus pontos exteriores, denotamo-lo $Ext(A)$.*

*Chamamos **fronteira** de A ao conjunto dos seus pontos fonteira, denotamo-lo $fr(A)$.*

*Chamamos **fecho** ou **aderência** de A ao conjunto A unido com a sua fronteira, $\bar{A} = A \cup fr(A)$. Denotamo-lo $\bar{A}$.*

Teorema 1.2 *Qualquer que seja $A \subset \mathbb{R}^n$ tem-se*

$$Int(A) \cup Ext(A) \cup fr(A) = \mathbb{R}^n,$$

sendo os conjuntos $Int(A)$, $Ext(A)$ e $fr(A)$ disjuntos dois a dois.

Definição 1.6 *Dizemos que x é um **ponto de acumulação** de A sse qualquer que seja a bola centrada em x e raio δ, $B(x,\delta)$, se tem $(B(x,\delta) \setminus \{x\}) \cap A \neq \varnothing$.*

*Chamamos **derivado** de A ao conjunto dos seus pontos de acumulação, denotamo-lo A'.*

*Dizemos que x é um **ponto isolado** de A sse for um ponto de A e não pertencer ao derivado A'.*

Definição 1.7 *Dizemos que o conjunto A é **aberto** sse $Int(A) = A$.*

*Dizemos que um conjunto A é **fechado** sse A^C for aberto.*

Resulta da definição que um conjunto é fechado sse $A = \bar{A}$.

Nota: Um conjunto pode não ser aberto nem fechado, por exemplo,

$$A = \{(x,y) \in \mathbb{R}^2 : 0 < x < 4 \wedge 0 \leq y \leq 5\}.$$

Por outro lado $\mathbb{R}^n$ e o conjunto vazio, $\varnothing$, são abertos e fechados, em simultâneo.

Definição 1.8 *Dizemos que o conjunto $A \subset \mathbb{R}^n$ é **limitado** sse existir $M > 0$ tal que $\|\mathbf{x}\|_n < M$ para todo o $\mathbf{x} \in A$*

Antes de iniciarmos o estudo das funções vectoriais, objecto principal deste semestre, recapitulemos as propriedades básicas de algumas funções reais de variável real, que vamos supôr que o aluno domina.

Em todos os exemplos que se seguem, tomemos $a, b, c \in \mathbb{R},\ a \neq 0$ e $y = f(x)$.

Exemplos

a) Função afim com parâmetros a e b reais:

$y = ax + b;$

O domínio é $D(y) = \mathbb{R};$

$y' = a;$

$\int y dx = a\frac{x^2}{2} + bx + c;$

É bijectiva e tem inversa $x = \frac{y-b}{a}$.

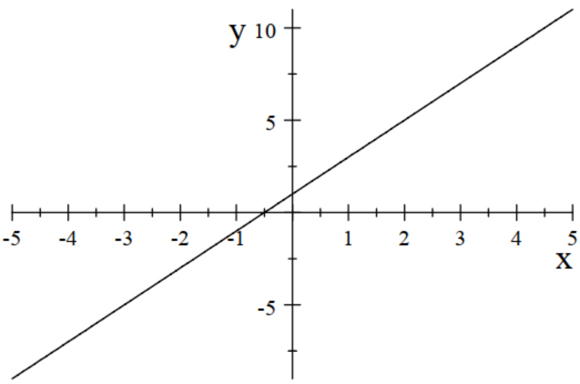


Figura 1.4: Parte do gráfico de $2x + 1$

b) Função quadrática simples:

$y = x^2;$

O domínio é $D(y) = \mathbb{R};$

$y' = 2x;$

$\int y\ dx = \frac{x^3}{3} + c;$

Não é bijectiva no seu domínio, mas a sua restrição a $\mathbb{R}^+$ é e tem inversa $x = \sqrt{y}$.

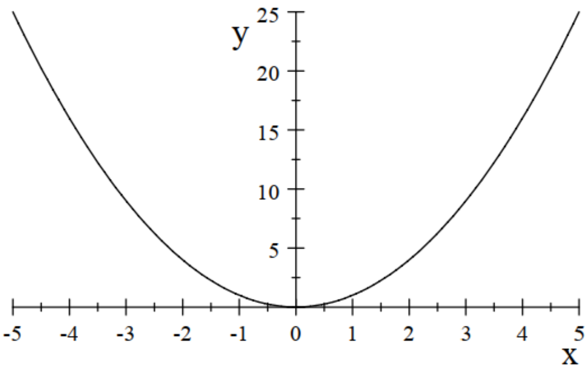


Figura 1.5: Parte do gráfico de x^2

c) Função seno:

$y = \operatorname{sen} x$;

O domínio é $D(y) = \mathbb{R}$;

$y' = \cos x$;

$\int y dx = -\cos x + c$;

Não é bijectiva no seu domínio, mas a sua restrição a $[-\frac{\pi}{2}, \frac{\pi}{2}]$ é e tem inversa $x = \operatorname{arcsen} y$ com $y \in [-1, 1]$.

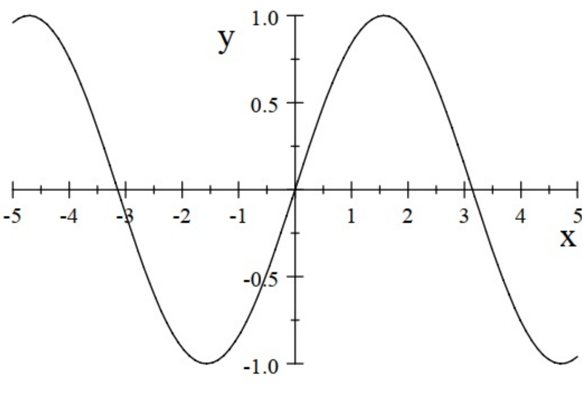


Figura 1.6: Parte do gráfico de sen x

d) Função arcseno:

$y = \operatorname{arcsen} x$;

O domínio é $D\,(y) = [-1, 1]$;

$y' = \frac{1}{\sqrt{1-x^2}} \quad (|x| < 1)\,;$

É bijectiva no seu domínio, e tem inversa, $x = \operatorname{sen} y$, com $y \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

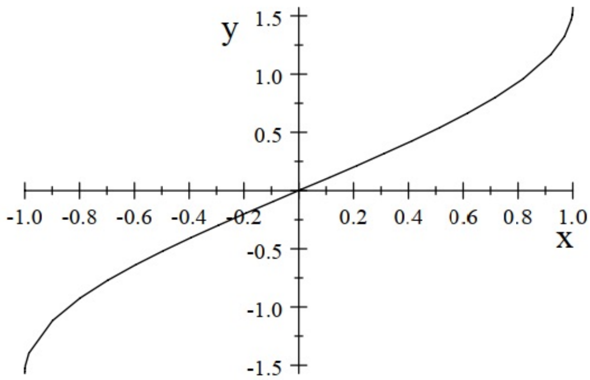


Figura 1.7: Gráfico de arcsen x

e) Função coseno:

$y = \cos x$;

O domínio é $D\,(y) = \mathbb{R}$;

$y' = -\operatorname{sen} x$;

$\int y\ dx = \operatorname{sen} x + c$;

Não é bijectiva no seu domínio, mas a sua restrição a $[0, \pi]$ é e tem inversa $x = \arccos y$ com $y \in [-1, 1]$.

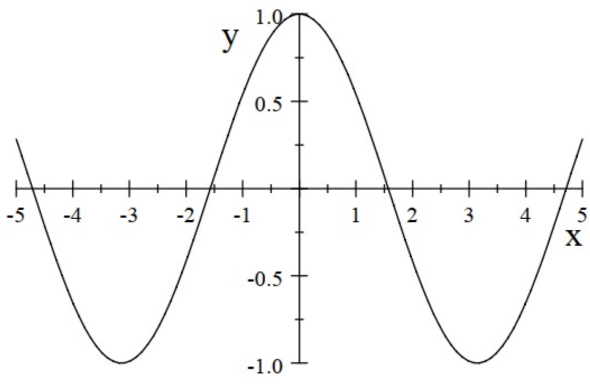


Figura 1.8: Parte do gráfico de cos x

f) Função arccoseno:

$y = \arccos x$;

O domínio é $D(y) = [-1, 1]$;

$y' = -\frac{1}{\sqrt{1-x^2}} \quad (|x| < 1)$;

É bijectiva no seu domínio e tem inversa

$x = \cos y,$ com $y \in [0, \pi]$.

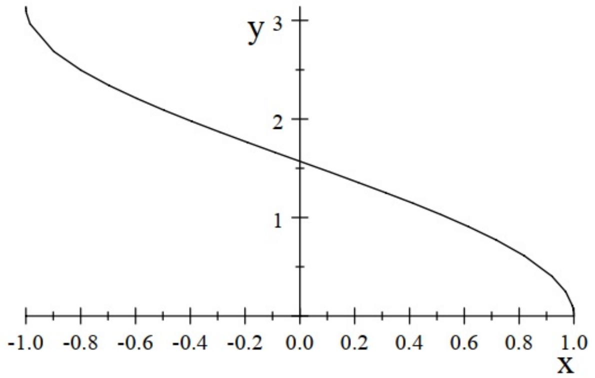


Figura 1.9: Gráfico de arccos x

g) Função tangente:

$y = \operatorname{tg} x$;

O domínio é $D(y) = \mathbb{R} \backslash \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$;

$y' = \frac{1}{\cos^2 x}$;

Não é bijectiva no seu domínio, mas a sua restrição a $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ é

e tem inversa $x = \operatorname{arctg} y$ com $y \in \mathbb{R}$.

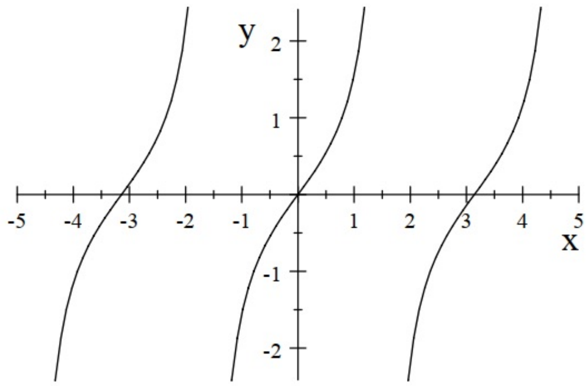


Figura 1.10: Parte do gráfico de tan x

g) Função arctangente:

$y = \operatorname{arctg} x$;

O domínio é $D(y) = \mathbb{R}$;

$y' = \frac{1}{1+x^2}$;

É bijectiva e tem inversa $x = \operatorname{tg} y,$ com $y \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

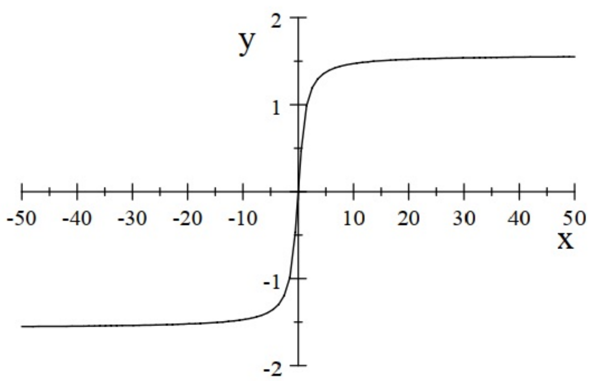


Figura 1.11: Parte do gráfico de arctg x

h) Função exponencial:

$y = e^x$;

O domínio é $D(y) = \mathbb{R}$;

$y' = e^x$;

$\int y dx = e^x + c$;

É bijectiva e tem inversa $x = \ln y$, com $y \in \mathbb{R}^+$.

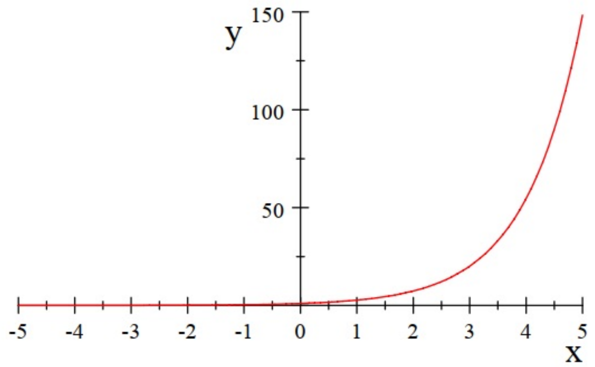


Figura 1.12: Parte do gráfico de e^x

i)Função logaritmo de base e:

$y = \ln x$;

O domínio é $D(y) = \mathbb{R}^+$;

$y' = \frac{1}{x}$;

É bijectiva e tem inversa $x = e^y$, com $y \in \mathbb{R}$.

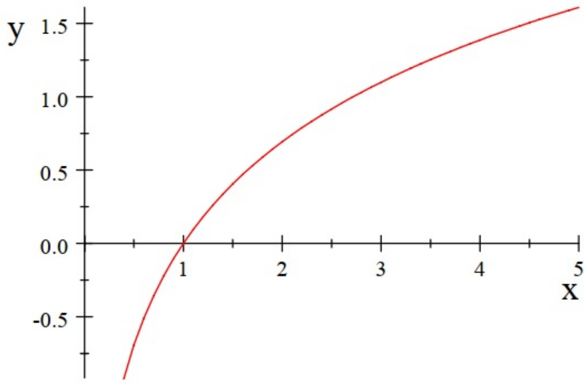


Figura 1.13: Parte do gráfico de log x

Capítulo 2

Funções de $\mathbb{R}^n$ em $\mathbb{R}^m$

2.1 Exemplos, domínios e gráficos

Vamos estudar, neste capítulo, uma família de funções, cujo domínio e conjunto de chegada são subconjuntos dos espaços euclideanos $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente, são as chamadas funções vectoriais. Comecemos por pensar no que será uma função de $\mathbb{R}^3$ em $\mathbb{R}$. Terá que ser uma função f que, a cada vector $(x_1, x_2, x_3) \in \mathbb{R}^3$, associa um número real $f(x_1, x_2, x_3)$, por exemplo

$$
\begin{aligned}
&f : \mathbb{R}^3 \to \mathbb{R} \\
&f(x_1, x_2, x_3) = x_1 + 2x_2 - 8x_3.
\end{aligned}
$$

Agora, pensemos no que será uma função de $\mathbb{R}$ em $\mathbb{R}^3$. Terá que ser uma regra que, a cada número real x, associe um vector com três coordenadas (funções de argumento x), por exemplo

$$
\begin{aligned}
&f : \mathbb{R} \to \mathbb{R}^3 \\
&f(x) = (x, 2x, 8x).
\end{aligned}
$$

Como se pode observar no exemplo, cada componente do vector que define a função é, ela própria, uma função de $\mathbb{R}$ em $\mathbb{R}$. Não é assim só no exemplo, de uma forma geral, uma função com valores em $\mathbb{R}^m$, terá sempre m componentes, sendo cada uma delas uma função com valores em $\mathbb{R}$. Temos então

$$
\begin{aligned}
&f : \mathbb{R}^n \to \mathbb{R}^m \\
&f(x_1, x_2, ..., x_n) = (f_1(x_1, x_2, ..., x_n), f_2(x_1, x_2, ..., x_n), ..., f_m(x_1, x_2, ..., x_n)).
\end{aligned}
$$

O **domínio de definição** de uma função vectorial, $f : \mathbb{R}^n \to \mathbb{R}^m$, será a intersecção dos domínios de definição das suas componentes, vistos como subconjuntos de $\mathbb{R}^n$.

A **composta** de duas funções vectoriais $f : D_f \subset \mathbb{R}^k \to \mathbb{R}^m$ e $g : D_g \subset \mathbb{R}^n \to \mathbb{R}^k$, tais que $g(D_g) \subset D_f$ é a função $h = f \circ g$, (lê-se f após g) com $h : D_g \subset \mathbb{R}^n \to \mathbb{R}^m$, definida por $h(x) = f(g(x))$.

Exemplos:

a) $f : D \subset \mathbb{R} \to \mathbb{R}^2$, $f(x) = (x, x+3)$;

b) $g : D \subset \mathbb{R}^3 \to \mathbb{R}^2$, $g(x_1, x_2, x_3) = (\text{sen}(x_1), x_2 + 5x_3 + 2)$;

c) $h : D \subset \mathbb{R}^2 \to \mathbb{R}$, $h(x_1, x_2) = x_1 - 2e^{x_2}$;

d) $g_1 : D \subset \mathbb{R}^4 \to \mathbb{R}^2$, $g_1(x_1, x_2, x_{3,} x_4) = (x_4, x_2 + 5\cos(x_3) + 2)$;

e) $g_2 : D \subset \mathbb{R}^3 \to \mathbb{R}^3$, $g_2(x_1, x_2, x_3) = (x_1, x_2 + x_3, 9)$;

f) $g_3 : D \subset \mathbb{R}^3 \to \mathbb{R}^3$, $g_3(x_1, x_2, x_3) = (x_1, \sqrt{x_2 + x_3}, 9)$;

g) $g_4 : D \subset \mathbb{R} \to \mathbb{R}^4$, $g_4(x) = (\sqrt{x}; x+5; 0; 2)$ (nota: usamos ponto e vírgula em vez da vírgula para não gerar confusão com a vírgula decimal);

h) $g_5 : D \subset \mathbb{R}^4 \to \mathbb{R}$, $g_5(x_1, x_2, x_{3,} x_4) = \frac{1}{x_1 + x_4} + x_2 + 5x_3 + 2$;

i) $g_6 : D \subset \mathbb{R}^2 \to \mathbb{R}^3$, $g_6(x_1, x_2) = (0, 0, 0)$;

j) $g_7 : D \subset \mathbb{R}^4 \to \mathbb{R}^2$, $g_7(x_1, x_2, x_{3,} x_4) = \left(\frac{1}{x_4}\ ,\ x_2 + 5\cos(x_3) + 2\right)$;

k) $g_8 : D \subset \mathbb{R}^4 \to \mathbb{R}^3$, $g_8(x_1, x_2, x_{3,} x_4) = \left(\ln(x_4), x_2 + \text{sen}(x_3) + 2e^{x_4}, \sqrt{2}\right)$;

l) $g_9 : D \subset \mathbb{R}^3 \to \mathbb{R}$, $g_9(x_1, x_2, x_3) = \frac{1}{(x_1-1)^2 + (x_2-3)^2 + x_3^2}$.

Exercícios 2.1:

1. Para cada uma das funções dos exemplos anteriores determine o seu domínio de definição.

2. Para cada uma das funções dos exemplos anteriores calcule o valor da função no vector nulo (se possível) e em mais um vector do domínio, à sua escolha.

3. Encontre a expressão da função composta e indique o conjunto de partida $\mathbb{R}^n$ e o conjunto de chegada $\mathbb{R}^m$:

a) $g_7 \circ g_4$;

b) $g_6 \circ g_1$;

c) $g_2 \circ g_6$;

d) $g_5 \circ g_4$;

e) $h \circ f$;

f) $g_1 \circ g_4$.

Vamos resolver apenas um exercícico de cada:

Resolução

Em 1. e 2. tomemos a função g_7 da alínea j).

1. O domínio de g_7 será a intersecção dos domínios das funções que constituem as suas componentes, tendo em conta que o conjunto de partida é $\mathbb{R}^4$. A primeira componente, $\frac{1}{x_4}$, tem domínio $\mathbb{R}^4 \backslash \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_4 = 0\}$ e a segunda, $x_2 + 5\cos(x_3) + 2$, tem domínio $\mathbb{R}^4$. Então o domínio é $D(g_7) = \mathbb{R}^4 \backslash \{(x_1, x_2, x_3, 0)\}$.

2. $g_7(0, 0, 0, 0)$ não faz sentido, porque o vector nulo não pertence ao domínio de g_7. Calculamos $g_7(1, 2, 0, 3) = (\frac{1}{3}, 2 + 5\cos(0) + 2) = (\frac{1}{3}, 9)$.

3. Resolvamos a alínea a): $(g_7 \circ g_4)$ será uma função de $\mathbb{R}$ em $\mathbb{R}^2$, cuja expressão é dada por

$$\begin{aligned}(g_7 \circ g_4)(x) &= g_7(g_4(x)) = g_7\left(\sqrt{x}; x+5; 0; 2\right) = \\ &= \left(\frac{1}{2},\ x + 5 + 5\cos 0 + 2\right) = \left(\frac{1}{2},\ x + 12\right).\end{aligned}$$

Como, em termos de domínios, mas não só, como veremos mais adiante, o que se estuda de facto são funções de $\mathbb{R}^n$ em $\mathbb{R}$, as chamadas **funções escalares**, dedicaremos grande parte do nosso estudo a este tipo de funções.

Na nossa vida deparamo-nos muitas vezes com funções escalares, por exemplo, a fórmula que dá a nota final de um aluno, tendo em conta a nota das frequências (f_1 e f_2), a nota dos trabalhos práticos (t_p) e a nota da oral (n_o)

$$n\left(f_1, f_2, t_p, o\right) = \frac{1}{6}f_1 + \frac{1}{6}f_2 + \frac{1}{3}t_p + \frac{1}{3}n_o.$$

Qual será a nota do aluno que obteve 10 na primeira frequência, 8 na segunda, 9 no trabalho prático e 13 na oral? Bastará calcular $n\left(10, 8, 9, 13\right)$.

Outro exemplo prático de aplicação é a função, conhecida da Física, que nos dá a pressão P no interior de um líquido em função da massa volúmica ρ, da aceleração da gravidade g e da altura h,

$$P\left(\rho, g, h\right) = \rho\ g\ h.$$

Outra função utilizada na Física é a função que dá a característica I_D de um díodo em função de 4 variáveis: a corrente de saturação inversa, i_s; a tensão nas extremidades do díodo, u_D; a tensão da temperatura, u_T e η é um parâmetro que caracteriza o semicondutor.

$$I_D\left(i_s, u_D, u_T, \eta\right) = i_s\left(e^{\frac{u_d}{\eta\ u_T}} - 1\right).$$

Geralmente, nas aplicações à Física, é costume chamar às funções com valores em $\mathbb{R}$, **campos escalares** e às funções com valores em $\mathbb{R}^n$, $n \neq 1$, **campos vectoriais**.

Sugerimos ao aluno que tente "inventar" outras funções que tenham aplicação na vida real.

O aluno já teve contacto com funções de $\mathbb{R}^n$ em $\mathbb{R}^m$ na disciplina de Álgebra Linear e Geometria, só que tratou apenas as que são lineares, logo representáveis através de matrizes, em bases dadas dos espaços vectoriais $\mathbb{R}^n$ e $\mathbb{R}^m$. As que vamos tratar agora poderão não ser lineares e a nossa abordagem a este tipo de funções terá outros objectivos.

Uma das dificuldades mais comuns para os alunos que agora iniciam o estudo de funções com mais do que uma variável é a dificuldade de visualização dos gráficos destas funções. O habitual sistema cartesiano, com o eixo dos xx e o eixo dos yy, fica sem aplicação quando se trata de uma função, por exemplo, de $\mathbb{R}^4$ em $\mathbb{R}^3$. No entanto é possível visualizar o gráfico de uma função de $\mathbb{R}^2$ em $\mathbb{R}$. Visualizamos uma função que, a cada ponto do plano (determinado pelas coordenadas x e y), associa uma altura (a coordenada z). O gráfico de uma tal função é uma figura de dimensão 3, identificada pelas coordenadas (x, y, z), com $z = f(x, y)$.

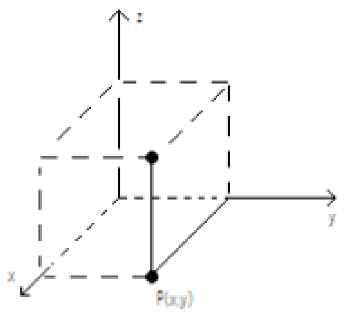


Figura 2.1: Ponto $P = (x, y)$ e a sua imagem $z = f(x, y)$

Vejamos algumas superfícies correspondentes a funções de $\mathbb{R}^2$ em $\mathbb{R}$:

$$z = f(x, y).$$

Exemplos

1. $z = x^2 - y^2$ (Parabolóide hiperbólico)

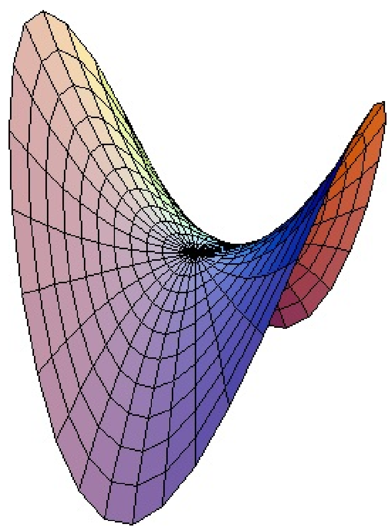


Figura 2.2: Parabolóide hiperbólico

2. $z = 2x + y$ (Plano)

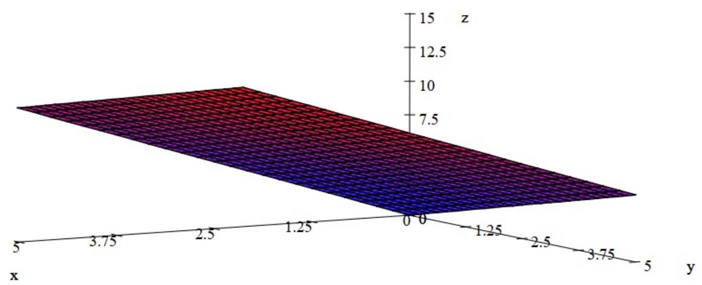


Figura 2.3: Plano

3. $z = x^2 + y^2$ (Parabolóide de revolução)

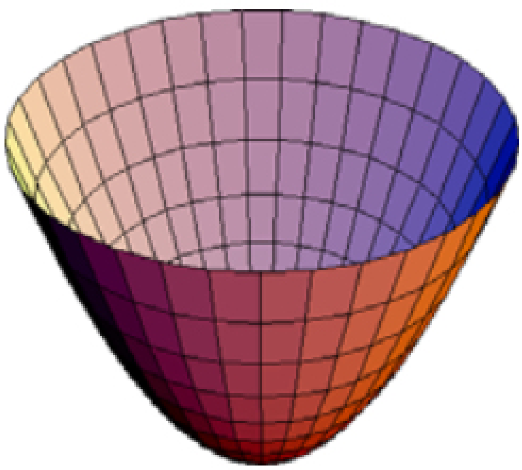


Figura 2.4: Parabolóide de revolução

4. $z = \sqrt{x^2 + y^2 + 1}$ (Hiperbolóide)

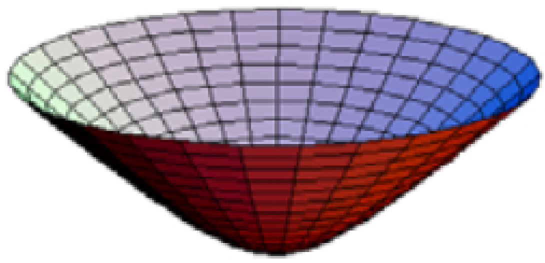


Figura 2.5: Hiperbolóide de revolução

5. $z = 2x + y^2$ (Superfície cilíndrica)

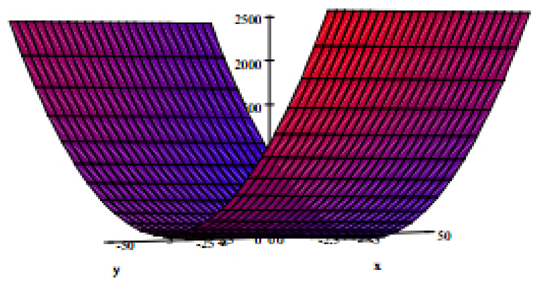


Figura 2.6: Superfície cilíndrica

6. $z = \sqrt{2x^2 + y^2}$ (Superfície cónica)

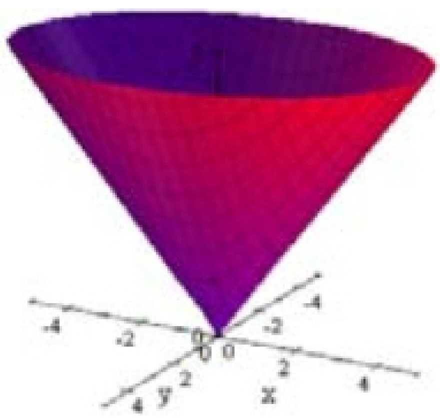


Figura 2.7: Superfície cónica

Para melhor raciocinar recorreremos bastantes vezes a estas funções de $\mathbb{R}^2$ em $\mathbb{R}$.

Ver é um importante passo para perceber, mas não é imprescindível, nem a única maneira.

Exercícios 2.2: Encontre o domínio de definição das seguintes funções $z = f(x, y)$ ou $u = f(x, y, z)$

a) $f(x,y) = \frac{1}{r^2-x^2-y^2},\ r > 0;$

b) $f(x,y) = \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\ ,\ a > b > 0$

c) $f(x,y) = ln(y^2 - 4x + 8);$

d) $f(x,y) = \sqrt{x \sin y};$

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

f) $f(x,y) = \ln x - \ln \sin y$

g) $f(x,y) = \arcsin \frac{y-1}{x};$

h) $f(x,y) = \ln xy + \pi y^2 \sqrt{x^2 - y^2};$

i) $f(x,y) = \ln(9 - x^2 - y^2) + \frac{1}{\sqrt{x^2+y^2-1}} + \arcsin \frac{y}{x};$

j)$f(x,y,z) = \frac{x}{|y|+|z|};$

k) $f(x,y,z) = \ln(xyz);$

l) $f(x,y,z) = \sqrt{4 - x^2 - y^2 - z^2}.$

Exercícios 2.3.

Esboce os gráficos das funções:

a) $z = x - y;$

b) $z = -x - y + 1;$

c) $z = 4x^2 + 9y^2;$

d) $z = x^2 - y^2;$

e) $z = 4 - x^2 - y^2;$

f) $z = \sqrt{1 - x^2 - y^2};$

g) $z = \sqrt{x^2 + y^2};$

$h)$ $z = 1 - y^2$.

2.2 Limites e Continuidade

A noção de limite para uma função de várias variáveis é muito semelhante à noção de limite para as já conhecidas funções reais de variável real. De facto, se se tratar de uma função vectorial de $\mathbb{R}$ em $\mathbb{R}^n$, o limite, se existir, é calculado como n limites de funções reais de variável real.

$$f : \mathbb{R} \to \mathbb{R}^m,$$
$$f(x) = (f_1(x), f_2(x), ..., f_m(x)),$$
$$\lim_{x \to a} f(x) = (\lim_{x \to a} f_1(x), \lim_{x \to a} f_2(x), ..., \lim_{x \to a} f_m(x)).$$

Para funções escalares, isto é, funções de n variáveis, com valores em $\mathbb{R}$, tem-se o seguinte

$$f : \mathbb{R}^n \to \mathbb{R},$$
$$f(x) = f(x_1, x_2, ..., x_n),$$
$$\lim_{x \to a} f(x) = \lim_{(x_1, x_2, ..., x_n) \to (a_1, a_2, ..., a_n)} f(x_1, x_2, ..., x_n),$$

não sendo trivial o significado de $x \to a$, com x e a elementos de $\mathbb{R}^n$. Menos trivial ainda é o limite de funções vectoriais com n variáveis. No fim tudo se resume à questão de distância, seja em $\mathbb{R}$, seja em $\mathbb{R}^n$: o conceito que queremos formalizar é o conceito de proximidade da imagem por f de vectores que se encontram "perto"de algum vector dado. Para isso necesitamos de uma definição mais geral que vai usar a definição de distância entre vectores que introduzimos atrás.

Definição 2.1 *Seja $f : \mathbb{R}^n \to \mathbb{R}^m$ uma função. Seja $a \in \mathbb{R}^n$ um ponto de acumulação do seu domínio D_f, seja $b \in \mathbb{R}^m$. Dizemos que*

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

sse

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 : (x \in D_f \;\wedge 0 < \|x - a\|_n < \delta) \Rightarrow \|f(x) - b\|_m < \varepsilon.$$

Por outras palavras: qualquer que seja a bola centrada em b, $B(b,\varepsilon)$, existe uma bola centrada em a, $B(a,\delta)$, tal que, para cada $x \in D_f \cap (B(a,\delta) \setminus \{a\})$, se tem $f(x) \in B(b,\varepsilon)$. Ou ainda: se me quero aproximar ε de b, então basta-me aproximar δ de a (ver Figura abaixo).

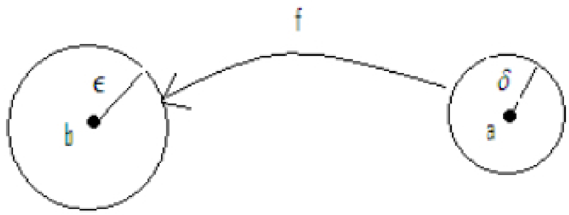


Figura 2.8: Relação entre as distâncias.

Se quisermos usar a definição de distância euclidiana ficaremos com a seguinte definição

$$\begin{aligned}
\forall\varepsilon &> 0 \;\; \exists\delta > 0 : \\
(x &\in D_f \;\wedge 0 < \sqrt{(x_1-a_1)^2+(x_2-a_2)^2+...+(x_n-a_n)^2} < \delta) \Rightarrow \\
&\Rightarrow \sqrt{(f_1(x)-b_1)^2+(f_2(x)-b_2)^2+...+(f_m(x)-b_m)^2} < \varepsilon.
\end{aligned}$$

O cálculo destes limites faz-se usando os métodos já estudados em Análise I, com as devidas adaptações e algumas novas técnicas. Os teorema gerais sobre operações algébricas sobre limites verificam-se para funções vectoriais.

Teorema 2.1 *O limite, quando existe, é único.*

A demonstração deste teorema é em tudo semelhante ao caso real.

Teorema 2.2 *Sejam f e $g : \mathbb{R}^n \to \mathbb{R}^m$ duas funções tais que $a \in \mathbb{R}^n$ é um ponto de acumulação dos domínios respectivos. Suponhamos que*

$$\lim_{x\to a} f(x) = b \;\; e \;\; \lim_{x\to a} g(x) = c.$$

Então tem-se

1. $\lim_{x\to a}[f(x)+g(x)] = b+c.$

2. $\lim_{x\to a} \lambda f(x) = \lambda b$ *para todo o escalar* λ.

3. $\lim_{x\to a} (f(x), g(x))_m = (b, c)_m$. *(Aqui* $(.,.)_m$ *é o produto interno em* $\mathbb{R}^m$*).*

4. $\lim_{x\to a} \|f(x)\|_m = \|b\|_m$.

Vejamos os seguintes exemplos onde se podem aplicar.

Exemplos:

a) $\lim\limits_{(x,y)\to(1,2)} (x^2 + y + 2) = 1^2 + 2 + 2 = 5;$

b) $L = \lim\limits_{(x,y)\to(0,0)} \frac{2-\sqrt{4-xy}}{xy} = \frac{0}{0}$(indeterminação)

$$L = \lim_{(x,y)\to(0,0)} \frac{2-\sqrt{4-xy}}{xy}\frac{2+\sqrt{4-xy}}{2+\sqrt{4-xy}} = \lim_{(x,y)\to(0,0)} \frac{4-(4-xy)}{xy\left(2+\sqrt{4-xy}\right)} =$$
$$= \lim_{(x,y)\to(0,0)} \frac{xy}{xy}\frac{1}{\left(2+\sqrt{4-xy}\right)} = \frac{1}{4}.$$

c) $L = \lim\limits_{(x,y)\to(0,0)} \frac{x^2+y^2}{\sqrt{x^2+y^2+1}-1} = \frac{0}{0}$(indeterminação)

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

d) $L = \lim\limits_{(x,y)\to(0,a)} \frac{\sin xy}{x} = \frac{0}{0}$(indeterminação)

$$L = \lim_{(x,y)\to(0,a)} y\frac{\sin xy}{xy} = a$$

Exercícios 2.4. Calcule os seguintes limites

a) $\lim\limits_{(x,y)\to(0,0)} \frac{1-cos(x^2+y^2)}{(x^2+y^2)x^2y^2};$

b) $\lim\limits_{(x,y)\to(4,0)} \frac{\operatorname{tg} xy}{y}.$

Enunciemos e demonstremos agora o teorema conhecido resumidamente por "o produto de um infinitésimo por uma limitada é um infinitésimo".

Teorema 2.3 *Sejam* f *e* $g : \mathbb{R}^n \to \mathbb{R}^m$ *duas funções tais que* $a \in \mathbb{R}^n$ *é um ponto de acumulação dos domínios respectivos. Suponhamos que*

$$\lim_{x\to a} f(x) = \mathbf{0}$$

e que g seja limitada numa vizinhança de a, isto é,

$$\exists L > 0 \ \exists \delta > 0 : \forall x \in D_g \ , \ \|x - a\|_n < \delta \Rightarrow \|g(x)\|_m < L.$$

Então

$$\lim_{x \to a} (f(x), g(x))_m = 0.$$

Demonstração:

Pelo facto de g ser limitada numa vizinhança-δ_1 de a, temos

$$\|g(x)\|_m < L$$

sempre que $x \in B(x, \delta_1)$.

Agora usamos a definição de limite,

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 : (x \in D_f \ \wedge 0 < \|x - a\|_n < \delta) \Rightarrow \|f(x)\|_m < \varepsilon,$$

para concluir que dado $\varepsilon > 0$ qualquer, existe $\delta_2 > 0$ tal que

$$\|f(x)\|_m < \frac{\varepsilon}{L}$$

para $x \in B(x, \delta_2)$. Assim, tomando $\delta = \min\{\delta_1, \delta_2\}$, temos, usando a conhecida desigualdade de Cauchy-Schwarz,

$$|(f(x), g(x))_m| \leq \|f(x)\|_m \|g(x)\|_m < \frac{\varepsilon}{L} L = \varepsilon$$

sempre que $x \in B(x, \delta)$ e, como ε era qualquer, temos o resultado pretendido.■

A generalização da definição de limite anterior a limites infinitos ou limites quando a variável tende para infinito (o que é isso em $\mathbb{R}^n$?) faz-se por analogia com as funções reais, novamente com as devidas adaptações.

Nota Usaremos a notação $x \to \infty$, com $x = (x_1, x_2, ..., x_n)$ para dizer que $\|x\|_n \to \infty$.

Definição 2.2 *Seja f uma função com domínio $D_f \subset \mathbb{R}^n$ contendo pontos arbitrariamente distantes da origem e com valores em $\mathbb{R}^m$. Seja $b \in \mathbb{R}^m$. Dizemos que*

$$\lim_{x \to \infty} f(x) = b$$

sse

$$\forall \varepsilon > 0 \ \exists M > 0 \ \forall x \in D_f \ , \ \|x\|_n > M \Rightarrow \|f(x) - b\|_m < \varepsilon.$$

Definição 2.3 *Seja f uma função com domínio $D_f \subset \mathbb{R}^n$ e com valores em $\mathbb{R}^m$. Seja $a \in \mathbb{R}^n$ ponto de acumulação de D_f. Dizemos que*

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

sse

$$\forall M > 0 \; \exists \delta > 0 \; \forall x \in D_f \;,\; 0 < \|x - a\|_n < \delta \Rightarrow \|f(x)\|_m > M.$$

Exercício 2.5. Formule a definição de

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

Coordenadas polares, cilíndricas e esféricas

Por vezes é útil efectuar uma mudança de variáveis. Vamos introduzir três exemplo típicos, que nos servirão também, quando iniciarmos o cálculo integral em $\mathbb{R}^n$. O primeiro exemplo aplica-se a $\mathbb{R}^2$, passando do sistema de coordenadas cartesianas, as habituais (x, y), para o sistema de **coordenadas polares** , que denotaremos (r, θ). Na seguinte figura temos ilustrada a situação no primeiro quadrante (x e y positivos):

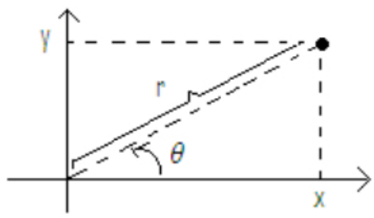


Figura 2.9: Correspondência entre coordenadas

As fórmulas que nos permitem passar de umas para as outras são as seguintes:

$$\begin{aligned} x &= r\cos\theta; \\ y &= r\,\mathrm{sen}\,\theta, \end{aligned}$$

e vice-versa

$$\begin{aligned} r &= \sqrt{x^2 + y^2}; \\ \theta &= \arctan\frac{y}{x}. \end{aligned}$$

O aluno facilmente as deduz da figura, mas atenção, a figura representa um ponto no primeiro quadrante, convém pensar um pouco sobre a passagem aos outros quadrantes. Para conseguirmos descrever todos os pontos do plano $\mathbb{R}^2$, necessitamos de $r \in \mathbb{R}_0^+$ e $\theta \in [0, 2\pi)$.

O segundo exemplo diz respeito a $\mathbb{R}^3$, trata-se das **coordenadas cilíndricas**, que não são mais do que uma generalização das coordenadas polares, (r, θ, z), acrescentando-lhes a "altura".

$$\begin{aligned} x &= r\cos\theta; \\ y &= r\,\mathrm{sen}\,\theta; \\ z &= z. \end{aligned}$$

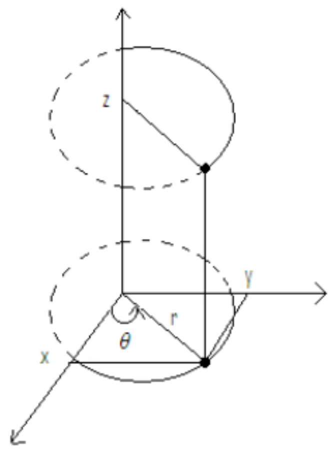


Figura 2.10: Coordenadas cilíndricas

O terceiro exemplo, que usaremos igualmente, é também uma mudança de variáveis em $\mathbb{R}^3$, são as **coordenadas esféricas** (r, ϕ, θ), dadas por

$$\begin{aligned} x &= r\,\mathrm{sen}\,\phi\cos\theta; \\ y &= r\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta; \\ z &= r\cos\phi \end{aligned}$$

e as inversas

$$\begin{aligned} r &= \sqrt{x^2+y^2+z^2}; \\ \theta &= \arctan \frac{y}{x}; \\ \phi &= \arctan \frac{\sqrt{x^2+y^2}}{z}. \end{aligned}$$

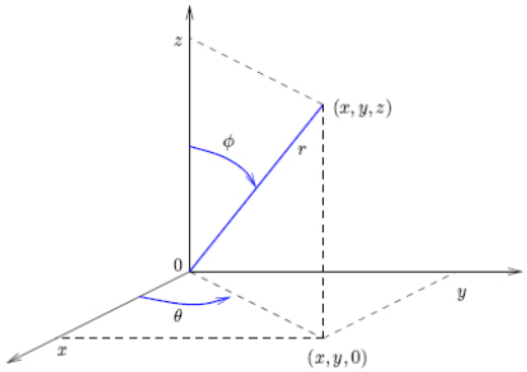


Figura 2.11: Coordenadas esféricas

Esta transformação é biunívoca (aceitemos uma ligeira ambiguidade em relação ao ponto $(0,0,0)$ que deveria ser tratado à parte) quando $r \in \mathbb{R}_0^+$ e $\theta \in [0, 2\pi)$ e (atenção!) $\phi \in (0, \pi)$.

Como é que podemos usar estas transformações? Vejamos num exemplo.

Exemplo. Seja $f : \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x,y) = \frac{x^4+y^4}{x^2+y^2}$. A mesma função[1],

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

em coordenadas polares, ficará

[1]**Nota:** Usámos aqui um abuso de linguagem que, não sendo grave, ajuda a compreensão do conceito de funções dadas em sistemas de coordenadas diferentes. Quando dizemos que f e $\overline{f}$ são a <u>mesma</u> função não estamos a ser correctos na linguagem. Duas funções são iguais se tiverem o mesmo domínio e se as imagens de cada ponto do domínio coincidirem. Ora aqui , olhando para as novas variáveis como funções de x e de y, a igualdade é no seguinte sentido

$$f(x,y) = \overline{f}(r(x,y), \theta(x,y)) = r^2(x,y) = \left(\sqrt{x^2+y^2}\right)^2 = x^2+y^2.$$

$$z = \bar{f}(r,\theta) = \frac{(r\cos\theta)^4 + (r\operatorname{sen}\theta)^4}{(r\cos\theta)^2 + (r\operatorname{sen}\theta)^2} = \frac{r^4(\operatorname{sen}^4\theta + \cos^4\theta)}{r^2(\operatorname{sen}^2\theta + \cos^2\theta)} = r^2(\overset{4}{\operatorname{sen}}\,\theta + \cos^4\theta).$$

Agora, usando os seguintes factos:

1. $(x,y) \to (0,0) \Leftrightarrow r \to 0$ e θ é qualquer (Porquê?)

2. $(\operatorname{sen}^4\theta + \cos^4\theta)$ é uma função limitada,

3. O produto de um infinitésimo por uma função limitada é um infinitésimo,

obtemos

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{r\to 0} \overline{f}\,(r,\theta) = \lim_{r\to 0} r^2(\overset{4}{\operatorname{sen}}\,\theta + \cos^4\theta) = 0.$$

Como vimos, com estas novas coordenadas, podemos resolver facilmente alguns limites de funções vectoriais. Experimente determinar os seguintes limites usando estas mudanças de variáveis.

Exercício 2.6.

a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2+y^2}}$;

b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x+y}{\sqrt{x^2+y^2}}$;

c) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{x^2+y^2-z^2}{\sqrt{x^2+y^2}+z}$.

No Capítulo do Cálculo integral será também necessário descrever certas regiões do plano ou do espaço em termos das novas coordenadas. Assim, por exemplo o conjunto $A = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 9\}$ pode exprimir-se nas coordenadas polares (r,θ) como $A' = \{(r,\theta) : 0 \le r \le 3, 0 \le \theta < 2\pi\}$.

Exercício 2.7. Descreva usando coordenadas polares as seguintes regiões do plano:

a) $A = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 4\}$;

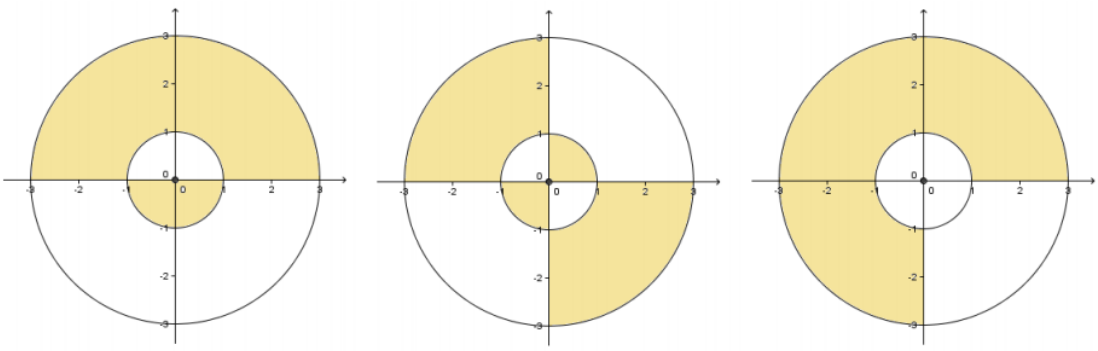
b) $B = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 4 \wedge y > 0\}$;

c) $C = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 5 \wedge xy > 0\}$;

d) $D = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 9 \wedge xy < 0\}$;

e) O disco situado entre as duas circunferências de raios 4 e 6;

f) As regiões a sombreado na figura:



Exercício 2.8. Esboce e descreva usando coordenadas cilíndricas as seguintes regiões do espaço:

a) $A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 4 \wedge 0 < z < 4\}$;

b) $B = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 < 4 \wedge y > 0 \wedge -2 < z < 2\}$;

c) $C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 > 5 \wedge xy > 0 \wedge -1 < z < 1\}$;

d) $D = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 < 9 \wedge xy < 0 \wedge z > 1\}$;

e) O cilindro com base no plano $z = -1$, altura 4 e raio da base igual a 5;

f) O cone com vértice no ponto $(0, 0)$ e base de raio 3 assente no plano $z = 3$.

Exercício 2.9. Esboce e descreva usando coordenadas esféricas as seguintes regiões do espaço:

a) $A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 4\}$;

b) $B = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 4 \wedge y > 0\}$;

c) $C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 25 \wedge xy > 0\}$;

d) $D = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 9 \wedge x^2 + y^2 + z^2 > 1\}$;

e) A coroa esférica situada entre as esferas centradas na origem e raios 2 e 3;

f) O volume delimitado pelo cone $x^2 + y^2 = z$ e pela esfera centrada na origem e raio 4.

Limites iterados e limites relativos

Observando o plano $0xy$ constatamos que, tal como em $\mathbb{R}$, podem existir várias maneiras de nos aproximarmos de um dado ponto. Em $\mathbb{R}$ falámos do limite à esquerda e do limite à direita. Em $\mathbb{R}^2$ existem "muito mais limites laterais", isto é limites de funções com restrições do domínio, vamos chamar-lhes **limites relativos** ou **limites**

segundo caminhos. Assim, por exemplo, podemos aproximar-nos tanto quanto queiramos do ponto $(0,0)$ através de pontos sobre a recta de equação $y = 2x$, ou através de pontos sobre a parábola $y = x^2$, ou ainda através da recta vertical $x = 0$, etc., etc.. Nestes casos transformamos a função de duas variáveis numa função de apenas uma variável, visto que y aparece como função de x e o cálculo de limites resume-se ao cálculo de limites de uma função real de variável real. Quando a restrição escolhida é uma recta não vertical $y = mx + b$, falamos de **limites direccionais ou rectilíneos**.

ATENÇÃO: o facto de existirem limites segundo alguns caminhos não implica que a função tenha limite, poderá convergir por um (ou mais) caminhos para um dado valor, mas por outros não. O que se está a fazer é a restringir a casos especiais de aproximação aos pontos, que não garantem que não haja outros caminhos que não levem a outros valores. Vejamos o que pode acontecer.

Exemplos:

a) Seja $f(x,y) = \frac{x+y}{3x}$. Queremos averiguar o seu comportamento em torno da origem. Se restringirmos o domínio aos pontos (x,y) com $y = 2x$, obtemos

$$\lim_{\substack{(x,y)\to(0,0)\\ y=2x}} \frac{x+y}{3x} = \lim_{x\to 0} \frac{x+2x}{3x} = \lim_{x\to 0} \frac{3x}{3x} = 1.$$

Mas se a restrição se fizer aos pontos (x,y) com $y = 5x$, obtemos

$$\lim_{\substack{(x,y)\to(0,0)\\ y=5x}} \frac{x+y}{3x} = \lim_{x\to 0} \frac{x+5x}{3x} = \lim_{x\to 0} \frac{6x}{3x} = 2.$$

Ou seja, podemos ter pontos arbitrariamente próximos de $(0,0)$ e imagens arbitrariamente próximas de 1 e pontos arbitrariamente próximos de $(0,0)$ com imagens arbitrariamente próximas de 2. Afinal qual é o limite? Evidentemente o limite não existe (Justifique!)

b) Seja $f(x,y) = \frac{x^2+y}{y}$. Queremos averiguar o seu comportamento novamente em torno da origem. Se restringirmos o domínio aos pontos (x,y) com $y = mx$, $m \in \mathbb{R}\backslash\{0\}$, obtemos

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

Ou seja, todos os limites direccionais são iguais, mas se a restrição se fizer aos pontos (x, y) com $y = x^2$ (pontos sobre a parábola), obtemos

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

Novamente o limite não existe!

Conclusão: Se encontrarmos limites diferentes segundo caminhos diferentes, então o limite não existe. Mas a recíproca não é verdadeira, se encontrarmos limites iguais segundo caminhos diferentes, nada podemos concluir.

Outra maneira de abordar limites em funções de duas variáveis (ou mais) é fixar uma das variáveis, como se se tratasse de um parâmetro, calcular o limite da função olhando apenas para a outra variável (obtendo funções dependentes do parâmetro) e, de seguida, calcular o limite para a variável que se fixou primeiro. A este tipo de abordagem chamamos **limites iterados**.

Definição 2.4 *Seja f uma função com domínio $D_f \subset \mathbb{R}^2$ e com valores em $\mathbb{R}^m$. Seja (x_0, y_0) ponto de acumulação de D_f. Suponhamos que para cada $x \neq x_0$, com $(x, y) \in D_f$, existe $\lim_{y\to y_0} f(x, y) =: g(x)$ e suponhamos que esta função tem limite no ponto x_0, então*

$$\lim_{x\to x_0} g(x) = \lim_{x\to x_0} \left(\lim_{y\to y_0} f(x, y) \right)$$

*é chamado **limite iterado** de $f(x, y)$ em relação a x e a y.*

Neste caso não há restrição ao domínio, mas pode acontecer que os limites unidimensionais (os que se calculam, fixando a outra variável) não existam. Vejamos o que pode acontecer.

Exemplo. Calculemos

$$\lim_{(x,y)\to(0,0)} (x+y) \operatorname{sen} \frac{1}{x} \operatorname{sen} \frac{1}{y}.$$

Usando o teorema 2.2. temos um infinitésimo a multiplicar por uma função limitada, logo é um infinitésimo. Mas, se quisermos calcular os limites iterados,

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} (\lim_{y\to 0} (x+y) \operatorname{sen} \frac{1}{x} \operatorname{sen} \frac{1}{y})$$

não conseguimos, porque

$$\lim_{y \to 0} \operatorname{sen} \frac{1}{y}$$

não existe.

No entanto, há situações em que o limite se pode substituir pelos limites iterados. Temos o seguinte teorema que enunciamos, por uma questão de simplicidade, para funções de $\mathbb{R}^2$ em $\mathbb{R}$.

Teorema 2.4 *Suponhamos que existe*

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

Suponhamos que, para cada $x \neq a$ fixo, existe

$$\lim_{y \to b} f(x,y) = g(x).$$

Então existe o limite

$$\lim_{x \to a} g(x) \text{ e verifica-se } \lim_{x \to a} g(x) = L.$$

Demonstração:

Da existência de limite obtemos que, para cada $\varepsilon > 0$, existe um $\delta > 0$ tal que $|f(x,y) - L| < \varepsilon$ sempre que $(x,y) \neq (a,b)$ pertença a uma dada vizinhança-δ do vector (a,b), $B((a,b),\delta)$. Daqui podemos concluir que esta desigualdade se verifica para todos os vectores do quadrado

$$Q =]a - \delta_1, a + \delta_1[\times]b - \delta_1, b + \delta_1[\subset B((a,b),\delta)$$

sem o vector (a,b), onde δ_1 é, por exemplo, $\frac{\delta}{\sqrt{2}}$.*(Pensar porquê!)*

Tomemos agora $x = x_0$ um número real fixo tal que $0 < |x_0 - a| < \delta_{1.}$ Então, para cada $y \in]b - \delta_1, b + \delta_1[$, $y \neq b$, temos

$$|f(x_0,y) - L| < \varepsilon.$$

Passando ao limite, quando $y \to b$ obtemos

$$|g(x_0) - L| \leq \varepsilon < 2\varepsilon.$$

Provámos então que para $x_0 \in]a - \delta_1, a + \delta_1[$, $x_0 \neq a$, se tem $|g(x_0) - L| < 2\varepsilon$ e, como ε é qualquer, isto significa que

$$\lim_{x_0 \to a} g(x_0) = L,$$

como queríamos demonstrar. ■

Este teorema diz, na prática, em que condições se pode calcular um limite, usando os limites iterados. Além disso fornece-nos uma ferramenta que serve para mostrar que não existe limite, no caso em que os limites iterados existem, mas dependem da ordem de iteração. Se, trocando a ordem em que fixamos as variáveis, no cálculo do limite iterado, obtivermos valores diferentes, então o limite não existe.

Exemplo: Considerando

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

obtemos

$$\lim_{x\to 0}\left(\lim_{y\to 0}\frac{2x+y}{x+y}\right) = \lim_{x\to 0}\frac{2x}{x} = 2$$

e

$$\lim_{y\to 0}\left(\lim_{x\to 0}\frac{2x+y}{x+y}\right) = \lim_{y\to 0}\frac{y}{y} = 1.$$

Se o limite existisse, usando o teorema duas vezes, teríamos o mesmo resultado (ver Exercício 2.7.2). Como tal não acontece, o limite não existe.

Além disto, nada nos garante que, pelo facto de os limites iterados existirem e serem iguais, o limite exista. Vejamos o seguinte exemplo:

Exemplo: Seja $f(x,y) = \dfrac{x^2y^2}{x^2y^2+(x-y)^2}$ sempre que $x^2y^2+(x-y)^2 \neq 0$. Os limites iterados verificam $\lim\limits_{x\to 0}\left[\lim\limits_{y\to 0} f(x,y)\right] = \lim\limits_{y\to 0}\left[\lim\limits_{x\to 0} f(x,y)\right] = 0$ mas, no entanto, $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ não existe (Calcular os limites direccionais com $y = x$ e com $y = 2x$).

Continuidade

Um dos momentos em que a maneira como é interiorizada a noção de continuidade no ensino secundário colide com a noção geral de continuidade é precisamente quando se introduz essa noção em funções que não têm domínio real nem valores reais. Como é que será possível traduzir essa ideia, se a compreensão deste conceito for baseada em "conseguir ou não conseguir desenhar o gráfico sem levantar o bico do lápis..."? Recorreremos novamente à Análise I e é da noção formal de continuidade estudada aí que partimos para a sua generalização. Uma função é contínua num ponto sse o limite nesse ponto existe e é igual ao valor da função. Matematicamente falando:

Definição 2.5 *Seja $f : \mathbb{R}^n \to \mathbb{R}^m$ uma função. Seja $a \in \mathbb{R}^n$ um ponto do seu domínio D_f. Dizemos que f é* ***contínua em*** *a sse existe o limite $\lim_{x\to a} f(x)$ e*

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

isto é, f é contínua em a sse

$$\forall\varepsilon > 0 \ \ \exists\delta > 0 : (x \in D_f \ \wedge 0 < \|x - a\|_n < \delta) \Rightarrow \|f(x) - f(a)\|_m < \varepsilon.$$

Dizemos que f ***é contínua num conjunto*** *$S \subset \mathbb{R}^n$ sse for contínua em cada ponto desse conjunto.*

Teorema 2.5 *A função $f : \mathbb{R}^n \to \mathbb{R}^m$ é contínua num ponto sse cada componente $f_k : \mathbb{R}^n \to \mathbb{R}$ for contínua nesse ponto.*

Demonstração: Resulta do Teorema 2.1., do facto de cada componente de f se escrever na forma de produto interno

$$f_k = \langle f, \mathbf{e}_k\rangle_m,$$

onde e_k é o k-ésimo vector da base canónica de R^n (tem todas as componentes nulas, excepto a $k-$ésima, que é igual a 1) e, além disso,

$$f(\mathbf{x}) = \sum_{k=1}^{m} f_k(\mathbf{x})\,\mathbf{e}_k.$$

■

Exemplos de funções contínuas:

- A função identidade, $f(\mathbf{x}) = \mathbf{x}$ é contínua em $\mathbb{R}^n$.

Daqui, pelo teorema anterior, também as projecções $f(\mathbf{x}) = x_i$ o são, qualquer que seja o $i = 1, 2, ..., n$.

- A transformação linear, $f(\mathbf{x}) = A\mathbf{x}$, sendo A uma matriz do tipo $m \times n$, é contínua em $\mathbb{R}^n$.

Basta ver que $f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + f(\mathbf{h})$. E que, sendo $\mathbf{h} = (h_1, h_2, ..., h_3) = h_1 e_1 + h_2 e_2 + ... + h_n e_n$, temos

$$f(\mathbf{h}) = h_1 f(\mathbf{e}_1) + h_2 f(\mathbf{e}_2) + ... + h_n f(\mathbf{e}_n).$$

Daqui sai que $f(\mathbf{h}) \to \mathbf{0}$, quando $\mathbf{h} \to \mathbf{0}$.

- Os polinómios de n variáveis são contínuos em $\mathbb{R}^n$.

Basta ver que são somas e produtos de projecções.

- Funções racionais: as funções escalares que resultam do quociente entre polinómios nas componentes de $\mathbf{x}$
$$f(\mathbf{x}) = \frac{P(\mathbf{x})}{Q(\mathbf{x})},$$
são contínuas em todos os pontos em que $Q(\mathbf{x}) \neq 0$.

Teorema 2.6 *Sejam f e g funções tais que $f \circ g$ está definida em a. Se g é contínua em a e f é contínua em $g(a)$, então a composta $f \circ g$ é contínua em a.*

Demonstração: Sejam $y = g(x)$ e $b = g(a)$. Como g é contínua em a, $y \to b$ quando $x \to a$.

Temos então

$$\lim_{x \to a}[f(g(x)) - f(g(a))] = \lim_{y \to b}[f(y) - f(b)] = 0$$

porque f é contínua em $b = g(a)$. Isto significa precisamente que

$\lim\limits_{x \to a} f(g(x)) = f(g(a))$, pelo que $f \circ g$ é contínua em a.■

Exemplos: Do teorema anterior sai imediatamente a continuidade em $\mathbb{R}^2$ das funções escalares seguintes, nos pontos em que estão definidas:

a) $\sin(x^2y)$; b) $\ln(x^2+y^2)$; c) $\dfrac{e^{xy}}{x+y}$; d) $e^{\cos(x+y)}$.

Exercícios 2.10.

1. Em cada uma das alíneas seguintes indique o conjunto de pontos de $\mathbb{R}^2$ onde f é contínua.

a) $f(x,y) = x^4 + y^4 - 4x^2y^2$;

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

c) $f(x,y) = \dfrac{1}{y}\cos x^2$;

d) $f(x,y) = \operatorname{tg}\dfrac{x^2}{y}$;

e) $f(x,y) = \operatorname{arctg}\dfrac{y}{x}$;

f) $f(x,y) = \arcsin\dfrac{x}{\sqrt{x^2+y^2}}$;

g) $f(x,y) = \arctan\dfrac{x+y}{1-xy}$;

h) $f(x,y) = \dfrac{x}{\sqrt{x^2+y^2}}$;

i) $f(x,y) = \arccos\sqrt{x/y}$

2. Se $\lim_{(x,y)\to(a,b)} f(x,y) = L$ e se existirem os limites unidimensionais

$$\lim_{x\to a} f(x,y) \qquad \text{e} \qquad \lim_{y\to b} f(x,y)$$

prove que existem os limites iterados e

$$\lim_{x\to a}[\lim_{y\to b} f(x,y)] = \lim_{y\to b}[\lim_{x\to a} f(x,y)] = L.$$

3. Seja $f(x,y) = \dfrac{x-y}{x+y}$ se $x+y \neq 0$. Prove que

$$\lim_{x\to 0}[\lim_{y\to 0} f(x,y)] = 1 \qquad \text{mas que} \qquad \lim_{y\to 0}[\lim_{x\to 0} f(x,y)] = -1.$$

Utilize este resultado e o do exercício 2. para mostrar que $f(x,y)$ não tem limite quando $(x,y) \to (0,0)$.

4. Seja

$$f(x,y) = \frac{x^2y^2}{x^2y^2 + (x-y)^2} \text{ sempre que } x^2y^2 + (x-y)^2 \neq 0.$$

Prove que

$$\lim_{x\to 0}[\lim_{y\to 0} f(x,y)] = \lim_{y\to 0}[\lim_{x\to 0} f(x,y)] = 0$$

mas que $f(x,y)$ não tem limite quando $(x,y) \to (0,0)$. (Sugestão: Examine f sobre a recta $y = x$). Este exercício mostra que a inversa do exercício 2. nem sempre é verdadeira.

5. Seja

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

Mostre que $f(x,y) \to 0$ quando $(x,y) \to (0,0)$, mas que

$$\lim_{x\to 0}[\lim_{y\to 0} f(x,y)] \neq \lim_{y\to 0}[\lim_{x\to 0} f(x,y)].$$

Explique por que razão isto não contradiz o exercício 2.

6. Se $(x,y) \neq (0,0)$, seja

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

Determine o limite de $f(x,y)$ quando $(x,y) \to (0,0)$, ao longo da recta $y = mx$. É possível definir $f(0,0)$ de modo que f seja contínua em $(0,0)$?

7. Seja

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

Prove que $f(x,y) \to 0$, quando $(x,y) \to (0,0)$ ao longo de qualquer recta passando pela origem. Determine uma curva, passando pela origem, ao longo da qual f (excepto na origem) seja constante igual a 1. Será f contínua na origem?

8. Se

$$f(x,y) = \frac{\operatorname{sen}(x^2+y^2)}{x^2+y^2}, \text{ quando } (x,y) \neq (0,0),$$

como deve ser definida $f(0,0)$ para que f seja contínua na origem?

Capítulo 3

Cálculo Diferencial em $\mathbb{R}^n$

3.1 Noção de derivada

(Ver sobretudo [Webb])

Para uma função real de variável real, $f : \mathbb{R} \to \mathbb{R}$, a noção de derivada num ponto a é definida por

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \left(= \lim_{t \to 0} \frac{f(a+t) - f(a)}{t}\right), \text{ sempre que este limite exista.}$$

Para funções vectoriais, ou mesmo escalares, esta definição não faz sentido, visto não estar definida a divisão por elementos de $\mathbb{R}^n$, com $n \neq 1$. De qualquer modo, a derivada em $\mathbb{R}$ fornece uma boa informação sobre o comportamento local da função, é, por exemplo, uma medida da taxa de variação da função. Para além disso é também o declive da recta tangente ao gráfico da função no ponto considerado. É a partir desta perspectiva que vamos construir a generalização da derivada às funções de variável em $\mathbb{R}^n$.

Quando uma função real de variável real, f, é diferenciável em a, o erro cometido ao aproximar a função pela recta tangente ao seu gráfico, num ponto x, "perto" de a, mas diferente de a é

$$R(x,a) = f(x) - [f(a) + f'(a)(x-a)] = (x-a)\left[\frac{f(x) - f(a)}{x-a} - f'(a)\right].$$

Pela definição de $f'(a)$, o termo dentro de [ ] tem limite 0, quando $x \to a$ e daí $\frac{R(x,a)}{(x-a)} \to 0$ quando $|x-a| \to 0$. Isto significa que $f(a) + f'(a)(x-a)$ é uma boa aproximação de $f(x)$ numa vizinhança de a. É esta ideia que vai servir para a generalização que pretendemos. Obviamente, não podemos usar a noção de linha tangente, mas podemos usar a linguagem dos operadores lineares. Observemos que $Lx := f'(a)\,x$ define uma função linear e uma boa aproximação para $f(x)$ é $f(a) + L(x-a)$.

Definição 3.1 *Seja $f : D_f \to \mathbb{R}^m$ uma função, com $D_f \subset \mathbb{R}^n$, aberto. Dizemos que f* ***é diferenciável num ponto*** $\mathbf{a} \in D_f$ *sse existir um operador linear $L : \mathbb{R}^n \to \mathbb{R}^m$ tal que*

$$f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x}-\mathbf{a}) + R(\mathbf{x},\mathbf{a}),\ \ \mathbf{x} \in D_f,$$

onde$\frac{\|R(\mathbf{x},\mathbf{a})\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} \to 0$, *quando* $\|\mathbf{x}-\mathbf{a}\|_n \to 0$. *Dizemos que f é diferenciável em D_f sse for diferenciável em cada ponto de D_f.*

Chamamos a L a ***derivada de f no ponto*** $\mathbf{a}$ *e denotamo-la $Df(\mathbf{a})$.*

Evidentemente, L depende de $\mathbf{a}$ e de f. Não dizemos que L existe para todo o $\mathbf{a}$, mas sim que, para cada $\mathbf{a}$ fixo, existe L. Além disso convém verificar que a definição está correcta, ou seja, que a derivada assim definida é única.

Nota:. Demonstra-se que a derivada é única, isto é, se M e L forem operadores lineares nas condições da Definição 3.1., então $M\mathbf{x} = L\mathbf{x}$ para todo o $\mathbf{x} \in \mathbb{R}^n$.

Apesar de ser possível demonstrar este resultado directamente da definição, a nossa abordagem vai levar-nos ao cálculo deste operador em termos de quantidades mais familiares. De facto, vamos aprender a determinar a matriz que representa $Df(\mathbf{a})$. Não definimos de princípio assim, porque a linguagem matricial depende das bases em que trabalhamos, apesar do operador linear não depender. É, por vezes, útil considerar uma mudança de bases, sabendo que a derivada não muda, é invariante.

Nota: Recordemos, de álgebra linear, que uma função linear $f : \mathbb{R}^n \to \mathbb{R}^m$, ou operador linear, é uma função que verifica as duas propriedades seguintes:

$$\begin{array}{ll} 1) & f(x+y) = f(x) + f(y), \\ 2) & f(cx) = cf(x). \end{array}$$

para todo o $c \in \mathbb{R}$ e $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Recordemos ainda que uma tal função pode ser representada, para uma dada escolha de bases dos espaços vectoriais $\mathbb{R}^n$ e $\mathbb{R}^m$, por uma matriz, cujas colunas constituem as imagens dos vectores da base de $\mathbb{R}^n$ expressas na base de $\mathbb{R}^m$.

Consideremos os dois exemplos seguintes em que a derivada se pode calcular directamente a partir da definição.

Exemplo: Seja $f : \mathbb{R}^n \to \mathbb{R}^m$ definida por $f(\mathbf{x}) = \mathbf{c}$, para todo o $\mathbf{x} \in \mathbb{R}^n$, sendo $\mathbf{c} \in \mathbb{R}^m$ (a função constante). Então $Df(\mathbf{a}) = 0$, o operador zero. Verifiquemos:

$$f(\mathbf{x}) = f(\mathbf{a}) + \mathbf{0} + R, \text{ onde } \frac{\|R\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} \to 0, \text{ quando } \|\mathbf{x}-\mathbf{a}\|_n \to 0,$$

verifica-se para $R = 0$.

Exemplo: Seja $f : \mathbb{R}^n \to \mathbb{R}^m$ uma função linear, então $Df(\mathbf{a}) = f$ para todo o $\mathbf{a}$. Basta ver que $f(\mathbf{x}) = f(\mathbf{a}) + f(\mathbf{x}-\mathbf{a}) + R$ com $R = 0$, por linearidade.

Exercício 3.2. Mostre que $f : \mathbb{R}^2 \to \mathbb{R}^3$, definida por

$$f(x, y) = (4x, x - y, 3x + y),$$

tem derivada $Df(x, y) = f(x, y)$, seguindo o raciocínio do exemplo anterior.

Como definimos derivada apenas em pontos de um conjunto aberto D_f, daqui para a frente, D_f denotará um aberto de $\mathbb{R}^n$, a menos que seja dito o contrário.

Uma consequência simples, mas importante, da diferenciabilidade é a continuidade.

Teorema 3.1 *Seja* $\dot{f} : D_f \subset \mathbb{R}^n \to \mathbb{R}^m$ *uma função diferenciável em* $\mathbf{a} \in D_f$*. Então* f *é contínua em* $\mathbf{a}$.

Demonstração: Seja $L = Df(a)$ *a derivada da função no ponto* $\mathbf{a}$*. Então*

$$f(\mathbf{x}) - f(\mathbf{a}) = L(\mathbf{x}-\mathbf{a}) + R(\mathbf{x}, \mathbf{a}).$$

Para cada $\varepsilon > 0$*, existe* $\delta > 0$ *tal que*

a) $B(\mathbf{a}, \delta) \subset D_f$,

b) $\|R(\mathbf{x}, \mathbf{a})\|_m < \varepsilon \|\mathbf{x}-\mathbf{a}\|_n$ *para* $0 < \|\mathbf{x}-\mathbf{a}\|_n < \delta$

e tal que

c) $\delta < \varepsilon / (\varepsilon + \|L\|)$.

Então, para $0 < \|\mathbf{x}-\mathbf{a}\|_n < \delta$*, temos*

$$\|f(\mathbf{x}) - f(\mathbf{a})\|_m \leq \|L\| \|\mathbf{x}-\mathbf{a}\|_n + \|R(\mathbf{x}, \mathbf{a})\|_m < \|L\| \delta + \varepsilon\delta < \varepsilon,$$

onde $\|L\|$ é a norma do operador L, definida em álgebra Linear[1].■

[1]Nota: Seja $L : \mathbb{R}^n \to \mathbb{R}^m$ um operador linear. Define-se norma de L por

3.2 Derivada segundo um vector e derivada parcial

Agora, seja $\mathbf{e} \in \mathbb{R}^n$ um vector. Consideremos $\mathbf{e}$ como uma direcção: ele determina uma recta que passa por $\mathbf{a} \in \mathbb{R}^n$, constituída pelos pontos da forma $\mathbf{x} = \mathbf{a} + t\mathbf{e}$, $t \in \mathbb{R}$. Se $f : D_f \subset \mathbb{R}^n \to \mathbb{R}^m$ for diferenciável em $\mathbf{a}$, teremos

$$f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + R(\mathbf{x}, \mathbf{a}),$$

ou seja

$$f(\mathbf{a} + t\mathbf{e}) = f(\mathbf{a}) + L((\mathbf{a} + t\mathbf{e}) - \mathbf{a}) + R(\mathbf{a} + t\mathbf{e}, \mathbf{a}).$$

Como L é um operador linear e $\mathbf{a} + t\mathbf{e} - \mathbf{a} = t\mathbf{e}$ tem norma t, obtemos

$$f(\mathbf{a} + t\mathbf{e}) = f(\mathbf{a}) + tL(\mathbf{e}) + R(t), \text{ onde } \|R(t)\|_m / t \to 0, \text{ quando } t \to 0.$$

Em particular, para cada componente f_k $(k = 1, ..., n)$ da função f

$$\lim_{t \to 0} \frac{f_k(\mathbf{a} + t\mathbf{e}) - f_k(\mathbf{a})}{t} = L_k(\mathbf{e}). \tag{3.1.}$$

Definição 3.2 *Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ uma função escalar.* ***A derivada de f segundo o vector*** *$\mathbf{e}$ (ou, no caso em que $\|\mathbf{e}\|_n = 1$,* ***derivada direccional****) num ponto $\mathbf{a} \in D_f$ é*

$$\lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{e}) - f(\mathbf{a})}{t},$$

sempre que este limite exista. Denotamo-la $D_e f(\mathbf{a})$ ou $\dfrac{df}{d\mathbf{e}}(\mathbf{a})$.

A relação (3.1) acima mostra que se f for uma função diferenciável em $\mathbf{a}$, então cada uma das suas componentes f_k tem derivada direccional em todas as direcções em $\mathbf{a}$. Veremos mais à frente que a afirmação inversa é falsa.

Algumas direcções são especiais, nomeadamente as da base canónica $\mathbf{e}_k$ que correspondem às direcções dos eixos ordenados. Às derivadas direccionais de uma função $f : D_f \to \mathbb{R}$ nas direcções $\mathbf{e}_k$ chamamos **derivadas parciais de** f em relação a x_k e denotamo-la habitualmente por $\dfrac{\partial f}{\partial x_k}(\mathbf{a})$ ou por $D_k f(\mathbf{a})$. Então, a derivada parcial em ordem a x_k no ponto $\mathbf{a} = (a_1, ..., a_{k-1}, a_k, a_{k+1}, ..., a_n)$ é dada por

$$\frac{\partial f}{\partial x_k}(\mathbf{a}) = \lim_{t \to 0} \frac{f(a_1, ..., a_{k-1}, a_k + t, a_{k+1}, ..., a_n) - f(a_1, ..., a_n)}{t}, \tag{3.2}$$

$\|L\| = \sup\{\|Lx\|_m : \|x\|_n = 1\}$.

$\|L\|$ é o mais pequeno número para o qual $\|Lx\|_m \leq \|L\| \|x\|_n$.

que é precisamente a derivada da função f, considerada como função de apenas uma variável x_k, estando as outras fixas.

No cálculo da k-ésima derivada parcial num ponto $\mathbf{a} \in \mathbb{R}^n$, estamos de facto a considerar a função g de apenas uma variável, definida por $g(x) = f(a_1, ..., a_{k-1}, x, a_{k+1}, ..., a_n)$ e calculamos

$$g'(a_k) = \lim_{t \to 0} \frac{g(a_k + t) - g(a_k)}{t},$$

como o fazemos para funções de variável real. Podemos aplicar todas as regras de derivação aprendidas em Análise I.

Exercícios 3.3. Calcule as derivadas parciais das seguintes funções escalares :

a) $f(x, y) = 3x^2y + 2x + y^3 - 1$;

b) $f(x, y) = xsen(y)$;

c) $f(x, y) = x^y$;

d) $f(x, y) = \frac{x-y}{x+y}$;

e) $f(x, y, z) = x(y)^z$;

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

Teorema 3.2 *Seja* $f : D_f \subset \mathbb{R}^n \to \mathbb{R}^m$ *uma função diferenciável em* $\mathbf{a}$. *Então* $Df(a)$ *é representada pela matriz* $m \times n$, $A = (a_{ij})$, *onde* $a_{ij} = D_j f_i(\mathbf{a}) = \dfrac{\partial f_i}{\partial x_j}(\mathbf{a})$, $1 \leq i \leq m,\ 1 \leq j \leq n$.

Demonstração: Seja $L = Df(\mathbf{a})$. *De (3.1) temos que*

$$L_i(\mathbf{e}_j) = \lim_{t \to 0} \frac{f_i(\mathbf{a} + t\mathbf{e}_j) - f_i(\mathbf{a})}{t} = D_j f_i(\mathbf{a}) = \frac{\partial f_i}{\partial x_j}.$$

Como L *é um operador linear, este fica completamente determinado pelas imagens dos vectores da base, que são precisamente as derivadas parciais da função* f.

A sua expressão matricial é

$$Df(\mathbf{a}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}, \tag{3.3}$$

com todas as derivadas parciais calculadas no ponto $\mathbf{a}$. A esta matriz chamamos **matriz jacobiana** da função f em $\mathbf{a}$.

Quando $f : \mathbb{R}^n \to \mathbb{R}^n$, a jacobiana é uma matriz quadrada e podemos calcular o seu determinante. A este determinante chamamos o **Jacobiano** de f em $\mathbf{a}$ e denotamo-lo $J_f(\mathbf{a})$. Como sabemos da álgebra Linear, se $J_f(\mathbf{a}) = 0$, então $Df(\mathbf{a})$ não é invertível e isso sugere que f tenha um comportamento irregular numa vizinhança de $\mathbf{a}$. A hipótese $J_f(\mathbf{a}) \neq 0$ previne esta situação. O Jacobiano, para funções vectoriais, vai jogar o mesmo papel que a derivada para funções reais de variável real em teoremas como o da função inversa, por exemplo (ver Cap. Função Inversa e Implícita mais à frente).

Exemplo: A Jacobiana da função

$$f(x, y, z) = \left(y \cos x, z^2 + 3x, 4x + 3y + z\right),$$

no ponto $a = (\pi/2, 0, 2)$, é

$$\begin{aligned} Df(\pi/2, 0, 2) &= \begin{bmatrix} -y \operatorname{sen} x & \cos x & 0 \\ 3 & 0 & 2z \\ 4 & 3 & 1 \end{bmatrix}_{|(x,y,z)=(\pi/2,0,2)} = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 4 \\ 4 & 3 & 1 \end{bmatrix}. \end{aligned}$$

O Jacobiano é

$$J_f(\pi/2, 0, 2) = 0.$$

Atenção, como já dissemos, f ser diferenciável implica que existem todas as derivadas parciais, mas o inverso não é verdade. Uma derivada parcial é apenas <u>parcial.</u> Apesar de podermos conseguir calcular a matriz (3.3), isso, por si só, não nos garante a diferenciabilidade de f. Para f ser diferenciável é necessário verificar a condição de "boa"aproximação linear.

O cálculo da matriz Jacobiana é uma tarefa simples: obtém-se aplicando os métodos do cálculo de funções de uma só variável.

Exercício 3.4: Determine a matriz Jacobiana e o Jacobiano da seguinte função no ponto $a = (0, 1, 2)$

$$f(x, y, z) = \left(x \ln y, e^x y^3 z, 3x + 2y - z\right).$$

A questão agora é saber, para além da existência da matriz Jacobiana, quais as condições adicionais necessárias para garantir a diferenciabilidade de f em a. Vejamos um exemplo que mostra que a simples existência das derivadas parciais não é suficiente para garantir a diferenciabilidade.

Exemplo: Seja $f : \mathbb{R}^2 \to \mathbb{R}$, dada por

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

As derivadas parciais $\frac{\partial f}{\partial x}(0,0)$ e $\frac{\partial f}{\partial y}(0,0)$ existem, mas, no entanto, a função não é diferenciável (na realidade, nem sequer é contínua).

Comecemos por calcular as derivadas parciais em $(x,y) \neq (0,0)$. Aplicamos as regras usuais para funções de uma só variável, neste caso a derivada do quociente.

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{(x^2+y^2)\,y - xy\,(2x)}{(x^2+y^2)^2} = \frac{y^3 - x^2 y}{(x^2+y^2)^2}; \\ \frac{\partial f}{\partial y}(x,y) &= \frac{(x^2+y^2)\,x - xy\,(2y)}{(x^2+y^2)^2} = \frac{x^3 - y^2 x}{(x^2+y^2)^2}. \end{aligned}$$

Não podemos repetir estes cálculos no ponto $(0,0)$, pelo que teremos que usar a definição:

$$\begin{aligned} \frac{\partial f}{\partial x}(0,0) &= \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} 0 = 0; \\ \frac{\partial f}{\partial y}(0,0) &= \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} 0 = 0. \end{aligned}$$

Daqui temos que a candidata mais forte a derivada de f no ponto $(0,0)$ é a matriz $(0,0)$.

Seja

$$\begin{aligned} R(x,y) &= f(x,y) - [f(0,0) + (0,0)\begin{pmatrix} x \\ y \end{pmatrix}], \text{ com } (x,y) \neq (0,0). \\ R(x,y) &= \frac{xy}{(x^2+y^2)}. \end{aligned}$$

Temos que verificar a condição de boa aproximação linear:

$$\lim_{\|(x,y)\|_2 \to 0} \frac{|R(x,y)|}{\|(x,y)\|} \text{ deveria ser zero.}$$

De facto, esta condição não se verifica:

$$\lim_{\|(x,y)\|_2 \to 0} \frac{|R(x,y)|}{\|(x,y)\|} = \lim_{\|(x,y)\|_2 \to (0,0)} \frac{xy}{(x^2+y^2)\sqrt{(x^2+y^2)}}$$

e neste último limite, se tomarmos $y = x$, obtemos o limite direccional

$$\lim_{x \to 0} \frac{x^2}{2x^2\sqrt{2x^2}} = \lim_{x \to 0} \frac{1}{2\sqrt{2}\,|x|} = +\infty,$$

que é diferente de zero, logo o limite pretendido não pode ser 0. Isto prova que f não é diferenciável, apesar de existirem as suas derivadas parciais em todos os pontos de $\mathbb{R}^2$.

Enunciemos agora o teorema que nos dá uma condição suficiente para dada função ser diferenciável. Mas atenção, esta condição não é necessária. Existem funções diferenciáveis com derivadas parciais não contínuas.

Teorema 3.3 *Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}^m$. Suponhamos que todas as derivadas parciais de f, $D_j f_i = \frac{\partial f_i}{\partial x_j}$, existem numa vizinhança de $a \in D_f$ e são contínuas em a. Então f é diferenciável em a.*

A demonstração deste resultado pode ser encontrada em [Webb].

Definição 3.3 *Se $f : D_f \subset \mathbb{R}^n \to \mathbb{R}^m$ tem todas as derivadas parciais $D_j f_i = \frac{\partial f_i}{\partial x_j}$ contínuas em D_f, então dizemos que f é continuamente diferenciável em D_f. Escrevemos $f \in C^1(D_f, \mathbb{R}^m)$ e dizemos, simplesmente, que f é de classe C^1 em D_f.*

Nesta definição, 1 significa a ordem da derivada e C a continuidade.

3.3 Derivada da composta

Tal como para o cálculo das derivadas simples, também para a derivada das funções compostas há regras semelhantes às regras válidas para funções compostas de uma só variável. Comecemos com as operações algébricas sobre as derivadas.

Teorema 3.4. Sejam f e $g : D \subset \mathbb{R}^n \to \mathbb{R}^m$ duas funções diferenciáveis em $\mathbf{a} \in D$. Então $f + g$ e λf (com $\lambda \in \mathbb{R}$) são diferenciáveis em $\mathbf{a}$, sendo

$$\begin{aligned} D(f+g)(\mathbf{a}) &= Df(\mathbf{a}) + Dg(\mathbf{a}) \text{ e} \\ D(\lambda f)(\mathbf{a}) &= \lambda Df(\mathbf{a}). \end{aligned}$$

Demonstração: Sejam f *e* $g : D \to R^m$ *diferenciáveis em* $\mathbf{a} \in D$. *Então*

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x}-\mathbf{a}) + R(\mathbf{x},\mathbf{a}), \\ g(\mathbf{x}) &= g(\mathbf{a}) + Dg(\mathbf{a})(\mathbf{x}-\mathbf{a}) + S(\mathbf{x},\mathbf{a}), \end{aligned}$$

com

$$\begin{aligned} \frac{\|R(\mathbf{x},\mathbf{a})\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} &\to 0, \text{ quando } \|\mathbf{x}-\mathbf{a}\|_n \to 0 \text{ e} \\ \frac{\|S(\mathbf{x},\mathbf{a})\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} &\to 0, \text{ quando } \|\mathbf{x}-\mathbf{a}\|_n \to 0. \end{aligned}$$

Então

$$f(\mathbf{x}) + g(\mathbf{x}) = f(\mathbf{a}) + g(\mathbf{a}) + [Df(\mathbf{a}) + Dg(\mathbf{a})](\mathbf{x}-\mathbf{a}) + R + S$$

e

$$\frac{\|R+S\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} \leq \frac{\|R\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} + \frac{\|S\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} \to 0, \text{ quando } \|\mathbf{x}-\mathbf{a}\|_n \to 0.$$

Como $Df(\mathbf{a}) + Dg(\mathbf{a})$ *é um operador linear, isto demonstra o resultado para* $f+g$. *O outro caso é deixado ao aluno.*$\blacksquare$

Vamos agora ver o que se passa com as funções compostas.

Teorema 3.5. (Regra da cadeia) Sejam duas funções vectoriais $f : D_f \subset \mathbb{R}^k \to \mathbb{R}^m$ e $g : D_g \subset \mathbb{R}^n \to \mathbb{R}^k$, tais que $g(D_g) \subset D_f$. Seja a função $h = f \circ g$, com $h : D_g \subset \mathbb{R}^n \to \mathbb{R}^m$, definida por $h(\mathbf{x}) = f(g(\mathbf{x}))$. Se g for diferenciável em $\mathbf{a} \in D_g$ e f for diferenciável em $\mathbf{y} = g(\mathbf{a}) \in D_f$, então h é diferenciável em $\mathbf{a} \in D_g$ e

$$Dh(\mathbf{a}) = Df(\mathbf{y})\, Dg(\mathbf{a}) = Df(g(\mathbf{a}))\, Dg(\mathbf{a})$$

onde a expressão anterior pode ser vista como a composição de operadores lineares ou como o produto de duas matrizes.

Demonstração: Sejam $L = Dg(\mathbf{a})$ *e* $M = Df(y)$, *onde* $\mathbf{y} = g(\mathbf{a})$. *Como, escrevendo* $h(\mathbf{x}) = h(\mathbf{a}) + Dh(\mathbf{a})(\mathbf{x}-\mathbf{a}) + R(\mathbf{x},\mathbf{a})$, *a condição para a diferenciabilidade de* h *é que*

$$\frac{\|R(\mathbf{x},\mathbf{a})\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} \to 0, \text{ quando } \|\mathbf{x}-\mathbf{a}\|_n \to 0,$$

basta-nos provar que

$$\lim_{\|\mathbf{x}-\mathbf{a}\|_n \to 0} \frac{\|h(\mathbf{x}) - h(\mathbf{a}) - Dh(\mathbf{a})(\mathbf{x}-\mathbf{a})\|_m}{\|\mathbf{x}-\mathbf{a}\|_n} = 0.$$

Para isso bastar-nos-á provar que para cada $\varepsilon > 0$, *existe* $\delta > 0$ *tal que*

$$\|h(\mathbf{x}) - h(\mathbf{a}) - Dh(\mathbf{a})(\mathbf{x}-\mathbf{a})\|_m < C\varepsilon \|\mathbf{x}-\mathbf{a}\|_n,$$

para uma certa constante C.

Escrevamos

$$\begin{aligned} & h(\mathbf{x}) - h(\mathbf{a}) - M\ L(\mathbf{x}-\mathbf{a}) = \\ = \ & f(g(\mathbf{x})) - f(g(\mathbf{a})) - M(g(\mathbf{x}) - g(\mathbf{a})) + M(g(\mathbf{x}) - g(\mathbf{a}) - L(\mathbf{x}-\mathbf{a})). \end{aligned}$$

Pela desigualdade triangular, temos então

$$\begin{aligned} & \|h(\mathbf{x}) - h(\mathbf{a}) - M\ L(\mathbf{x}-\mathbf{a})\|_m \leq \\ \leq \ & \|f(g(\mathbf{x})) - f(g(\mathbf{a})) - M(g(\mathbf{x}) - g(\mathbf{a}))\|_m + \|M(g(\mathbf{x}) - g(\mathbf{a}) - L(\mathbf{x}-\mathbf{a}))\|_m. \end{aligned}$$

Como g *é diferenciável, para o segundo termo temos a existência de* $\delta_1 > 0$ *tal que*

$$\begin{aligned} \|M(g(\mathbf{x}) - g(\mathbf{a}) - L(\mathbf{x}-\mathbf{a}))\|_m & \leq \|M\| \, \|g(\mathbf{x}) - g(\mathbf{a}) - L(\mathbf{x}-\mathbf{a})\|_k < \varepsilon \|M\| \, \|\mathbf{x}-\mathbf{a}\|_n, \\ \textit{para } \|\mathbf{x}-\mathbf{a}\|_n & < \delta_1. \end{aligned}$$

Tratemos do primeiro termo.

Na demonstração da continuidade de uma função diferenciável, usa-se implicitamente o seguinte facto (tome-se nessa demonstração $\varepsilon = 1$*):*

para um certo $\delta_2 > 0$ *tem-se* $\|g(\mathbf{x}) - g(\mathbf{a})\|_k \leq (\|L\| + 1)\|x - a\|_n$ *para* $x \in B(\mathbf{a},\delta_2)$.

Agora, como f *é diferenciável em* y, *existe* $\delta_3 > 0$ *tal que*

$$\|f(z) - f(y) - M(z-y)\|_m < \varepsilon \|z-y\|_k, \ \textit{para } \|z-y\|_k < \delta_3.$$

Escolhendo δ_4 *tal que* $\delta_4 < \delta_2$ e $(\|L\| + 1)\,\delta_4 < \delta_3$ *vemos que é possível para* $x \in B(a, \delta_4)$ *tomar* $z = g(x)$ *e* $y = g(a)$ *para obter*

$$\|f(g(x)) - f(g(a)) - M(g(x) - g(a))\|_m < \varepsilon\,\|g(x) - g(a)\|_k < \varepsilon\,(\|L\| + 1)\,\|x - a\|_n$$

para $\|x - a\|_n < \delta_4$. *Tomando* $\delta = \min\{\delta_1, \delta_4\}$ *obtemos o resultado pretendido para* $C = \|M\| + \|L\| + 1.\blacksquare$

Exemplo. Seja $g(x, y) = (x^2 - y^2 + xy, y^2 - 1)$ e $f(u, v) = (u + v, 2u, v^2)$. Prove que f e g são diferenciáveis, que $f \circ g$ existe e calcule a derivada de $f \circ g$ no ponto $(1, 1)$:

$a)$ directamente;

$b)$ pela regra da cadeia.

Resolução: $g(x, y) = (g_1(x, y), g_2(x, y))$, com $g_1(x, y) = x^2 - y^2 + xy$ e $g_2(x, y) = y^2 - 1$. Então $\frac{\partial g_1}{\partial x} = 2x + y$, $\frac{\partial g_1}{\partial y} = -2y + x$, $\frac{\partial g_2}{\partial x} = 0$, $\frac{\partial g_2}{\partial y} = 2y$. Todas as derivadas parciais são contínuas, logo a função é diferenciável, $g \in C^1(\mathbb{R}^2, \mathbb{R}^2)$. Igualmente para $f \in C^1(\mathbb{R}^2, \mathbb{R}^3)$. Então $h = f \circ g$ está definida em todo o $\mathbb{R}^2$ e tem valores em $\mathbb{R}^3$. De facto

$$\begin{aligned} (f \circ g)(x, y) &= f(g_1(x, y), g_2(x, y)) = \left(g_1 + g_2, 2g_1, g_2^2\right) = \\ &= \left(x^2 + xy - 1, 2x^2 - 2y^2 + 2xy, y^4 - 2y^2 + 1\right) = (h_1, h_2, h_3). \end{aligned}$$

$a)$ Calculando directamente

$$\begin{aligned} Dh(x, y) &= \begin{bmatrix} \frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial y} \\ \frac{\partial h_2}{\partial x} & \frac{\partial h_2}{\partial y} \\ \frac{\partial h_3}{\partial x} & \frac{\partial h_3}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x + y & x \\ 4x + 2y & 2x - 4y \\ 0 & 4y^3 - 4y \end{bmatrix} \\ Dh(1, 1) &= \begin{bmatrix} 3 & 1 \\ 6 & -2 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

$b)$ Agora usamos a regra da cadeia.

Como $g(1, 1) = (1, 0)$ temos

$$D(f \circ g)(1, 1) = Df(1, 0)\,Dg(1, 1)$$

ou seja

$$\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \\ \frac{\partial f_3}{\partial u} & \frac{\partial f_3}{\partial v} \end{bmatrix}_{|(u,v)=(1,0)} \cdot \begin{bmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{bmatrix}_{|(x,y)=(1,1)} =$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 2v \end{bmatrix}_{|(u,v)=(1,0)} \cdot \begin{bmatrix} 2x+y & -2y+x \\ 0 & 2y \end{bmatrix}_{|(x,y)=(1,1)} =$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 6 & -2 \\ 0 & 0 \end{bmatrix}.$$

Exemplo: Sejam $g : \mathbb{R}^3 \to \mathbb{R}^3$ e $f : \mathbb{R}^3 \to \mathbb{R}$. Seja $h = f \circ g : \mathbb{R}^3 \to \mathbb{R}$. A derivada de h, será a matriz 1×3,

$$Dh(x) = Df(g(x))\ Dg(x),$$

produto da matriz 1×3, $Df(g(x))$ pela matriz 3×3, $Dg(x)$. Componente a componente obtemos, se chamarmos y_j a $g_j(x)$,

$$\frac{\partial h}{\partial x_i} = \frac{\partial f}{\partial y_1}\frac{\partial y_1}{\partial x_i} + \frac{\partial f}{\partial y_2}\frac{\partial y_2}{\partial x_i} + \frac{\partial f}{\partial y_3}\frac{\partial y_3}{\partial x_i}.$$

Obtendo assim uma espécie de cadeias, cujos elos são as componentes de $g = y$: se quero derivar em ordem a x_i, derivo f em ordem a y_j e y_j em ordem a x_i e estabeleço todas as cadeias possíveis que levam à variável considerada, x_i.

Exemplo: Sejam $g(x,y,z) = (x+y, 2z, 4x+5z)$ e $f(u,v,w) = u + 2v + 3w$. A composta $h = f \circ g$ verifica, considerando $u = x + y$, $v = 2z$ e $w = 4x + 5z$,

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} = 1 \times 1 + 2 \times 0 + 3 \times 4 = 13; \\ \frac{\partial h}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial y} = 1 \times 1 + 2 \times 0 + 3 \times 0 = 1; \\ \frac{\partial h}{\partial z} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial z} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial z} = 1 \times 0 + 2 \times 2 + 3 \times 5 = 19. \end{aligned}$$

Exercícios 3.5:

1. Calcule:

   a) $\frac{dz}{dt}$, se $z = \frac{x}{y}$, onde $x = e^t, y = \ln t$;

   b) $\frac{du}{dt}$, se $u = \ln \sin \frac{x}{\sqrt{y}}$, onde $x = 3t^2, y = \sqrt{t^2+1}$;

   c) $\frac{du}{dt}$, se $u = \frac{z}{\sqrt{x^2+y^2}}$, onde $x = R\cos t, y = R\sin t, z = H$;

   d) $\frac{\partial z}{\partial u}$ e $\frac{\partial z}{\partial v}$, se $z = \operatorname{Arctg} \frac{x}{y}$, onde $x = u\sin v, y = u\cos v$;

   e) $\frac{\partial z}{\partial x}$ e $\frac{\partial z}{\partial y}$, se $z = \cos u$, onde $u = xy + \frac{y}{x}$

2. Verifique a regra da cadeia para as seguintes funções:

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

   b) $f(x,y) = xe^{x^2+y^2}$, $x = t$, $y = -t$;

   c) $f(x,y) = x + y^2 + z^3$, $x = \cos t$, $y = \sin t$, $z = t$;

   d) $u = \sin(xy)$, $x = t^2 + t$, $y = t^3$.

3. (Exercício de teste) Considere as duas seguintes funções de $\mathbb{R}^2$ em $\mathbb{R}^2$

$$\begin{aligned} f(u,v) &= \left(uv, e^{u+v}\right); \\ g(x,y) &= (2x+3y, x-y). \end{aligned}$$

   Considere a função composta $h : \mathbb{R}^2 \to \mathbb{R}^2$, dada por $h = (h_1, h_2) = f \circ g$. Assinale a opção verdadeira

   ☐ *a)* $\dfrac{\partial h_1}{\partial x} = 3e^{3x+2y}$ ☐ *c)* $\dfrac{\partial h_2}{\partial y} = 3e^{3x+2y}$

   ☐ *b)* $\dfrac{\partial h_1}{\partial y} = 3e^{3x+2y}$ ☐ *d)* $\dfrac{\partial h_2}{\partial x} = 3e^{3x+2y}$

3.4 Gradiente, divergência e rotacional

Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ uma função escalar diferenciável num certo ponto $x \in D_f$. A derivada de f será um operador linear, identificado com a matriz $1 \times n$, a que habitualmente chamamos vector (de $\mathbb{R}^n$). A este vector chamamos **gradiente de** f

em x e denotamo-lo *grad* $f(x)$ ou $\nabla f(x)$. A forma como este operador age sobre vectores de $\mathbb{R}^n$ é definida tal como se se tratase do produto interno entre vectores (caso especial do produto de matrizes). Mais precisamente

$$Df(x)\,h = (\nabla f(x), h)_n, \text{ para todo o } h \in \mathbb{R}^n.$$

Ao escalar $Df(a)\,h = (\nabla f(a), h)_n$, para $h \in \mathbb{R}^n$ chamamos **diferencial** ou **diferencial total** de f no ponto a, relativo ao vector $h = (h_1, h_2, ..., h_n)$. Geralmente o vector h significa um acréscimo infinitesimal nas variáveis $(x_1, x_2, ..., x_n)$, $h = (dx_1, dx_2, ..., dx_n)$.

Exemplo: O gradiente da função

$$f(x, y, z, w) = e^x + \cos(y + 2z) + 2xw$$

é o vector

$$\nabla f(x, y, z, w) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial w}\right) = (e^x + 2w, -\operatorname{sen}(y + z), -2\operatorname{sen}(y + z), 2x).$$

Calculado no ponto $(0, \pi/4, \pi/8, 3)$ fica

$$\nabla f(0, \pi/4, \pi/8, 3) = (1 + 2 \times 3, -\operatorname{sen}(\pi/2), -2\operatorname{sen}(\pi/2), 0) = (7, -1, -2, 0).$$

Para saber como age sobre o vector $(0.1, 0.1, 0.1, 0.1)$, calcula-se

$$\nabla f(0, \pi/4, \pi/8, 3) \cdot (0.1, 0.1, 0.1, 0.1) = (7, -1, -2, 0) \cdot (0.1, 0.1, 0.1, 0.1) = 0.4.$$

Usando o vector gradiente, é possível calcular as derivadas direccionais de forma mais eficiente do que o cálculo do limite, segundo a definição.

Teorema 3.6. Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ diferenciável em $\mathbf{a} \in D_f$. Então, para cada direcção $\mathbf{e}$, a derivada na direcção de $\mathbf{e}$, no ponto a, $D_e f(\mathbf{a})$, existe e além disso

$$D_{\mathbf{e}} f(\mathbf{a}) = Df(\mathbf{a})\,\mathbf{e} = \nabla f(\mathbf{a}) \cdot \mathbf{e}.$$

Demonstração: Como f *é diferenciável, podemos escrever*

$$f(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a}) + R(\mathbf{x}, \mathbf{a}),$$

com $\frac{|R(\mathbf{x},\mathbf{a})|}{\|\mathbf{x}-\mathbf{a}\|_n} \to 0$, *quando* $\|\mathbf{x}-\mathbf{a}\|_n \to 0$. *Quando escolhemos pontos* x *sobre a recta* $\mathbf{a}+t\mathbf{e}$, $t \in \mathbb{R}$, *esta relação fica (ver definição de derivada direccional)*

$$f(\mathbf{a}+te) = f(\mathbf{a}) + tDf(\mathbf{a})\,\mathbf{e} + R(t)\,,\ \text{onde}\,\|R(t)\|_n/t \to 0,\ \text{quando}\ t \to 0$$

então

$$\frac{f(\mathbf{a}+t\mathbf{e}) - f(\mathbf{a})}{t} = Df(\mathbf{a})\,\mathbf{e} + \frac{R(t)}{t}$$

ou seja, passando ao limite,

$$D_{\mathbf{e}}f(\mathbf{a}) = Df(\mathbf{a})\,\mathbf{e} = \nabla f(\mathbf{a}) \cdot \mathbf{e}.\blacksquare$$

ATENÇÃO: As derivadas segundo um vector podem existir, sem que a função seja diferenciável. O Teorema 3.6. dá-nos apenas uma ferramenta mais simples para o seu cálculo, quando aplicável.

Exercício 3.6.: Calcule a derivada direccional de cada função na direcção e e no ponto P indicados:

a) $f(x,y) = x^2 + y^2 - 3xy^3, \quad P = (1,2)\,, \quad \mathbf{e} = \left(1/2, \sqrt{3}/2\right);$

b) $f(x,y) = e^x \cos y, \quad P = (0, \pi/4)\,, \quad \mathbf{e} = \left(1/\sqrt{10}, 3/\sqrt{10}\right);$

c) $f(x,y) = 17x^y, \quad P = (1,1)\,, \quad \mathbf{e} = \left(1/\sqrt{2}, 1/\sqrt{2}\right);$

d) $f(x,y) = e^{x^2} \cos y, \quad P = (1, \pi/2)\,, \quad \mathbf{e} = (3/5, 4/5)\,;$

e) $f(x,y,z) = x^2 - 2xy + 3z^2, \quad P = (1,1,2)\,, \quad \mathbf{e} = \left(1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}\right);$

f) $f(x,y,z) = e^{-(x^2+y^2+z^2)}, \quad P = (1,10,100)\,, \quad \mathbf{e} = (1,-1,-1)/\sqrt{3};$

g) $f(x,y,z) = \sin(xyz), \quad P = (1,1,\pi/4)\,, \quad \mathbf{e} = \left(1/\sqrt{2}, 0, -1/\sqrt{2}\right);$

h) $f(x,y,z) = 1/(x^2+y^2+z^2), \quad P = (2,3,1)\,, \quad \mathbf{e} = (1,1,-2)/\sqrt{6}.$

Uma aplicação do gradiente de uma função escalar é determinar as direcções de maior e menor crescimento de f.

A derivada direccional $D_e f$ mede a taxa de crescimento de f na direcção de $\mathbf{e}$ (sabemos isso das funções de variável real). Que direcção dará a maior taxa de crescimento? Temos, pela desigualdade de Cauchy-Schwartz,

$$\nabla f(\mathbf{a}) \cdot \mathbf{e} \leq \|\nabla f(\mathbf{a})\|_n \|\mathbf{e}\|_n$$

verificando-se a igualdade apenas no caso de $\mathbf{e}$ ser um múltiplo escalar de $\nabla f(\mathbf{a})$[2]. Ou seja, quando o coseno do ângulo formado entre o gradiente de f e o vector $\mathbf{e}$, é 1, isto é, quando o ângulo é 0.

Isto significa que $D_e f(\mathbf{a})$ é máximo, quando $\mathbf{e} = \dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_n}$ (lembremos que $\|\mathbf{e}\|_n = 1$). Por outras palavras, a direccção de maior crescimento é a determinada pelo gradiente. Similarmente a direcção de menor crescimento é a determinada por $-\nabla f(\mathbf{a})$.

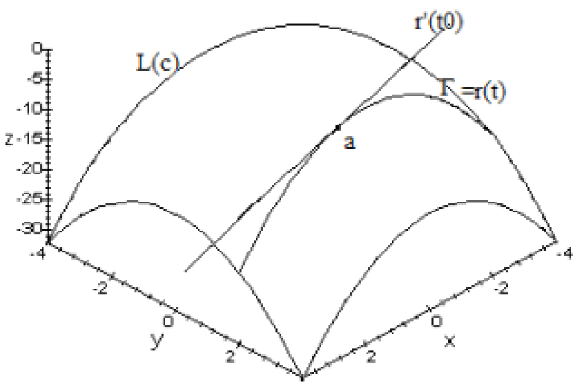
Conjuntos de nível e planos tangentes A regra da cadeia pode utilizar-se para deduzir propriedades geométricas do vector gradiente. Tomemos $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ uma função escalar e tomemos o conjunto dos pontos onde f toma um certo valor c constante

$$L(c) = \{x \in D_f : f(x) = c\}.$$

A estes conjuntos chamamos **conjuntos de nível de** f. Se $D_f \subset \mathbb{R}^2$, chamamos curva de nível e se $D_f \subset \mathbb{R}^3$, chamamos superfície de nível.

Suponhamos que f é diferenciável em $D_f \subset \mathbb{R}^3$ e examinemos uma superfície de nível, $L(c)$. Seja $\mathbf{a}$ um ponto de $L(c)$ e Γ uma curva situada em $L(c)$ passando por $\mathbf{a}$. Suponhamos que Γ é parametrizada pela função r, definida num intervalo real J e derivável. Como $\Gamma \subset L(c)$, temos

$$f(r(t)) = c \text{ para todo o } t \in J.$$



[2]Uma forma de definir ângulo entre dois vectores de um espaço vectorial com produto interno é $\theta = \measuredangle(uv) = \arccos \dfrac{(u,v)_n}{\|u\|_n \|v\|_n}$, de tal forma que $(u,v)_n = \|u\|_n \|v\|_n \cos\theta$.(ver[LM])

Se definirmos g como a função composta $g(t) = f(r(t))$, a regra da cadeia dá

$$g'(t) = \nabla f(r(t)) \cdot r'(t).$$

Sendo g constante em J, $g'(t) = 0$ em J e, para t_0 escolhido tal que $r(t_0) = \mathbf{a}$, temos

$$\nabla f(\mathbf{a}) \cdot r'(t_0) = 0.$$

Por outras palavras, o gradiente de f em a é perpendicular ao vector tangente à curva Γ no ponto a, $r'(t_0)$

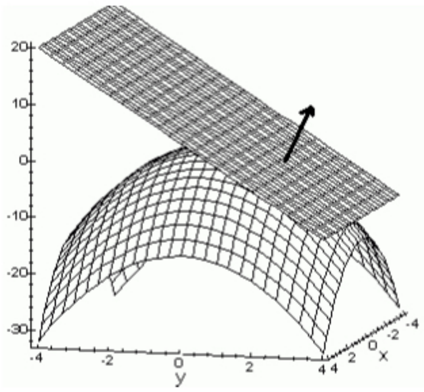
Tomando uma família de curvas na superfície $L(c)$, passando por $\mathbf{a}$, podemos, de acordo com o que foi feito acima, concluir que os vectores tangentes a estas curvas são todos perpendiculares a $\nabla f(\mathbf{a})$. Se $\nabla f(\mathbf{a})$ não for o vector nulo, os vectores tangentes a $L(c)$ em a definem um plano, para o qual $\nabla f(\mathbf{a})$ é o vector normal. Chamamos a este plano o **plano tangente a** $L(c)$ em $\mathbf{a}$.

Sabemos da geometria analítica que a equação de um plano passando por $\mathbf{a}$, com vector normal $\overrightarrow{\mathbf{n}}$ é dada por

$$\overrightarrow{\mathbf{n}} \cdot (\mathbf{x} - \mathbf{a}) = 0$$

ou seja, neste caso

$$\nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = 0.$$



Se quisermos em termos de coordenadas, sendo $\mathbf{x} = (x, y, z)$, $\mathbf{a} = (a_1, a_2, a_3)$

$$\frac{\partial f}{\partial x}(a)(x - a_1) + \frac{\partial f}{\partial y}(a)(y - a_2) + \frac{\partial f}{\partial z}(a)(z - a_3) = 0.$$

Exemplo: Calculemos a equação do plano tangente à superfície definida por $z = 3 - 2x^2 - y^2$ no ponto $(1, -1, 0)$.

Escrevamos a superfície na forma $f(x,y,z)=0$, com $f(x,y,z)=3-2x^2-y^2-z$. Calculemos ∇f:

$$\frac{\partial f}{\partial x} = -4x, \qquad \frac{\partial f}{\partial y} = y+1, \qquad \frac{\partial f}{\partial z} = -1$$

Um vector normal à superfície é $\nabla f(1,-1,0)=(-4,2,-1)$ e a equação do plano tangente será

$$\nabla f(1,-1,0)\cdot(x-1,y+1,z)=0 \Leftrightarrow (-4,2,-1)\cdot(x-1,y+1,z)=0 \Leftrightarrow z=-4x+2y+6$$

Definamos agora outros dois operadores diferenciais: a divergência e o rotacional.

Definição 3.4 *Seja* $f: D_f \subset \mathbb{R}^3 \to \mathbb{R}^3$ *uma função vectorial diferenciável.*

Ao campo escalar, definido por

$$\operatorname{div} f = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

*chamamos **divergência de** f e denotamo-lo* $\operatorname{div} f$.

Ao campo vectorial, definido por

$$\operatorname{rot} f = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right)$$

*chamamos **rotacional de** f e denotamo-lo* $\operatorname{rot} f$.

Outra forma, equivalente, de definir rotacional é através do determinante de uma matriz especial, que representa o produto externo entre ∇ e f, olhando para ∇ como o "vector"$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$:

$$\operatorname{rot} f = \begin{vmatrix} \overrightarrow{e_1} & \overrightarrow{e_2} & \overrightarrow{e_3} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{vmatrix} = \overrightarrow{e_1}\left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\right) - \overrightarrow{e_2}\left(\frac{\partial f_3}{\partial x} - \frac{\partial f_1}{\partial z}\right) + \overrightarrow{e_3}\left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right).$$

A divergência é o traço da matriz Jacobiana, ou seja, a soma dos elementos da diagonal principal.

Exercícios 3.7.:

1. Calcule o gradiente das seguintes funções:

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

b) $f(x,y,z) = xy + yz + xz$;

c) $f(x,y) = \ln\left(\sqrt{x^2+y^2}\right)$;

d) $f(x,y) = xe^{x^2+y^2}$;

e) $f(x,y) = xe^{xy^3+3}$.

2. Calcule a divergência dos campos vectoriais obtidos nas perguntas 1.a) e 1.b). Ao resultado desta operação chamamos Laplaciano de f.

3. Calcule o rotacional dos campos vectoriais obtidos nas perguntas 1.a) e 1.b).

3.5 Derivadas parciais de ordem superior

Definimos as derivadas parciais de uma função escalar. Se estas derivadas existirem, elas são também funções escalares, que poderão igualmente ser diferenciáveis. Às derivadas parciais das derivadas parciais de uma função, chamamos **derivadas parciais de segunda ordem** da função. Denotamo-las

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Se $i = j$, então escrevemos

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_i^2}.$$

Exemplo: Calculemos as derivadas parciais de 2^a ordem da função escalar $f(x,y) = xy^2 + xe^y$. Comecemos pelas de 1^a ordem:

$$\begin{aligned} \frac{\partial f}{\partial x} &= y^2 + e^y, & \frac{\partial f}{\partial y} &= 2xy + xe^y, \\ \frac{\partial^2 f}{\partial x^2} &= 0, & \frac{\partial^2 f}{\partial y^2} &= 2x + xe^y, \\ \frac{\partial^2 f}{\partial x \partial y} &= 2y + e^y, & \frac{\partial^2 f}{\partial y \partial x} &= 2y + e^y. \end{aligned}$$

Como podemos observar as últimas derivadas de segunda ordem, chamadas derivadas parciais mistas, por envolverem derivadas parciais em ordem a diferentes variáveis, são iguais. Tal igualdade nem sempre se verifica, como é possível ver no exemplo seguinte.

Exemplo: A função

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

é diferenciável em $(0,0)$, mas as derivadas parciais mistas não são iguais em $(0,0)$.

Calculemos as de 1^a ordem:

$$\begin{aligned}\frac{\partial f}{\partial x}(0,0) &= \lim_{t\to 0}\frac{f((0,0)+t(1,0))-f(0,0)}{t}=\\ &= \lim_{t\to 0}\frac{f(t,0)-0}{t}=\lim_{t\to 0}0=0\end{aligned}$$

e

$$\begin{aligned}\frac{\partial f}{\partial y}(0,0) &= \lim_{t\to 0}\frac{f((0,0)+t(0,1))-f(0,0)}{t}=\\ &= \lim_{t\to 0}\frac{f(0,t)-0}{t}=\lim_{t\to 0}0=0.\end{aligned}$$

Para calcular as derivadas de segunda ordem, precisamos de calcular as derivadas de $1\ ^a$ ordem em $(x,y)\neq(0,0)$.

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \frac{\left(x^2+y^2\right)\frac{\partial}{\partial x}\left(x^3y-xy^3\right)-\left(x^3y-xy^3\right)\frac{\partial}{\partial x}\left(x^2+y^2\right)}{\left(x^2+y^2\right)^2}=\\ &= \frac{\left(x^2+y^2\right)\left(3x^2y-y^3\right)-\left(x^3y-xy^3\right)2x}{\left(x^2+y^2\right)^2}.\\ \frac{\partial f}{\partial y}(x,y) &= \frac{\left(x^2+y^2\right)\frac{\partial}{\partial y}\left(x^3y-xy^3\right)-\left(x^3y-xy^3\right)\frac{\partial}{\partial y}\left(x^2+y^2\right)}{\left(x^2+y^2\right)^2}=\\ &= \frac{\left(x^2+y^2\right)\left(x^3-3xy^2\right)-\left(x^3y-xy^3\right)2y}{\left(x^2+y^2\right)^2}.\end{aligned}$$

Assim

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

e

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

Antes de passarmos ao teorema que nos dá as condições suficientes para que a igualdade se verifique, introduzamos a seguinte notação:

Notação: Dizemos que f é de classe C^2 e escrevemos $f \in C^2(D_f)$ se $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ e existirem todas as derivadas parciais de segunda ordem e forem contínuas. Note que $f \in C^2(D_f) \Rightarrow f \in C^1(D_f)$, visto que $f \in C^2$ quer dizer $\dfrac{\partial f}{\partial x_i} \in C^1$, então $\dfrac{\partial f}{\partial x_i}$ é diferenciável , logo contínua.

Teorema 3.7. Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R},\ f \in C^2(D_f)$. Então, para todo o $a \in D_f$ e para todo o i, j com $0 \leq i, j \leq n$, $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$.

A demonstração completa deste resultado baseia-se no torema do valor médio para funções de variável real e pode ser consultado em [Webb]. Em jeito de esboço o que se faz é definir uma função de apenas duas variáveis, aquelas em ordem às quais se quer derivar. Em seguida definir duas funções, correspondentes ao acréscimo em cada uma das duas variáveis pré-fixadas e aplicar a cada uma delas o Teorema do valor médio para funções de variável real. Depois usa-se a continuidade das derivadas parciais de $2^{\ a}$ ordem.

Exercícios 3.8.: Comprove a validade do Teorema 3.7. nos seguintes casos:

a) $f(x, y) = 3x^2y + 2x + y^3 - 1$;

b) $f(x, y) = sen(xy^2)$;

c) $f(x, y) = x^y$;

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

Tendo definido as derivadas de segunda ordem, podemos definir o operador Laplaciano e a matriz Hessiana.

Definição 3.5 *Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ uma função escalar para a qual existem todas as derivadas parciais de segunda ordem*

1. *Definimos o **Laplaciano** de f por*

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}.$$

2. *Definimos a **matriz Hessiana** de f por*

$$H(f) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

3. *Chamamos **Hessiano** de f ao determinante da matriz Hessiana.*

Exercícios 3.9.: Quais das seguintes funções satisfazem a equação de Laplace $\Delta f = 0$? Tais funções dizem-se funções harmónicas.

a) $f(x,y) = x^2 - y^2$;

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

c) $f(x,y) = xy$;

d) $f(x,y) = \sin x \ \cosh y$;

e) $f(x,y) = e^x \sin y$.

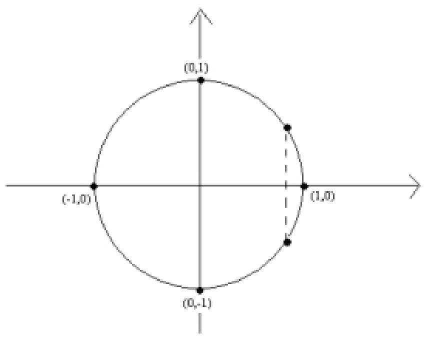
3.6 Funções inversa e implícita

(Seguimos principalmente o texto de [Webb], cap.4) Neste capítulo vamos tratar de dois teoremas muito importantes da Análise e que respondem à questão de ser ou não possível obter uma variável y explicitamente em função de uma outra x quando estas são relacionadas por uma expressão do tipo

$$F(x,y) = 0. \tag{3.6.1}$$

Se tal for possível dizemos que $F(x,y)$ define y implicitamente como função de x. Formalmente isto significa que existe uma função f tal que $F(x,y) = 0$ é equivalente a $y = f(x)$. Evidentemente não se pode esperar que sejamos capazes de escrever uma fórmula para f, mas devemos ser capazes de encontrar para cada x um e um só y que lhe corresponda. Se estivermos a trabalhar com uma curva em $\mathbb{R}^2$ o que se pretende é averiguar se o conjunto de pontos que satisfaz a equação (3.6.1) é ou não o gráfico de uma função real de variável real f. Observemos o exemplo seguinte:

Exemplo 1. Consideremos o conjunto de pontos $(x, y) \in \mathbb{R}^2$ que verificam $x^2 + y^2 - 1 = 0$, ou seja a circunferência centrada na origem e raio 1.



Este exemplo mostra que nem sempre é possível estabelecer essa relação: para cada x , com $-1 < x < 1$ existem dois y's que satisfazem a relação. A curva $x^2 + y^2 = 1$ é composta pelos gráficos de duas funções distintas $f_1(x) = \sqrt{1 - x^2}$ e $f_2(x) = -\sqrt{1 - x^2}$, com $-1 \leq x \leq 1$. No entanto, se tomarmos qualquer ponto da circunferência e nos concentrarmos numa vizinhança desse ponto, estaremos perante uma parte do gráfico de uma das funções f_1 ou f_2, excepto nos pontos $(x, y) = (-1, 0)$ e $(x, y) = (1, 0)$, que são pontos que estão na fronteira das duas curvas. Portanto estaremos sempre a pensar num resultado local e temos que excluir os pontos onde a tangente é vertical. A escolha desses pontos problemáticos pode fazer-se usando a derivada. O Teorema da Função Implícita é um resultado geral que permite responder à questão mesmo para um número de variáveis superior a dois. Este teorema fornece as condições apropriadas para que uma equação do tipo $F(x, y) = 0$ seja localmente equivalente a $y = f(x)$, para alguma função f, e para que as propriedades de diferenciabilidade de F passem para f.

3.6.1 Teorema da função inversa

O Teorema da Função Inversa é, de facto, um caso particular do Teorema da Função Implícita[1] mas é geralmente demonstrado à parte e usado para demonstrar o Teorema da Função Implícita. Aqui o problema é saber se dada função f tem inversa f^{-1} e saber em que medida é que as propriedades de diferenciabilidade de uma implicam

[1]Se considerarmos $F(x, y) = f(x) - y = 0$ e resolvermos para x em termos de y o que estamos a fazer é determinar a inversa de f.

as da outra. Como se sabe da Análise real, uma função tem inversa sse for injectiva num dado conjunto S e, nesse caso, a função inversa f^{-1} será definida de $f(S)$ em S. Se considerarmos a função $y = (x^2+1)e^x$ é fácil encontrar y, dado um certo x, mas já o contrário não é evidente, por exemplo, encontrar x tal que $(x^2+1)e^x = 1$. Na realidade precisamos saber se a função é injectiva e, nesse caso, a inversa existirá. A discussão sobre a diferenciabilidade da inversa pode nem sequer fazer sentido, uma vez que $f(S)$ pode nem ter pontos interiores (e então?). Mas se admitirmos que a função é diferenciável, é bem aproximada pela derivada e podemos esperar que as propriedades da derivada sejam usadas como propriedades locais da função. De facto o Teorema da Função Inversa estabelece que a invertibilidade da derivada de f num ponto, $Df(a)$, implica invertibilidade local de f e a diferenciabilidade da inversa f^{-1}, com derivada igual a $[Df(a)]^{-1}$.

A demonstração destes dois resultados, que enunciaremos em seguida, não é elementar e ultrapassa o âmbito da disciplina. No entanto, para quem tiver interesse, recomenda-se a leitura de [Webb].

Teorema da função inversa*:* Seja V um aberto de $\mathbb{R}^n$ e seja $f \in C^1(V, \mathbb{R}^n)$. Seja $a \in V$ e suponhamos que o Jacobiano $J_f(a)$ é diferente de zero. Então existe uma vizinhança $U \subseteq V$ de a, tal que $f|_U$ é invertível. A função inversa $g := (f|_U)^{-1}$ é de clase C^1 num conjunto aberto $W \subseteq f(U)$ e $Dg(y) = [Df(x)]^{-1}$, onde $y = f(x)$.

Exemplo 2. Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x,y) = (x + y + e^{x+y}, x^2 + 2y)$. Queremos saber se esta função é invertível numa vizinhança de $(a_1, a_2) \in \mathbb{R}^2$ ou seja, saber se a equação

$$(u,v) = f(x,y)$$

tem solução única numa vizinhança de $f(a_1,a_2)$. Saber se dado (u,v) em $V(f(a_1,a_2))$ existe $(x,y) \in V(a_1,a_2)$ tal que

$$(u,v) = \left(x + y + e^{x+y}, x^2 + 2y\right).$$

Esta última igualdade pode escrever-se na forma de um sistema de equações

$$\begin{aligned} u &= x + y + e^{x+y} \\ v &= x^2 + 2y. \end{aligned}$$

Como aplicar o teorema?

Verifiquemos as condições: f é claramente uma função diferenciável em todo o $\mathbb{R}^2$ e com derivadas contínuas. Tomemos $a = (a_1, a_2) \in \mathbb{R}^2$ qualquer e calculemos o Jacobiano nesse ponto :

$$\begin{aligned}
f &= (f_1, f_2) = \left(x + y + e^{x+y}, x^2 + 2y\right). \\
J_f(a) &= \det \begin{bmatrix} \frac{\partial f_1}{\partial x}(a) & \frac{\partial f_1}{\partial y}(a) \\ \frac{\partial f_2}{\partial x}(a) & \frac{\partial f_2}{\partial y}(a) \end{bmatrix} = \\
&= \det \begin{bmatrix} 1 + e^{x+y} & 1 + e^{x+y} \\ 2x & 2 \end{bmatrix}_{|(x,y)=(a_1,a_2)} = \\
&= 2 + 2e^{a_1+a_2} - 2a_1 - 2a_1 e^{a_1+a_2} = \\
&= 2 - 2a_1 + (2 - 2a_1) e^{a_1+a_2} = \\
&= (2 - 2a_1)(1 + e^{a_1+a_2}).
\end{aligned}$$

O primeiro termo da última expressão é igual a zero apenas quando $a_1 = 1$ e o segundo é sempre diferente de zero. Isto quer dizer que para todos os pontos (a_1, a_2), com $a_1 \neq 1$, existem uma vizinhança U de (a_1, a_2) e uma vizinhança W de $(a_1 + a_2 + e^{a_1+a_2}, a_1^2 + 2a_2)$ em que $f_{|U} : U \to W$ é invertível. Por exemplo $(a_1, a_2) = (0, 0)$ garante a existência de inversa com domínio igual a uma vizinhança de $f(0, 0) = (1, 0)$. A derivada dessa função inversa $\left(f_{|U}\right)^{-1} := g$ em $(1, 0)$ é dada por:

$$\begin{aligned}
Dg(u, v) &= \begin{bmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u}(a) & \frac{\partial g_2}{\partial v} \end{bmatrix} \overset{\text{usamos o teorema}}{=} [Df(x, y)]^{-1} = \\
&= \begin{bmatrix} 1 + e^{x+y} & 1 + e^{x+y} \\ 2x & 2 \end{bmatrix}^{-1} \overset{\text{usamos a Nota 1 abaixo}}{=} \\
&= \frac{1}{(2 - 2x)(1 + e^{x+y})} \begin{bmatrix} 2 & -1 - e^{x+y} \\ -2x & 1 + e^{x+y} \end{bmatrix}.
\end{aligned}$$

Se concretizarmos nos pontos $a = (0, 0)$ e $f(a) = (1, 0)$ obtemos

$$Dg(1, 0) = \frac{1}{4} \begin{bmatrix} 2 & -2 \\ 0 & 2 \end{bmatrix}.$$

Nota 1: Aqui usámos a fórmula para a matriz inversa:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Exemplo 3. Mostre que as equações

$$\begin{aligned} u &= x^3y+y^2 \\ v &= \ln(x+y) \end{aligned}$$

têm uma solução única $x = x(u,v)$, $y = y(u,v)$ numa vizinhança de $(u,v) = (6, \ln 3)$, que verifica $x(6,\ln 3) = 1$ e $y(6,\ln 3) = 2$.

Encontre $\partial x/\partial u$, $\partial x/\partial v$, $\partial y/\partial u$ e $\partial y/\partial v$ no ponto $(u,v) = (6,\ln 3)$.

Solução: Seja $f(x,y) = (u,v) = (x^3y+y^2, \ln(x+y))$. Encontrar x e y em função de u e v é mostrar que f^{-1} existe, porque então teremos $(x,y) = f^{-1}(u,v)$.

Apliquemos então o teorema. A diferenciabilidade de f é evidente, trata-se de funções polinomiais e logarítmicas. Como as derivadas

$$\frac{\partial f_1}{\partial x} = 3x^2y, \qquad \frac{\partial f_1}{\partial y} = x^3+2y,$$

$$\frac{\partial f_2}{\partial x} = \frac{1}{x+y} \quad \text{e} \quad \frac{\partial f_2}{\partial y} = \frac{1}{x+y}$$

existem e são contínuas para todo o (x,y) tal que $x+y \neq 0$, $Df(x,y)$ existe numa vizinhança de $(1,2)$ e

$$\begin{aligned} Df(1,2) &= \begin{bmatrix} 6 & 5 \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \\ J_f(1,2) &= \frac{1}{3} \neq 0. \end{aligned}$$

Temos então garantida a existência de f^{-1} numa vizinhança de $f(1,2) = (6,\ln 3)$. Além disso $Df^{-1}(u,v) = [Df(x,y)]^{-1}$ e logo

$$Df^{-1}(6,\ln 3) = \begin{bmatrix} 6 & 5 \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}^{-1} \overset{\text{usamos a Nota 1}}{=} 3\begin{bmatrix} \frac{1}{3} & -5 \\ -\frac{1}{3} & 6 \end{bmatrix} = \begin{bmatrix} 1 & -15 \\ -1 & 18 \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}, \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$

3.6.2 Teorema da função implícita

O teorema da função implícita trata o problema de saber quando é que uma relação entre varáveis permite uma delas ser resolvida em termos das outras. Este problema analisado em certas situações é de abordagem simples e intuitiva, mas nem sempre isso acontece. Vejamos a relação entre três variáveis x, y e z :

$$y + 3x - 2z + 3 = 0.$$

Evidentemente conseguimos exprimir qualquer delas em função das outras duas e ainda calcular as derivadas parciais de cada uma dessas funções. Assim

$$\begin{aligned} x(y,z) &= \frac{1}{3}(2z - y - 3); \\ y(x,z) &= -3x + 2z - 3; \\ z(x,y) &= \frac{1}{2}(3x + y + 3) \end{aligned}$$

e

$$\begin{aligned} \frac{\partial x}{\partial y} &= -1/3, & \frac{\partial x}{\partial z} &= 2/3; \\ \frac{\partial y}{\partial x} &= -3, & \frac{\partial y}{\partial z} &= 2; \\ \frac{\partial z}{\partial x} &= 3/2, & \frac{\partial z}{\partial y} &= 1/2. \end{aligned}$$

Outra relação simples é por exemplo a relação entre duas variáveis x e y dada por:

$$y - e^{3x} - 2 = 0.$$

Podemos novamente resolver cada uma em função da outra, sem grandes problemas (cuidado com os domínios!):

$$\begin{aligned} y(x) &= e^{3x} + 2, \\ x(y) &= \frac{1}{3}\ln(y - 2). \end{aligned}$$

Tente calcular as derivadas y' e x' e verá que não encontrará dificuldades.

Tentemos agora encontrar as derivadas $y'(x)$ ou $x'(y)$ para as respectivas funções dadas implicitamente por

$$3x^3y^2 + x\cos y - 3 - \cos 1 = 0 \tag{3.6.2}$$

numa certa vizinhança de $(1, 1)$.

Não, não é tarefa fácil. Mas um aluno mais persistente poderá chegar lá usando a derivada da função composta. Se considerarmos $F(x, y) = 3x^3y^2 + x\cos y - 3 - \cos 1$ e supusermos que y é função de x, temos, derivando a igualdade $F(x,y)=0$ de ambos os lados

$$0 = F'(x, y(x)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} y'$$

o que dá

$$y' = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}}$$

nos pontos em que $\dfrac{\partial F}{\partial y} \neq 0$. Como, neste caso,

$$\frac{\partial F}{\partial x} = 9x^2y^2 + \cos y \quad \text{e} \quad \frac{\partial F}{\partial y} = 6x^3y - x\sin y,$$

obtemos

$$y'(1) = -\frac{9x^2y^2 + \cos y}{6x^3y - x\sin y}_{|(x,y)=(1,1)} = \frac{9 + \cos 1}{6 - \operatorname{sen} 1}.$$

Qual é então o problema? Será que seríamos capazes de afirmar, sem dúvidas, que a expressão (3.6.2) define y como função de x? Decerto, mesmo o aluno mais persistente, não conseguiria encontrar a expressão dessa função. Vejamos outro exemplo que nos vem fazer desejar o Teorema da Função Implícita. Consideremos a relação

$$3x^2 + 4y^2 + 7 = 0. \tag{3.6.3}$$

Podemos repetir os raciocínios feitos acima para descobrir, por exemplo, a derivada da função $y(x)$, definida implicitamente por (3.6.3)? Alguma coisa sairia, mas completamente sem sentido visto que (3.6.3) não define nenhuma curva em $\mathbb{R}^2$, é o conjunto vazio em pessoa.

Poderemos dar outros exemplos em que "as contas"nos levariam a absurdos:

Exemplo 1 A relação $x^2 + 7y^2 = 0$ não determina nenhuma função implícita, o conjunto em que se verifica é o conjunto singular $\{(0,0)\}$.

Exemplo 2 A relação $\sqrt{x^2y^2} - xy = 0$ satisfaz qualquer função $y = f(x)$, cujo gráfico se situe no I ou II quadrantes.

Já vimos que nem todas as relações do tipo $F(x,y)=0$ definem implicitamente uma variável em função da outra numa vizinhança de um ponto do domínio. Para enunciar o teorema que é título desta secção necessitamos ainda da seguinte notação:

Notação: No que se segue seguiremos a seguinte notação. Se $\mathbf{x} = (x_1, x_2, ..., x_k) \in \mathbb{R}^k$ e $\mathbf{y} = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$, escrevemos $(\mathbf{x}, \mathbf{y})$ para denotar o vector

$$\mathbf{z} = (x_1, x_2, ..., x_k, y_1, y_2, ..., y_n) \in \mathbb{R}^{k+n}.$$

Supondo que a nossa função F vai de $\mathbb{R}^{k+n}$em $\mathbb{R}^n$. A relação $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ é um sistema de n equações e $k+n$ incógnitas:

$$F_i(x_1, x_2, ..., x_k, y_1, y_2, ..., y_n) = 0, \qquad 1 \leq i \leq n$$

onde as funções F_i são as n componentes da função F. Recordemos que $DF(z)$ é a matriz do tipo $n \times (k+n)$, $[\frac{\partial F_i}{\partial z_j}]$, ou seja

$$DF(z) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_k} & \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_k} & \frac{\partial F_n}{\partial y_1} & \cdots & \frac{\partial F_n}{\partial y_n} \end{bmatrix}.$$

Definimos a derivada parcial em bloco, relativa à componente $\mathbf{y}$, por

$$D_y F(z) = \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial y_1} & \cdots & \frac{\partial F_n}{\partial y_n} \end{bmatrix}.$$

Teorema da Função Implícita Seja V um conjunto aberto de $\mathbb{R}^{k+n}$ e seja $F \in C^1(V, \mathbb{R}^n)$. Suponhamos que

1. $F(a, b) = 0$ para algum $(a, b) \in V$;
2. $\det[D_y F(a, b)] \neq 0$ (equivalente a $D_y F(a, b)$ ser invertível).

Então existe uma vizinhança $U \subset \mathbb{R}^k$ de a, existe uma vizinhança $W \subset \mathbb{R}^n$ de b e existe uma única função $f : U \to W$ tal que:

a) f é de classe C^1 em U,

b) $f(a) = b$,

c) para todo o $x \in U$, $(x, f(x)) \in V$ e $F(x, f(x)) = 0$

d) $Df(a) = -(D_y F(a, b))^{-1} D_x F(a, b)$.

Nota: Repare que na alínea *d*) se trata do produto de duas matrizes $n \times n$ e $n \times k$, que dará uma matriz $n \times k$.

Decerto já reparou que o teorema trata de funções de um domínio mais vasto do que apenas as de $\mathbb{R}^2$ em $\mathbb{R}^2$ como as dos exemplos que nos serviram de motivação. Continuamos a recorrer ao que nos está mais à mão: tentar ver algo de $\mathbb{R}^8$ em $\mathbb{R}^3$ é para os nossos simples cérebros uma tarefa pouco motivadora...

A demonstração deste resultado pode ser consultada em [Webb].

Apresentamos agora alguns exemplos de aplicação do teorema. Começamos por um exemplo que seria tratável, mesmo sem o teorema da função implícita, mas que nos dá a ideia do problema. Se F for uma função linear, a sua aproximação linear será a própria função F.

Exemplo 3 Consideremos o seguinte sistema de equações:

$$\begin{aligned} x_1 + x_2 + y_1 + y_2 &= 0 \\ 2x_1 + 3x_2 - y_1 + 2y_2 &= 0 \end{aligned} \tag{3.6.4}$$

que corresponde a

$$F(\mathbf{x}, \mathbf{y}) = 0$$

com

$$F(\mathbf{x}, \mathbf{y}) = F(x_1, x_2, y_1, y_2) = (x_1 + x_2 + y_1 + y_2, 2x_1 + 3x_2 - y_1 + 2y_2).$$

Temos $n = k = 2$ e verificam-se as condições do teorema, por exemplo no ponto $(a, b) = (0, 0, 0, 0)$. Neste caso

$$DF = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & -1 & 2 \end{bmatrix}$$

e

$$D_y F = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}, \qquad D_x F = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$$

em todos os pontos de $\mathbb{R}^4$. Usando novamente a Nota2, obtemos

$$(D_y F)^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}^{-1} = 1/3 \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}.$$

Podemos garantir que existe uma função f, definida numa vizinhança U de $(0, 0) \in \mathbb{R}^2$, com valores numa vizinhança W de $(0, 0) \in \mathbb{R}^2$, que verifica as alíneas $a)$ a $d)$ do teorema. Temos para a alínea $d)$ a fórmula para a sua derivada no ponto $(0,0)$, dada explicitamente por

$$Df(0,0) = (D_yF)^{-1} \times D_xF = 1/3 \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1/3 \\ -1 & -4/3 \end{bmatrix}.$$

Se quisermos fazer as contas "ao contrário"podemos neste caso (e este caso é raro!) encontrar f, senão vejamos: podemos resolver o sistema de equações (3.6.4) em ordem a y_1 e y_2. Obtemos

$$\begin{aligned} y_1 &= \frac{1}{3}x_2 \\ y_2 &= -x_1 - \frac{4}{3}x_2 \end{aligned}$$

ou seja

$$\mathbf{y} = (y_1, y_2) = \left(\frac{1}{3}x_2, -x_1 - \frac{4}{3}x_2\right) = f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2)) = f(\mathbf{x})$$

e podemos calcular directamente

$$Df(0,0) = \begin{bmatrix} 0 & 1/3 \\ -1 & -4/3 \end{bmatrix}.$$

Exemplo 4 Sejam $n = 2,\ k = 3$ e consideremos a aplicação $F = (F_1, F_2)$ de $\mathbb{R}^5$ em $\mathbb{R}^2$ dada por

$$\begin{aligned} f_1(x_1, x_2, x_3, y_1, y_2) &= x_1\ y_2 - 4x_2 + 2e^{y_1} + 3 \\ f_2(x_1, x_2, x_3, y_1, y_2) &= 2x_1 - x_3 - 6y_1 + y_2\ \cos y_1. \end{aligned}$$

Se $a = (3,2,7)$ e $b = (0,1)$, temos $F(a,b) = 0$. Calculamos as derivadas em (a,b)

$$DF = \begin{bmatrix} 1 & -4 & 0 & 2 & 3 \\ 2 & 0 & -1 & -6 & 1 \end{bmatrix}$$

e

$$D_yF = \begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix} \qquad , D_xF = \begin{bmatrix} 1 & -4 & 0 \\ 2 & 0 & -1 \end{bmatrix}$$

$$\det[D_yF(a,b)] = 20 \neq 0.$$

Podemos então garantir a existência de uma função f definida numa vizinhança de $(3,2,7)$ e com valores numa vizinhança de $(0,1)$, tal que $f \in C^1$, $f(3,2,7) = (0,1)$, $F(x, f(x)) = 0$ e podemos calcular a sua derivada

$$\begin{aligned} Df(3,2,7) &= -\begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} 1 & -4 & 0 \\ 2 & 0 & -1 \end{bmatrix} = \\ &= -\frac{1}{20}\begin{bmatrix} 1 & -3 \\ 6 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & -4 & 0 \\ 2 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/5 & -3/20 \\ -1/2 & 6/5 & 1/10 \end{bmatrix}. \end{aligned}$$

Exercícios 3.10:

1) Mostre que o sistema de equações

$$\begin{aligned} 3x + y - z + u^2 &= 0 \\ x - y + 2z + u &= 0 \\ 2x + 2y - 3z + 2u &= 0 \end{aligned}$$

pode ser resolvido para x, y, u em termos de z, para x, z, u em termos de y, para y, z, u em termos de x, mas não para x, y, z em termos de u.

2) Defina f em $\mathbb{R}^3$ por

$$f(x, y_1, y_2) = x^2 y_1 + e^x + y_2.$$

Mostre que $f(0,1,-1) = 0$, $D_1 f(0,1,-1) \neq 0$ e que então existe uma função diferenciável g nalguma vizinhança de $(1,-1)$ de $\mathbb{R}^2$, tal que $g(1,-1) = 0$ e

$$f(g(y_1, y_2), y_1, y_2) = 0.$$

Encontre $D_1 g(1,-1)$ e $D_2 g(1,-1)$.

3.7 Optimização livre e condicionada

3.7.1 Extremos e pontos críticos

O objectivo deste capítulo é estudar os pontos em que dada função, com valores reais, tenha máximos ou mínimos. Obviamente, falamos apenas de funções escalares, visto não existir ordem em $\mathbb{R}^n$, com $n \neq 1$. Esses pontos, onde se atingem valores máximos ou mínimos locais são chamados maximizantes ou minimizantes locais e não têm de existir. Para dar uma ideia desta situção basta pensar em conjuntos que não têm máximo e mínimo, apesar de terem supremo e ínfimo. Para que uma função f, definida num certo conjunto A tenha máximo ou mínimo são necessárias algumas propriedades em f e em A. Antes de mais, definamos formalmente estes conceitos.

Definição 3.6 *Seja $f : A \subset \mathbb{R}^n \to \mathbb{R}$ uma função. Dizemos que f atinge o seu máximo no conjunto A no ponto $a \in A$ se $f(x) \leq f(a)$ para todo o $x \in A$. Ao ponto a chamamos* ***maximizante*** *e ao valor $f(a)$ chamamos* ***máximo de*** *f no conjunto A. Se $f(x) < f(a)$ para todo o $x \in A$, dizemos que f tem um* ***máximo estrito*** *em a. O conceito de* ***minimizante, mínimo e mínimo estrito*** *definem-se de forma análoga, exigindo que $f(x) \geq f(a)$ ou $f(x) > f(a)$, respectivamente, para todo o $x \in A$. Máximos e mínimos são chamados* ***extremos****.*

Tal como em $\mathbb{R}$ temos o seguinte resultado de existência.

Teorema 3.4 *Se $f : K \subset \mathbb{R}^n \to \mathbb{R}$ é contínua num conjunto limitado*[3] *e fechado K, então f atinge o seu máximo e o seu mínimo em K.*

Este resultado garante-nos a existência, mas não nos fornece nenhuma pista para o seu cálculo. É aqui que a derivada vai ter um papel essencial.

Teorema 3.5 *Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ diferenciável em D_f aberto. Seja $A \subset D_f$. Se f atinge o seu máximo no conjunto A num ponto interior $a \in A$ então $Df(a) = \nabla f(a) = \mathbf{0}$.*

Demonstração: Como $a \in Int(A)$, existe uma vizinhança-δ de a toda contida em A. Então, para cada direcção e ($\|e\| = 1$), $a + te \in A$ para $|t| < \delta$. Daqui resulta que

$$D_e f(a) = \lim_{t \to 0} \frac{f(a + te) - f(a)}{t}$$

tem de ser zero, visto que este quociente tem numerador sempre negativo ($f(a)$ é máximo) e, por isso, será negativo se $t > 0$ e positivo se $t < 0$. Assim, o limite será sempre ≥ 0 e ≤ 0, ou seja, tem que ser igual a 0. Se isto se verifica em todas as direcções, verificar-se-á também nas direcções dos eixos, ou seja, todas as derivadas parciais são nulas e o gradiente é 0.■

[3]Um cojunto A diz-se limitado sse existir $M > 0$ tal que para todo o $x \in A$, se tem $\|x\| < M$.

Este resultado facilmente se adapta ao mínimo, visto que o mínimo de f será o máximo de $-f$.

Exercício: Formule e demonstre este resultado para o mínimo.

Definição 3.7 *Chamamos **ponto crítico** ou **ponto de estacionaridade** a um ponto $a \in D_f$ tal que $\nabla f(a) = \mathbf{0}$.*

Corolário 3.1 *Seja $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ diferenciável em D_f aberto e seja $A \subset D_f$. Se f atinge um máximo (ou mínimo) no conjunto A no ponto $a \in A$, então ou a é um ponto crítico de f, ou $a \in fr(A)$.*

Demonstração: Se $a \in A$, então, das duas uma: ou $a \in Int(A)$ ou $a \in fr(A)$. Se $a \in Int(A)$, então, pelo Teorema 3.5, a é um ponto crítico.■

Encontrar pontos críticos é uma tarefa fácil, basta calcular as derivadas parciais, igualá-las a 0, e resolver o sistema de equações correspondente. O problema é que nem todos os pontos críticos são pontos onde há extremos. Por analogia com o que se passa com a função real $f(x) = x^3$, também em $\mathbb{R}^n$ podem existir pontos em que a derivada é zero, mas nem máximos nem mínimos são atingidos. Em $\mathbb{R}^2$ podemos pensar em máximos como cimos de montanhas e mínimos como vértices de vales. Mas também podem existir pontos em que se atinjam máximos numa dada direcção e mínimos noutra, são os chamados pontos de sela, por ser a sela de um cavalo a forma que melhor ilustra esta situação.

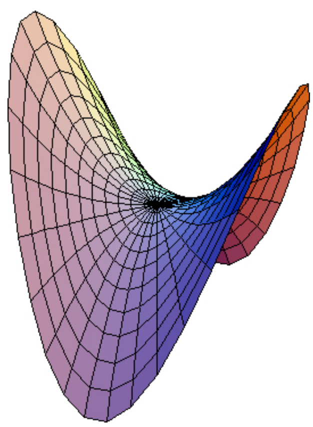


Figura 3.1: Ponto de sela.

Para espaços de dimensões superiores esta imagem já não é válida, mas usaremos o mesmo termo, **ponto sela**, para designar um ponto crítico que não seja maximizante nem minimizante.

Continuando nas imagens, podemos pensar em cadeias montanhosas, com vários cumes, onde um é o mais alto: os restantes são também chamados máximos, mas locais.

Definição 3.8 *Um ponto a é chamado* ***maximizante local*** *de f se existir um $\delta > 0$ tal que $f(a) \geq f(x)$ para todo o $x \in B(a, \delta)$. A $f(a)$ chamamos, nesse caso,* ***máximo local****. Por analogia se pode definir* ***máximo estrito local, mínimo local*** *e* ***mínimo estrito local***

Corolário 3.2 *Extremos locais ocorrem sempre em pontos críticos.*

Demonstração: Por definição estes são sempre pontos interiores.

Exercícios 3.11. Encontre os pontos críticos das seguintes funções:

1) $f(x, y) = x^3 + 3xy^2 - 15x - 12y$;

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

3) $f(x, y) = (x-1)^2 - 2y^2$;

4) $f(x, y) = (x^2 + y^2)\, e^{-\left(x^2+y^2\right)}$;

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

3.7.2 Teorema de Taylor

Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável num ponto a. Então, podemos escrever

$$f(a+h) = f(a) + Df(a)\, h + R(a, h)\,, \text{ onde } \frac{|R(a,h)|}{\|h\|_n} \to 0, \text{ quando } \|h\|_n \to 0.$$

Se a for um ponto crítico de f, $Df(a) = 0$ e

$$f(a+h) - f(a) = R(a, h)\,. \tag{3.7.5}$$

Isto significa que, para determinar se a é um ponto minimizante ou maximizante local, teremos que saber algo mais sobre o resto $R(a, h)$. Veremos que, desde que f tenha derivadas de ordens superiores, o resto da expansão em série de Taylor vai determinar $R(x, a)$.

Comecemos por enunciar o Teorema de Taylor, com resto de Lagrange, para funções reais de variável real.

Teorema 3.6 *Seja* $f :]\alpha, \beta[\rightarrow \mathbb{R}$ *uma função* p *vezes diferenciável em* $]\alpha, \beta[$ *e suponhamos que* $a \in]\alpha, \beta[$. *Então, para cada ponto* $x \in]\alpha, \beta[\backslash \{a\}$, *existe um* $\xi \in]a, x[$ *tal que*

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2 + ... + \frac{1}{(p-1)!} f^{(p-1)}(a)(x-a)^{p-1} + R_p,$$

onde

$$R_p = \frac{1}{p!} f^{(p)}(\xi)(x-a)^p.$$

Com este resto, conhecido como resto de Lagrange, o primeiro termo a ser desprezado estabelece que a derivada de ordem p é calculada num ponto ξ e não em a. Este ponto pode ser escrito na forma $\xi = (1-\theta)a + \theta x$ com $0 < \theta < 1$. Isto tem a vantagem de cobrir o caso em que $x < a$, com as devidas adaptações na forma dos intervalos $[a, x]$ considerados.

Podemos escrever a fórmula de Taylor anterior da seguinte forma equivalente

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2!} f''(a)h^2 + ... + \frac{1}{(p-1)!} f^{(p-1)}(a)h^{p-1} + \frac{1}{p!} f^{(p)}(\xi)h^p,$$

com $\xi = a + th, \ 0 < t < 1$.

Este teorema é usado para demonstrar a versão $n-$dimensional.

Da mesma forma que definimos a segunda derivada parcial de uma função, como sendo a derivada parcial de uma derivada parcial, definimos derivada parcial de ordem n como sendo a derivada parcial da derivada parcial de ordem $n-1$, definindo assim, de forma recorrente as derivadas parciais de ordem superior a 2. Igualmente se adapta a definição de C^2 a C^p, com $p > 2$, o espaço das funções diferenciáveis até à ordem p, com derivadas parciais de ordem p contínuas.

Enunciemos o Teorema de Taylor para funções escalares, $f : D_f \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

Teorema 3.7 *Seja* $D_f \subset \mathbb{R}^n$ *um conjunto aberto e seja* $f \in C^p(D_f)$. *Suponhamos que* $\mathbf{a} \in D_f$ *e que* $\mathbf{a} + t\mathbf{h} \in D_f$ *para todo o* $t \in [0,1]$. *Então existe* θ, $0 < \theta < 1$, *tal que*

$$\begin{aligned} f(a+h) &= f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + ... + \\ &\quad + \frac{1}{(p-1)!} \sum_{i_1, i_2, ..., i_{p-1}=1}^{n} \frac{\partial^{p-1} f}{\partial x_{i_1} \partial x_{i_2} ... \partial x_{i_{p-1}}}(a) h_{i_1} h_{i_2} ... h_{i_{p-1}} + R_p(h), \end{aligned}$$

onde

$$R_p(h) = \frac{1}{p!} \sum_{i_1, i_2, \ldots, i_p = 1}^{n} \frac{\partial^p f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_p}} (a + \theta h)\, h_{i_1} h_{i_2} \ldots h_{i_p}.$$

Demonstração: Faremos apenas um esboço.

Considera-se a função de variável real

$$\phi(t) = f(a + th),$$

que está definida para todo o t tal que $a+th \in D_f$. A esta função aplicamos o Teorema 3.10. e, para calcular as derivadas usa-se a regra da cadeia. Para a primeira derivada temos:

$$\phi'(t) = Df(a + th)\, h = (\nabla f(a + th), h)_n = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} (a + th)\, h_i.$$

Procede-se igualmente para as derivadas de ordem superior, obtendo o resultado pretendido com $x = 1$ e *$a = 0$,*

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2!}\phi''(0) + \ldots + \frac{1}{(p-1)!}\phi^{(p-1)}(0) + R_p$$

Nota: Se $n = 2$, o Teorema para $f \in C^2$ estabelece o seguinte

$$f(a_1 + h_1, a_2 + h_2) = f(a_1, a_2) + \frac{\partial f}{\partial x}(a_1, a_2)\, h_1 + \frac{\partial f}{\partial y}(a_1, a_2)\, h_2 +$$
$$+\frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}(a_1 + \theta h_1, a_2 + \theta h_2)\, h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a_1 + \theta h_1, a_2 + \theta h_2)\, h_1 h_2 + \frac{\partial^2 f}{\partial y^2}(a_1 + \theta h_1, a_2 + \theta h_2)\, h_2^2\right)$$

Se a for um ponto crítico, temos $Df(a) = 0$ e

$$f(a + h) = f(a) + \frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}(a + \theta h)\, h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a + \theta h)\, h_1 h_2 + \frac{\partial^2 f}{\partial y^2}(a + \theta h)\, h_2^2\right)$$

resultando assim que em (3.7.5) se tenha

$$R(a, h) = \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(a + \theta h)\, h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a + \theta h)\, h_1 h_2 + \frac{\partial^2 f}{\partial y^2}(a + \theta h)\, h_2^2\right),$$

para um certo θ, $0 < \theta < 1$. O ponto a será então um maximizante ou minimizante, consoante o sinal de $R(a, h)$, que, por sua vez, depende das segundas derivadas parciais de f calculadas no ponto $a + \theta h$. Como θ é desconhecido, deveríamos demonstrar que a função $R(a, h)$ mantém o mesmo sinal numa certa vizinhança de a (ou seja para

h perto de zero) e esse facto resulta (não trivialmente) da continuidade das derivadas parciais de 2^a ordem. Vamos então admitir que o sinal de $R(a,h)$ é o sinal de

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

donde sai trivialmente o critério estabelecido, para $n = 2$, no teorema seguinte.

Teorema 3.8 *Seja $f : D_f \subset \mathbb{R}^2 \to \mathbb{R}$, $f \in C^2(D_f)$ e seja $a \in D_f$ um ponto crítico de f. Denotando por*

$$\begin{aligned} f_{xx} &= \frac{\partial^2 f}{\partial x^2}(a), \qquad f_{yy} = \frac{\partial^2 f}{\partial y^2}(a) \qquad e \qquad f_{xy} = \frac{\partial^2 f}{\partial x \partial y}(a). \\ \Delta &= \det \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}, \end{aligned}$$

temos:

a) Se $\Delta < 0$, f admite um ponto sela em a.

b) Se $\Delta > 0$, $\begin{cases} e\ f_{xx} > 0 & f \text{ admite um mínimo local em } a \\ e\ f_{xx} < 0 & f \text{ admite um máximo local em } a \end{cases}$

c) Se $\Delta = 0$ este critério não é conclusivo.

A demonstração deste teorema baseia-se no que já foi dito sobre o sinal da função $R(a,h)$ e da seguinte igualdade

$$f_{xx}h_1^2 + 2f_{xy}h_1h_2 + f_{yy}h_2^2 = f_{xx}\left(h_1 + \frac{f_{xy}}{f_{xx}}h_2\right)^2 + \left(f_{yy} - \frac{f_{xy}^2}{f_{xx}}\right)h_2^2.$$

Como definimos atrás, a matriz $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$, é a **matriz Hessiana** de f.

Este teorema é um caso particular do caso geral, de uma função com n variáveis, que se resume no seguinte critério:

Seja $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ uma matriz $n \times n$. Chamamos menores principais aos determinantes das matrizes quadradas $k \times k$, que se obtêm a partir do canto superior esquerdo, eliminando em cada linha e em cada coluna os últimos $n - k$

elementos. Assim

$$\begin{aligned}
d_1 &= a_{11}; \\
d_2 &= \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \\
d_3 &= \det\begin{bmatrix} a_{11} & a_{12} & a_{132} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}; \\
& ... \\
d_n &= \det\begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & & & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & & & a_{nn} \end{bmatrix}.
\end{aligned}$$

Teorema 3.9 (Critério da Hessiana para extremos de funções de n variáveis:) *Seja $H(a)$ a matriz Hessiana da função f, calculada no ponto a, $H(a) = (h_{ij})_{ij} = (\frac{\partial^2 f}{\partial x_i \partial x_j}(a))_{ij}$. Então:*

a) se todos os menores principais forem positivos, $f(a)$ é um mínimo local;

b) se os menores de ordem ímpar forem negativos e os menores de ordem par forem positivos, isto é,

$$(-1)^k d_k > 0, \ \forall k \in \{1, 2, ..., n\},$$

$f(a)$ é um máximo local.

c) se se verificar uma destas ordenações até certa ordem, mas a partir daí todos os menores são nulos, nada se pode concluír;

d) em todos os restantes casos, o ponto a é um ponto sela.

Exercício: Especifique este critério para $n = 3$.

Exercícios: Calcule, se existirem, os extremos locais das seguintes funções.

1) $f(x, y) = 2x^2 + 2xy - \frac{1}{3}y^3 + 2y$.

2) $f(x, y) = e^x \cos y$.

3.7.3 Multiplicadores de Lagrange

É frequente encontrarmos problemas de optimização com condições suplementares. Por exemplo, determinar as dimensões de uma caixa cúbica, feita a partir de uma chapa de dimensão dada de forma a conter o maior volume possível, ou determinar o ponto, sobre uma curva dada, que está mais próximo da origem, ou ainda determinar a produção máxima de um campo agrícola de dimensão constante.

Todos este problemas se resumem a encontrar os extremos de uma função $f(x)$ sujeitos à condição $g(x) = k$ ou mais simplesmente $g(x) = 0$.

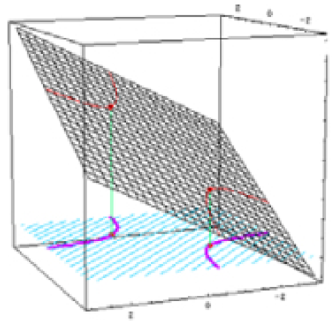


Figura 3.2: $f(x, y) = x + y$ e $g(x, y) = xy - 1$

Exemplo: Optimizar a funcão

$$f(x, y) = x^3 - xy + y^2 + 3$$

sujeita à condição $x^2 + 2y^2 = 1$. Assim, a condição $g(x, y) = 0$ é dada pela equação

$$g(x, y) = x^2 + 2y^2 - 1 = 0$$

Este problema está representado na figura seguinte

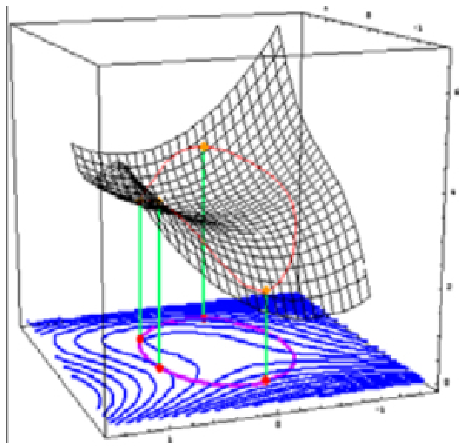


Figura 3.3: $f(x,y) = x^3 - xy + y^2 + 3$ e $g(x,y) = x^2 + 2y^2 - 1 = 0$

Há várias formas de resolver este tipo de problemas e, geralmente, existe uma abordagem diferente para cada caso. Nesta secção tentaremos dar um método geral, que explicaremos geometricamente, partindo do princípio que o nosso espaço é o $\mathbb{R}^3$.

Seja então $f(x)$ a função que queremos maximizar (ou minimizar) e $g(x) = 0$ a condição suplementar.

Tipicamente $g(x) = 0$ define uma superfície , $S = \{x : g(x) = 0\}$, chamada habitualmente conjunto de nível da função $g(x)$ e o nosso objectivo é maximizar $f(x)$ com $x \in S$.

Se $a \in S$ for um maximizante local de f restringida a S, não podemos afirmar que $f(a)$ é um máximo em todas as direcções possíveis, mas apenas nas direcções que correspondem a pontos de S. Por outras palavras, as derivadas direccionais de f devem ser zero para todas as direcções que sejam tangentes a S em a. Se e for um vector tangente a S em a, deveríamos ter $(\nabla f(a), e) = 0$, ou seja, $\nabla f(a)$ deverá ser paralelo a $\nabla g(a)$, por outras palavras, deverá existir um escalar λ tal que $\nabla f(a) = \lambda \nabla g(a)$. Considerações semelhantes se aplicam ao mínimo.

Comecemos por formular dois lemas para funções definidas em $\mathbb{R}^2$.

Lema 3.1 *Seja $a \in \mathbb{R}^2$ um ponto onde f, condicionada a $g(x,y) = 0$, admite um extremo local. Suponhamos que se verificam as duas seguintes condições:*

a) $\|\nabla g(a)\| > 0$;

b) As funções f e g pertencem a C^1 numa vizinhança de a.

Então existe um real λ (chamado multiplicador de Lagrange), o qual, juntamente *com o ponto a, verificam o seguinte sistema de equações.*

$$\begin{aligned} \nabla f(x,y) &= \lambda \nabla g(x,y); \\ g(x,y) &= 0. \end{aligned}$$

Lema 3.2 *Suponhamos que os números x_0, y_0, λ_0 são uma solução do sistema de equações:*

$$\begin{aligned} \nabla f(x,y) &= \lambda \nabla g(x,y); \\ g(x,y) &= 0. \end{aligned}$$

Se a função de Lagrange

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$$

admite um extremo no ponto $a = (x_0, y_0)$ para $\lambda = \lambda_0$, então a função $z = f(x,y)$ tem no ponto a um extremo condicionado a $g(x,y) = 0$.

Estes dois lemas generalizam-se a mais do que duas variáveis e a mais do que uma condição. É esse o conteúdo do teorema dos multiplicadores de Lagrange na sua forma geral.

Teorema 3.10 (dos multiplicadores de Lagrange) *Sejam f, g_1, g_2, ..., g_m funções escalares definidas em $\mathbb{R}^n$ que são continuamente diferenciáveis num aberto D de $\mathbb{R}^n$. Suponhamos que a Jacobiana de $g = (g_1, g_2, ..., g_m)$, do tipo $m \times n$ (com $m < n$),*

$$Dg = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

tem característica m.

Suponhamos que o sistema de $n + m$ equações, relativo à função Lagrangeana

$$L(x,\lambda) = f(x) + \lambda_1 g_1(x) + ... + \lambda_m g_m(x),$$

$$\nabla L(x,\lambda) = \mathbf{0} \Longleftrightarrow \begin{cases} \nabla f(x) + \lambda_1 \nabla g_1(x) + ... + \lambda_m \nabla g_m(x) = \mathbf{0} \\ g(x) = \mathbf{0} \end{cases}$$

tem solução $\left(\widehat{x}, \widehat{\lambda}\right) = (\widehat{x}_1, \widehat{x}_2, ..., \widehat{x}_n, \widehat{\lambda}_1, \widehat{\lambda}_2, ..., \widehat{\lambda}_m)$.

Se a função Lagrangeana L tiver para $\widehat{\lambda} = \left(\widehat{\lambda}_1, \widehat{\lambda}_2, ..., \widehat{\lambda}_m\right)$, no ponto crítico $\widehat{x} = (\widehat{x}_1, \widehat{x}_2, ..., \widehat{x}_n)$, um extremo local, então a função f tem em $\widehat{x}$ um extremo do mesmo tipo, sujeito às condições $g(x) = 0$.

Exemplo. Calcule os extremos de $f(x,y) = (x-1)^2 + y^2$, sujeito à condição $y^2 - 4x = 0$.

Resolução: Formamos a função Lagrangeana

$$L(x,y,\lambda) = x^2 - 2x + y^2 + 1 + \lambda\left(y^2 - 4x\right)$$

e o respectivo sistema de equações

$$\begin{array}{l} \frac{\partial L}{\partial x} = 0 \Longleftrightarrow 2x - 2 - 4\lambda = 0 \\ \frac{\partial L}{\partial y} = 0 \Longleftrightarrow 2y + 2\lambda y = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \Longleftrightarrow y^2 = 4x \end{array}$$

Da segunda equação tiramos que ou $\lambda = -1$ ou $y = 0$. Se λ fosse -1 teríamos, da primeira equação $x = -1$ e a terceira equação seria impossível. Asssim, $y = 0$, $x = 0$ e $\lambda = -\frac{1}{2}$. Pelo lema 3.1 sabemos que extremos condicionados só podem existir em $(0,0)$. Para sabermos se de facto este ponto é um maximizante ou minimizante recorremos à matriz Hessiana

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y} \\ \frac{\partial^2 L}{\partial x \partial y} & \frac{\partial^2 L}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2+2\lambda \end{bmatrix}$$

Tomamos $\lambda = -\frac{1}{2}$ e

$$\Delta = \det H(0,0) = \det \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = 2 > 0.$$

$$\text{e sendo } \frac{\partial^2 L}{\partial x^2}(0,0) = 2 > 0,$$

obtemos então que $(0,0)$ é um minimizante e $f(0,0) = 1$ um mínimo de f sujeito à condição $y^2 - 4x = 0$.

Nota: Por vezes, depois de encontrarmos os pontos críticos da Lagrangeana, deparamo-nos com $\Delta = 0$, o que nos impede de usar o critério da Hessiana. Nesses casos podemos proceder à análise do problema sem recurso a "receitas"e nem sempre é fácil concluir.

No caso da condição $g(x, y) = 0$, definir um conjunto limitado e fechado, podemos usar o teorema 3.8., que estabelece a existência de máximo e mínimo nesse conjunto. Estes extremos só podem ser atingidos nos pontos críticos, que se obtêm pelo método dos multiplicadores de Lagrange, por isso, bastará tomar o maior como máximo e o menor como mínimo.

Bibliografia

[Apos] T.M.Apostol, Cálculo, Vol.2, Editora reverté,lda.

[Dem] B. Demidovich, Exercícios de Análise Matemática, Mac Graw Hill.

[FerAm] M.A.Ferreira e I. Amaral, Matemática, Cálculo diferencial em Rn, Edições Sílabo.

[LM] L. Magalhães, Álgebra Linear como introdução à matemática aplicada, 3^a edição, Texto Editora.

[Piskounov] Cálculo Diferencial e Integral, Vol.II, Lopes da Silva Editora, 2002.

[Rudin] Walter Rudin, *Principles of Mathematical Analysis,* third edition, International Series in Pure ans Applied Mathematics, McGraw-Hill.

[SkTi] Základy aplikované matematiky I, Josef Skrásek, Zdenek Tichý, SNTL, Praha 1989.

[Webb] J.R.L. Webb, *Functions of several real variables*, Ellis Horwood series in Mathematics and its applications, King's College, University of London.

[Inline1] http://valle.fciencias.unam.mx/intermat/ArticuloLag/articuloPDFb.pdf