

Exercícios de Cálculo Diferencial e Integral de Funções
Definidas em \mathbb{R}^n

Diogo Aguiar Gomes, João Palhoto Matos e João Paulo Santos

24 de Janeiro de 2000

Conteúdo

1	Introdução	5
1.1	Explicação	5
1.2	Futura introdução	5
2	Complementos de Cálculo Diferencial	7
2.1	Preliminares	7
2.1.1	Exercícios suplementares	10
2.1.2	Sugestões para os exercícios	11
2.2	Cálculo diferencial elementar	12
2.2.1	Exercícios suplementares	17
2.2.2	Sugestões para os exercícios	18
2.3	Derivadas parciais de ordem superior à primeira	19
2.3.1	Exercícios suplementares	22
2.3.2	Sugestões para os exercícios	22
2.4	Polinómio de Taylor	23
2.4.1	Exercícios suplementares	25
2.4.2	Sugestões para os exercícios	26
3	Extremos	27
3.1	Extremos	28
3.1.1	Exercícios suplementares	33
3.1.2	Sugestões para os exercícios	34
3.2	Testes de Segunda Ordem	34
3.2.1	Exercícios suplementares	43
3.2.2	Sugestões para os exercícios	43
4	Teoremas da Função Inversa e da Função Implícita	47
4.1	Invertibilidade de funções	47
4.1.1	Exercícios Suplementares	50
4.1.2	Sugestões para os exercícios	51
4.2	Teorema do valor médio para funções vectoriais	52
4.3	Teorema da Função Inversa	53
4.3.1	Exercícios Suplementares	60
4.3.2	Sugestões para os exercícios	61
4.4	Teorema da Função Implícita	62
4.4.1	Exercícios suplementares	67
4.4.2	Sugestões para os exercícios	67
	Bibliografia	69

Capítulo 1

Introdução

1.1 Explicação

Está a ler uma versão parcial e preliminar de um texto em elaboração. Os autores agradecem quaisquer notificações de erros, sugestões, . . . , para ecdi@math.ist.utl.pt. Estima-se que o texto final terá uma extensão cerca de três a quatro vezes maior e incluirá capítulos que nesta versão foram excluídos.

A secção seguinte desta introdução tem carácter preliminar e tem como pressuposto a existência do material que aqui ainda não foi incluído.

Partes deste texto foram distribuídas separadamente por cada um dos autores no passado. Tendo descoberto que os diversos textos tinham carácter algo complementar decidimos reuni-los. A presente versão idealmente não mostra de uma maneira óbvia as adaptações e correcções que foram necessárias para chegar ao formato actual.

Novas versões deste texto irão aparecendo sempre que os autores considerarem oportuno em <http://www.math.ist.utl.pt/~jmatos/AMIII/temp.pdf>. Para evitar a proliferação de textos obsoletos a maioria das páginas apresenta a data de revisão corrente em pé de página.

1.2 Futura introdução

Este texto nasce da nossa experiência a leccionar a disciplina de Análise Matemática III no Instituto Superior Técnico. Por um lado reúne um número considerável de enunciados de problemas de exame e por outro serve de propaganda à nossa maneira de ver os assuntos aqui tratados. Análise Matemática III é uma disciplina do primeiro semestre do segundo ano de todos os currículos de licenciatura leccionados no Instituto Superior Técnico (IST) excepto Arquitectura.

Se se perguntar a um aluno de um dos dois primeiros anos do IST que tipo de “folhas” mais deseja que lhe sejam disponibilizadas pelos seus professores temos como resposta mais que provável: “folhas de exercícios resolvidos de Análise Matemática”. No entanto tal resposta costuma suscitar como reacção da parte dos docentes essencialmente *preocupação*. De facto a resolução de exercícios de Análise Matemática não é geralmente única e o processo de aprendizagem está mais ligado à tentativa de resolução dos mesmos quando se possui um conjunto de conhecimentos mínimo do que à absorção acéfala de um número finito de receitas.

O que se segue é uma tentativa de compromisso entre a procura e a oferta neste mercado *sui generis*. São incluídos exercícios de exame dos últimos anos com modificações do enunciado quando tal foi julgado conveniente e muitos outros com um carácter mais ou menos trivial, ou de complemento de resultados citados, ou de comentário de uma resolução de um exercício, sugestão de extensões, etc. Por vezes um exercício embora incluído numa secção inclui uma questão que só é tratada numa secção posterior. Tais exercícios estão assinalados com um asterisco *. Foram incluídos esboços de resolução e sugestões em número considerável.

O leitor deverá ter em consideração que o programa de Análise Matemática III tem variado ao longo do tempo. É consensual no Departamento de Matemática do IST e na escola em geral que a introdução à análise em \mathbb{R}^n e o cálculo diferencial em \mathbb{R}^n deverão ser tratados em grande parte no primeiro ano do curso. Daí a existência de secções correspondentes a revisão de material coberto no primeiro ano do curso.

Outro facto a ter em conta é a diferença de programa para os cursos de Matemática Aplicada e Computação e Engenharia Física Tecnológica. Nestes cursos são introduzidos o formalismo das formas diferenciais e a respectiva versão do teorema fundamental do cálculo em vez da formulação clássica do teorema de Stokes. Aconselha-se os alunos destes dois cursos a comparar os enunciados de exercícios deste tema com as formulações clássicas dos mesmos. Tais comparações estão indicadas em nota de pé de página.

A notação utilizada é *clássica* tanto quanto possível, embora obviamente não universal, e nem sempre será isenta de incoerências. Por exemplo: usaremos a notação de Leibniz para derivadas parciais mas de acordo com a notação geral para operadores, isto é, $\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial y} \right)$; usaremos \iint , \iiint sempre que tal for considerado sugestivo.

Citaremos os resultados essenciais de cada tema mas não necessariamente com a sua formulação mais geral remetida por vezes para observações marginais ou problemas. O enunciado de tais resultados por vezes é seguido de uma “demonstração” que mais não faz que relembrar sinteticamente a dependência em relação a outros resultados e os métodos utilizados.

Faz-se notar que não seguimos a ordenação de material geralmente adoptada durante a exposição dos cursos no IST devido a razões como a conveniência em apresentar problemas sobre a introdução do conceito de variedade como complemento do estudo do teorema da função implícita.

Um último aviso: este texto não pretende substituir os excelentes livros de texto disponíveis sobre os assuntos aqui abordados. Diria mesmo que é provavelmente incompreensível se um ou mais desses livros não for consultado. Os textos adoptados no IST são [6, 3, 5].

Lisboa, Outubro de 1999

DG, JPM, JPS

Capítulo 2

Complementos de Cálculo Diferencial

O conceito de *função diferenciável* é uma das noções chave da análise. Por exemplo, se $f : \mathbb{R} \rightarrow \mathbb{R}$ for diferenciável em x_0 , o cálculo de $f'(x_0)$ permite aproximar f pela fórmula de Taylor perto de x_0 , i.e.,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0),$$

onde $\lim_{x \rightarrow x_0} \frac{o(x-x_0)}{x-x_0} = 0$. Esta fórmula tem a seguinte interpretação geométrica: $f'(x_0)$ é o declive da recta tangente a f em x_0 e $y = f(x_0) + f'(x_0)(x - x_0)$ é a equação dessa recta.

Outras aplicações do conceito de derivada familiares a um estudante que conheça Análise Matemática ao nível de um primeiro ano de licenciatura são, por exemplo, a determinação de pontos de extremo: se $f : \mathbb{R} \rightarrow \mathbb{R}$ for diferenciável, os seus máximos ou mínimos são zeros de f' ¹. Outra aplicação que deve ser familiar é a mudança de coordenadas na integração através de:

$$\int_a^b g(x)dx = \int_{f^{-1}(a)}^{f^{-1}(b)} g(f(y))f'(y)dy.$$

Esta presença ubíqua da diferenciação no estudo de funções reais de variável real faz com que seja natural, quando se estudam funções de várias variáveis, generalizar a noção de derivada. Para funções de \mathbb{R}^n em \mathbb{R} , a interpretação geométrica da derivada será o “declive” do “plano” tangente ao gráfico da função, mais precisamente $y = f(x_0) + Df(x_0)(x - x_0)$ é a equação desse “plano” tangente².

Neste capítulo resumiremos alguns resultados de cálculo diferencial, para funções reais de mais do que uma variável real. Em particular trataremos questões importantes sobre a continuidade e diferenciabilidade de funções de \mathbb{R}^n em \mathbb{R}^m . Para além disso estudaremos a fórmula de Taylor.

2.1 Preliminares

Esta secção relembra alguns dos conceitos e resultados sobre funções de \mathbb{R}^n em \mathbb{R}^m que se supõem conhecidos nas secções seguintes. Aconselha-se o leitor a consultar [1] para relemburar, com detalhe, os resultados, supostos já conhecidos, que a seguir se enumeram de uma forma necessariamente breve.

Tanto a definição de continuidade como a de diferenciabilidade dependem do conceito de distância entre dois pontos, definida por sua vez à custa da noção de *norma*:

¹Note, no entanto, que o facto de a derivada se anular num ponto, não implica que este seja um máximo ou mínimo; pode ser ponto de sela! Veja o capítulo 3.

²Designações técnicas para um tal conjunto são de um subespaço afim de dimensão n de \mathbb{R}^{n+1} ou *hiperplano*

Definição 2.1.1 Seja $\eta : \mathbb{R}^n \rightarrow \mathbb{R}$. Diz-se que η é uma norma se verificar as seguintes propriedades:

- i) $\eta(\mathbf{x}) > 0$ se $\mathbf{x} \neq 0$ e $\eta(0) = 0$;
- ii) $\eta(\lambda \mathbf{x}) = |\lambda| \eta(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$;
- iii) $\eta(\mathbf{x} + \mathbf{y}) \leq \eta(\mathbf{x}) + \eta(\mathbf{y})$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Para designarmos uma norma genérica utilizaremos a notação $\|\mathbf{x}\| = \eta(\mathbf{x})$. Em \mathbb{R}^n é usual considerar a norma euclideana, definida por

$$\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Porém, em certas situações, pode ser útil trabalhar com normas diferentes.

Exercício 2.1.1 Prove que as seguintes funções são normas em \mathbb{R}^2 :

1. $\eta(x, y) = |x| + |y|$
2. $\eta(x, y) = \max\{|x|, |y|\}$
3. $\eta(x, y) = 2\sqrt{x^2 + y^2}$
4. $\eta(x, y, z) = |x| + \sqrt{y^2 + z^2}$.

Exercício 2.1.2 Mostre que $\eta(x, y) = |x + y|$ não é uma norma mas satisfaz ii e iii em 2.1.1.

Definição 2.1.2 Em \mathbb{R}^n , a bola (aberta) centrada em \mathbf{x} e de raio r , relativa à norma $\|\cdot\|$, é o conjunto $B(\mathbf{x}, r)$ (ou $B_r(\mathbf{x})$) definido por

$$B(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}.$$

Se a norma em questão for a norma euclideana as bolas serão “redondas”, caso contrário poderão ter formatos mais ou menos inesperados, como se pode ver no exercício seguinte.

Exercício 2.1.3 Esboce as bolas $B_1(0)$ em \mathbb{R}^2 para as seguintes normas:

1. $\|(x, y)\| = \sqrt{x^2 + y^2}$
2. $\|(x, y)\| = |x| + |y|$
3. $\|(x, y)\| = \max\{|x|, |y|\}$

Exercício 2.1.4 Mostre que uma bola será sempre um conjunto convexo, isto é, dados dois quaisquer dos seus pontos, o segmento de recta que os une está contido na bola.

Daqui para a frente vamos sempre supor que a norma em \mathbb{R}^n é a norma euclideana, a não ser que seja dito algo em contrário. Além disso a notação não distinguirá as normas euclidianas em diferentes espaços \mathbb{R}^n para $n \geq 2$.

Definição 2.1.3 Diz-se que um conjunto $A \subset \mathbb{R}^n$ é aberto se verificar a seguinte propriedade:

$$\forall \mathbf{x} \in A, \exists r > 0 : B(\mathbf{x}, r) \subset A.$$

Exemplo 2.1.1 O conjunto $]0, 1[\subset \mathbb{R}$ é aberto. Com efeito, para qualquer número real $0 < x < 1$ temos $x > 1/2$ ou $x \leq 1/2$. No primeiro caso $B(x, x/2) \subset]0, 1[$, no segundo $B(x, (1-x)/2) \subset]0, 1[$.

Exercício 2.1.5 Mostre que as bolas abertas são conjuntos abertos.

Temos reunidos todos os ingredientes necessários à definição de *função contínua*:

Definição 2.1.4 Diz-se que uma função $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é contínua num ponto $\mathbf{x} \in A$ se:

$$\forall \epsilon > 0 \exists \delta > 0 \text{ tal que } \|\mathbf{x} - \mathbf{y}\| < \delta, \mathbf{y} \in A \Rightarrow \|f(\mathbf{x}) - f(\mathbf{y})\| < \epsilon.$$

Diz-se que f é contínua num subconjunto do seu domínio se for contínua em todos os pontos desse conjunto.

Exemplo 2.1.2 Suponhamos $f(x, y) = x + y$. Provemos que f é contínua. Seja $\epsilon > 0$ arbitrário. Reparemos que, para todo o (x_1, y_1) e (x_2, y_2) , se tem

$$|x_1 + y_1 - x_2 - y_2| \leq |x_1 - x_2| + |y_1 - y_2|,$$

sendo que $|x_1 - x_2| \leq \|(x_1, y_1) - (x_2, y_2)\|$ e $|y_1 - y_2| \leq \|(x_1, y_1) - (x_2, y_2)\|$. Portanto, fixando $\epsilon > 0$, e escolhendo $\delta < \frac{\epsilon}{2}$ teremos:

$$|x_1 + y_1 - x_2 - y_2| \leq 2\delta < \epsilon,$$

se $\|(x_1, y_1) - (x_2, y_2)\| < \delta$. Logo f é contínua.

Exercício 2.1.6 Mostre que a função definida por

$$f(x, y) = \begin{cases} 1, & \text{se } x + y > 0, \\ 0, & \text{se } x + y \leq 0 \end{cases}$$

não é contínua.

Muitas vezes, para mostrar continuidade (ou a falta dela), utiliza-se a caracterização de continuidade através de sucessões:

Teorema 2.1.1 (Continuidade à Heine)

Seja $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. f é contínua em $\mathbf{x}_0 \in A$ se e somente se para toda a sucessão $(\mathbf{x}_k)_{k \in \mathbb{N}} \subset A$ que converge para \mathbf{x}_0 (isto é, $\lim_{k \rightarrow +\infty} \|\mathbf{x}_k - \mathbf{x}_0\| = 0$) a sucessão $(f(\mathbf{x}_k))_{k \in \mathbb{N}}$ converge para $f(\mathbf{x}_0)$.

Exemplo 2.1.3 Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$, f e g contínuas. Provemos que $g \circ f$ é contínua. Seja $\mathbf{x}_0 \in \mathbb{R}^n$ e $(\mathbf{x}_k) \subset \mathbb{R}^n$ uma sucessão convergente para \mathbf{x}_0 . Definindo $\mathbf{y}_k = f(\mathbf{x}_k)$ obtemos uma sucessão $(\mathbf{y}_k) \subset \mathbb{R}^m$ que converge para $\mathbf{y}_0 = f(\mathbf{x}_0)$, uma vez que f é contínua. A sucessão $(\mathbf{z}_k) \subset \mathbb{R}^p$, definida por $\mathbf{z}_k = g(\mathbf{y}_k)$, converge para $\mathbf{z}_0 = g(\mathbf{y}_0)$, uma vez que g é contínua. Resta observar que $\mathbf{z}_k = g \circ f(\mathbf{x}_k) \rightarrow \mathbf{z}_0 = g \circ f(\mathbf{x}_0)$, pelo que $g \circ f$ é contínua.

Exercício 2.1.7 Refaça o exemplo anterior usando a definição 2.1.4.

Exercício 2.1.8 Prove o teorema 2.1.1.

Exercício 2.1.9 Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Prove que f é contínua se e somente se para todo o aberto $A \subset \mathbb{R}^m$ se tem $f^{-1}(A) \subset \mathbb{R}^n$ aberto, onde o conjunto $f^{-1}(A)$ é definido como sendo:

$$f^{-1}(A) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \in A\}.$$

Generalize este resultado para funções definidas num subconjunto arbitrário de \mathbb{R}^n .

Definição 2.1.5 Diz-se que um conjunto $F \subset \mathbb{R}^n$ é fechado se o seu complementar F^c for aberto.

Teorema 2.1.2 (Caracterização dos fechados via sucessões)

$F \subset \mathbb{R}^n$ é fechado se e só se dada uma qualquer sucessão convergente de termos em F esta converge para um elemento de F .

Exercício 2.1.10 Dê dois exemplos distintos de subconjuntos de \mathbb{R}^n que sejam, cada um deles, simultaneamente aberto e fechado (isto só se verifica para dois conjuntos muito especiais!).

Definição 2.1.6 A união de todos os abertos contidos num conjunto A será designada por interior de A e abrevia-se $\text{int } A$. A intersecção de todos os fechados contendo A chamar-se-á fecho de A e abrevia-se \bar{A} . A fronteira de A , ∂A , é definida por $\partial A = \bar{A} \setminus \text{int } A$.

Definição 2.1.7 Diz-se que um conjunto $K \subset \mathbb{R}^n$ é compacto se dada uma qualquer sucessão de termos em K esta possui uma subsucessão convergente para um elemento de K .

Teorema 2.1.3 (Caracterização dos compactos de \mathbb{R}^n)

$K \subset \mathbb{R}^n$ é compacto se e só se K é limitado e fechado.

Exercício 2.1.11 O conjunto vazio é compacto? E o conjunto dos números racionais de valor absoluto menor que 1?

Exercício 2.1.12 Dê um exemplo de uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

1. $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq 1\}$ seja um conjunto compacto.
2. $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < 1\}$ seja um conjunto compacto não vazio. Observação: se f for contínua então este conjunto é necessariamente aberto (porquê?) portanto se escolher f contínua o conjunto será necessariamente vazio (porquê?).
3. Seja K um conjunto compacto. Construa uma função f tal que $K = \{x : f(x) = 1\}$. Escolhendo f não contínua o problema é trivial. No entanto pode tornar o problema bem mais interessante tentando construir f contínua!

2.1.1 Exercícios suplementares

Exercício 2.1.13 Diz-se que duas normas em \mathbb{R}^n , $\|\cdot\|_\alpha$ e $\|\cdot\|_\beta$, são equivalentes se existirem constantes positivas, a e b tais que

$$a\|\mathbf{x}\|_\alpha \leq \|\mathbf{x}\|_\beta \leq b\|\mathbf{x}\|_\alpha$$

para todo o $\mathbf{x} \in \mathbb{R}^n$. Prove que as seguintes normas são todas equivalentes entre si:

1. $\|(x_1, \dots, x_n)\|_1 = |x_1| + \dots + |x_n|$
2. $\|(x_1, \dots, x_n)\|_2 = \sqrt{|x_1|^2 + \dots + |x_n|^2}$
3. $\|(x_1, \dots, x_n)\|_\infty = \text{máx}\{|x_1|, \dots, |x_n|\}$

Exercício 2.1.14 Prove que as seguintes funções são contínuas:

1. $f(x) = 1$ se $-\infty < x \leq 1$ e $f(x) = x$ se $x \geq 1$;
2. qualquer polinómio em n variáveis.

Exercício 2.1.15 Prove que

$$f(x) = \begin{cases} 0, & \text{se } x < 0, \\ 1, & \text{se } x \geq 0, \end{cases}$$

não é contínua.

Exercício 2.1.16 Diz-se que uma função $f : J \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é semicontínua inferior se para toda a sucessão $\mathbf{x}_k \rightarrow \mathbf{x} \in J$ se tem $\liminf_{j \rightarrow +\infty} f(\mathbf{x}_k) \geq f(\mathbf{x})$ (recorde que o \liminf de uma sucessão $(y_k)_{k \in \mathbb{N}}$ é definido como sendo $\liminf_{k \rightarrow +\infty} y_k = \lim_{n \rightarrow +\infty} \inf_{k > n} \{y_k\}$).

1. Mostre que o \liminf existe sempre (eventualmente pode ser igual a $-\infty$, quando?).
2. Mostre que qualquer função contínua é semicontínua inferior.
3. Dê um exemplo de uma função semicontínua inferior que não seja contínua.
4. Mostre que qualquer função semicontínua inferior f definida num compacto K é limitada inferiormente, isto é $\exists C \in \mathbb{R}$ tal que $f(x) \geq C$ sempre que $x \in K$.
5. Mostre que uma função semicontínua inferior definida num compacto tem sempre mínimo.
6. Utilizando as ideias das alíneas anteriores mostre que qualquer função contínua definida num compacto tem máximo e mínimo.

Exercício 2.1.17 As definições de aberto e função contínua dependem aparentemente de usarmos a norma euclidiana. Uma dúvida legítima é saber se tivéssemos usado outra norma chegaríamos às mesmas conclusões relativamente a que conjuntos são abertos e que funções são contínuas. Mostre que:

1. Todas as normas em \mathbb{R}^n são contínuas.
2. Qualquer norma em \mathbb{R}^n tem um mínimo positivo na fronteira da bola $B(0,1)$.
3. Todas as normas em \mathbb{R}^n são equivalentes.
4. Conclua que as noções de aberto e função contínua são independentes da norma utilizada.

2.1.2 Sugestões para os exercícios

2.1.13 Observe que $\forall \mathbf{x} \in \mathbb{R}^n$

1. $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n\|\mathbf{x}\|_\infty$;
2. $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty$.

Usando 1 e 2 deduza as restantes desigualdades. ◀

2.1.14 Utilize a definição 2.1.4 e o teorema 2.1.1. ▶

2.1.15 Note que $f(-\frac{1}{n}) \rightarrow 0 \neq f(0)$. ▶

2.1.16

1. Note que a sucessão $z_n = \inf_{k>n} \{y_k\}$ é monótona crescente.
2. Se f é contínua e $\mathbf{x}_k \rightarrow \mathbf{x}$ então $f(\mathbf{x}_k) \rightarrow f(\mathbf{x})$.
3. Por exemplo

$$f(x) = \begin{cases} 0 & \text{se } x \leq 0, \\ 1 & \text{se } x > 0. \end{cases}$$

4. Se f não fosse limitada inferiormente existiria uma sucessão $\mathbf{x}_k \in K$ tal que $f(\mathbf{x}_k) \rightarrow -\infty$. Como K é compacto poder-se-ia extrair uma subsucessão convergente $\mathbf{x}_{k_j} \rightarrow \mathbf{x} \in K$. Consequentemente ter-se-ia $-\infty = \lim f(\mathbf{x}_{k_j}) = \liminf f(\mathbf{x}_{k_j}) \geq f(\mathbf{x}) > -\infty$ o que é absurdo.
5. Seja $f : K \rightarrow \mathbb{R}$, onde $K \subset \mathbb{R}^n$ é compacto, semicontínua inferior. Note que, pela alínea anterior, f é minorada. Defina-se $m = \inf_{y \in K} f(y)$. Então existe uma sucessão $\mathbf{x}_k \in K$ tal que $f(\mathbf{x}_k) \rightarrow m$. Como K é compacto, existe uma subsucessão \mathbf{x}_{k_j} que converge para algum $x \in K$. Por semicontinuidade inferior tem-se

$$m = \lim_{j \rightarrow +\infty} f(\mathbf{x}_{k_j}) = \liminf_{j \rightarrow +\infty} f(\mathbf{x}_{k_j}) \geq f(\mathbf{x})$$

mas por outro lado $f(\mathbf{x}) \geq \inf_{y \in K} f(y) = m$ portanto $f(\mathbf{x}) = m$.

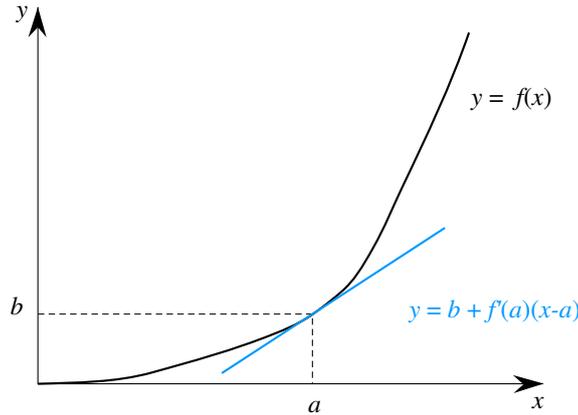


Figura 2.1: A interpretação geométrica de derivada para funções reais de variável real.

6. Se f é contínua então f e $-f$ são semicontínuas inferiores.

2.2 Cálculo diferencial elementar

Vamos começar por definir *função diferenciável*.

Definição 2.2.1 Seja $U \subset \mathbb{R}^n$ um aberto. Diz-se que uma função $f : U \rightarrow \mathbb{R}^m$ é diferenciável no ponto $\mathbf{x}_0 \in U$ se existir uma aplicação linear A de \mathbb{R}^n em \mathbb{R}^m , para a qual se tem

$$\lim_{\mathbf{h} \rightarrow 0, \mathbf{h} \in \mathbb{R}^n} \frac{\|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - A\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

Será a aplicação linear A na definição anterior que chamaremos *derivada*³ de f no ponto \mathbf{x}_0 . No entanto poderia existir mais do que uma aplicação linear nestas condições. . .

Problema 2.2.1 Mostre que a aplicação linear A da definição 2.2.1 se existir é única.

Definição 2.2.2 A aplicação linear A da definição 2.2.1 designa-se por *derivada* de f em \mathbf{x}_0 escrevendo-se $Df(\mathbf{x}_0)$.

Esta definição de derivada coincide com a definição usual de derivada para funções reais de variável real. Para este caso, a aplicação linear A referida na definição anterior é simplesmente multiplicação por um escalar.

Exercício 2.2.1 Suponha $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é diferenciável num ponto $\mathbf{x}_0 \in \text{int } U$. Prove que

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + Df(\mathbf{x}_0)(\mathbf{h}) + o(\mathbf{h}),$$

onde $\lim_{\mathbf{h} \rightarrow 0, \mathbf{h} \in \mathbb{R}^n} \frac{o(\mathbf{h})}{\|\mathbf{h}\|} = 0$.

Definição 2.2.3 Diz-se que uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Se U for aberto dizemos que f é diferenciável em U se o for em todos os pontos do domínio U . Se U não for aberto dizemos que f é diferenciável em U se existir um prolongamento \bar{f} de f a um aberto V contendo U tal que \bar{f} seja diferenciável em V .

³Tal aplicação será muitas vezes identificada com a matriz real $m \times n$ que a representa ou com um vector se n ou m for igual a 1. Se $n = 1$ é comum usar $f'(x_0)$ em vez de $Df(x_0)$.

Exemplo 2.2.1 Seja f definida em \mathbb{R} por $f(x) = x^3$. Mostremos que ela é diferenciável em qualquer ponto de $x \in \mathbb{R}$ e que a sua derivada é $3x^2$.

Com efeito temos

$$\lim_{h \rightarrow 0} \frac{|(x+h)^3 - x^3 - 3x^2h|}{|h|} = \lim_{h \rightarrow 0} \frac{|3xh^2 + h^3|}{|h|} = 0.$$

A verificação da diferenciabilidade usando directamente a definição pode ser, mesmo em casos simples, penosa. Isso não acontece, no entanto, no caso ilustrado no próximo exercício.

Exercício 2.2.2 Mostre que uma transformação linear $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$, dada por $f(x) = Mx$, onde M é uma matriz $n \times m$, é diferenciável e que $Df = M$.

As funções diferenciáveis formam um subconjunto estrito das funções contínuas. Com efeito:

Exercício 2.2.3 Mostre que qualquer função diferenciável é contínua.

Consideremos uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ e fixemos um vector $v \in \mathbb{R}^n$. Dado um ponto $x_0 \in U$, podemos restringir a função f à recta que passa por x_0 e com sentido definido por v . A derivada “ao longo” desta recta chama-se *derivada dirigida*:

Definição 2.2.4 Define-se a derivada dirigida da função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ no ponto $x_0 \in U$, segundo o vector $v \in \mathbb{R}^n$ como sendo

$$D_v f(x_0) = \lim_{\lambda \rightarrow 0} \frac{f(x_0 + \lambda v) - f(x_0)}{\lambda}.$$

se o limite existir.

Este uma relação simples entre derivadas dirigidas relativamente a vectores com a mesma direcção (qual?). Daí “normalizarmos” as derivadas dirigidas considerando muitas vezes v como sendo unitário. Nesse caso designamos a derivada dirigida como *derivada direcciona*.

A definição de derivada dirigida é mais fraca do que a definição de função diferenciável. Com efeito há funções que não são diferenciáveis num determinado ponto mas que admitem derivadas dirigidas. Pode mesmo acontecer que uma função admita algumas (ou todas!) as derivadas dirigidas num determinado ponto mas que não seja sequer contínua nesse ponto.

Exemplo 2.2.2 Consideremos a função definida por

$$f(x, y) = \begin{cases} 1, & \text{se } x \notin \mathbb{Q}, \\ 0, & \text{se } x \in \mathbb{Q}. \end{cases}$$

Claramente esta função não é contínua. No entanto, ela admite derivada dirigida na direcção $(0, 1)$. Fixemos um ponto (x_0, y_0) . Se x_0 for racional teremos $f(x_0, y_0 + h) = 0$, para qualquer $h \in \mathbb{R}$. Deste modo

$$D_{(0,1)} f(x_0, y_0) = 0.$$

Analogamente se x_0 for irracional teremos $f(x_0, y_0 + h) = 1$, para todo o $h \in \mathbb{R}$. Pelo que também se terá

$$D_{(0,1)} f(x_0, y_0) = 0.$$

As derivadas direccionais de funções $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ na direcção dos eixos coordenados e no sentido crescente da coordenada são frequentemente utilizadas e por isso têm um nome especial: *derivadas parciais*.

Definição 2.2.5 Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. A derivada parcial de f em relação a x_i é definida, caso o limite exista, por

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = D_{\mathbf{e}_i} f(\mathbf{x}) = \lim_{h \rightarrow 0} \frac{f(\mathbf{x} + \lambda \mathbf{e}_i) - f(\mathbf{x})}{\lambda},$$

com $\mathbf{x} = (x_1, \dots, x_n)$ e sendo \mathbf{e}_i o versor da direção i . Por vezes usaremos a notação $D_i f$ em vez de $\frac{\partial f}{\partial x_i}$.

Analisando a definição facilmente se conclui que, em termos práticos, a derivada parcial de f em ordem a x_i é calculada coordenada a coordenada se $m > 1$, o que permite lidar só com funções escalares, e, para cada uma destas, fixando todas as variáveis excepto x_i e derivando cada f_j em ordem a x_i como se esta fosse uma função real de variável real.

Exemplo 2.2.3 Seja $g(x, y) = (x^2 y^2, x)$. As derivadas parciais de g em ordem a x e y são

$$\frac{\partial g}{\partial x} = (2xy^2, 1) \quad \frac{\partial g}{\partial y} = (2x^2 y, 0).$$

Exercício 2.2.4 Calcule a derivada parcial em ordem a y das seguintes funções

1. $f(x, y, z) = xyz$;
2. $f(x, y) = x^2 + \text{sen}(xy)$;
3. $f(x, y, z, w) = 0$.

Se uma função é diferenciável as derivadas parciais permitem construir facilmente a matriz representando a derivada.

Proposição 2.2.1

Se uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é diferenciável em \mathbf{a} então a derivada $Df(\mathbf{a})$ satisfaz $Df(\mathbf{a})(h) = J_f(\mathbf{a})h$ em que é a matriz jacobiana de f no ponto \mathbf{a} definida por

$$J_f(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_m}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial f_n}{\partial x_m}(\mathbf{a}) \end{bmatrix}.$$

A diferenciabilidade de uma função pode ser estabelecida facilmente à custa da continuidade das derivadas parciais:

Definição 2.2.6 Diz-se que uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ com U aberto é de classe $C^1(U)$ se existirem as derivadas parciais

$$\frac{\partial f_j}{\partial x_i}, \quad 1 \leq j \leq m, \quad 1 \leq i \leq n$$

e forem contínuas. Se U não for aberto dizemos que $f \in C^1(U)$ se existir um aberto $V \supset U$ e uma função $g : V \rightarrow \mathbb{R}^m$ tal que $g|_U = f$ e $g \in C^1(V)$.

Exemplo 2.2.4 A função $f(x, y) = x^2 y^2$ é de classe C^1 pois as suas derivadas parciais são contínuas (veja exemplo 2.2.3).

Exemplo 2.2.5 Calculemos a derivada da função

$$f(x, y, z, w) = (f_1, f_2, f_3) = (x + y, x + y + z^2, w + z).$$

Aplicando os resultados e observações anteriores temos

$$J_f = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} & \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} & \frac{\partial f_2}{\partial w} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} & \frac{\partial f_3}{\partial w} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 2z & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

pelo que a função é C^1 , logo diferenciável e a derivada é representada pela matriz J_f .

Proposição 2.2.2 (C^1 implica diferenciabilidade)

Uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ de classe $C^1(U)$ com U aberto é diferenciável em U .

Ideia da demonstração. Claro que basta supor $m = 1$. Além disso consideramos $n = 2$ pois tal permite usar notação mais simples e quando terminarmos será óbvio como generalizar para $n > 2$.

Seja $(x, y) \in U$. Basta provar que

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(x+h, y+k) - f(x, y) - h \frac{\partial f}{\partial x}(x, y) - k \frac{\partial f}{\partial y}(x, y)}{(h^2 + k^2)^{1/2}} = 0.$$

Para tal decomponemos a diferença $f(x+h, y+k) - f(x, y)$ como uma soma de parcelas de diferenças de valores de f em que em cada parcela os argumentos de f só diferem numa coordenada. Uma escolha possível é

$$f(x+h, y+k) - f(x, y) = [f(x+h, y+k) - f(x, y+k)] + [f(x, y+k) - f(x, y)].$$

Podemos assim lidar separadamente com cada coordenada reduzindo o nosso objectivo a provar

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(x+h, y+k) - f(x, y+k) - h \frac{\partial f}{\partial x}(x, y)}{(h^2 + k^2)^{1/2}} = 0, \quad (2.1)$$

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(x, y+k) - f(x, y) - k \frac{\partial f}{\partial y}(x, y)}{(h^2 + k^2)^{1/2}} = 0. \quad (2.2)$$

Para lidar com (2.1) use o teorema de Lagrange, aplicado a $g(t) = f(x+t, y+k) - f(x, y+k)$, para obter que existe θ , $0 < \theta < 1$, tal que $f(x+h, y+k) - f(x, y+k) = h \frac{\partial f}{\partial x}(x+\theta h, y+k)$ e use a continuidade da derivada parcial. Para lidar com (2.2) pode usar um raciocínio análogo ou simplesmente a definição de derivada parcial.

Problema 2.2.2 Verifique que a demonstração da proposição 2.2.2 permite enunciar o resultado sob hipóteses mais gerais. Dê um exemplo de uma função que satisfaça tais hipóteses e não seja C^1 . Altere a demonstração para obter o caso $n > 2$.

Exercício 2.2.5 Mostre que são diferenciáveis e calcule a derivada das seguintes funções:

1. $f(x, y, z) = (x^2 - y^2, xy)$
2. $f(x, y) = (x - y, x + y, 2x + 3y)$
3. $f(x, y) = (\text{sen}(x + y), \cos(x - y))$
4. $f(x, y) = (e^{x+y+z}, \log(1 + e^y), z^2 + x)$

No caso de funções escalares ($m = 1$) a derivada é representada por uma matriz linha que se identifica a um vector de \mathbb{R}^n que merece um nome especial pela sua importância no cálculo diferencial e nas aplicações.

Definição 2.2.7 Suponha que uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ possui todas as derivadas parciais num ponto $\mathbf{a} \in U$. Defina-se o gradiente de f em \mathbf{a} , $\nabla f(\mathbf{a})$, via

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

Exercício 2.2.6 Verifique que se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável em $\mathbf{a} \in U$ então:

1. $Df(\mathbf{a})(\mathbf{h}) = D_{\mathbf{h}}f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{h}$;
2. $\sup_{\|\mathbf{h}\|=1} D_{\mathbf{h}}f(\mathbf{a}) = \|\nabla f(\mathbf{a})\|$.

Exercício 2.2.7 Mostre que a derivada da composição $f \circ g$ das transformações lineares $f(y) = Ay$, $g(x) = Bx$, onde $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^p \rightarrow \mathbb{R}^n$ e A, B são matrizes reais $m \times n$ e $n \times p$, respectivamente, é a matriz AB .

O próximo teorema fornece um método de cálculo da derivada de funções obtidas por composição. Note que para aplicações lineares a demonstração é trivial (exercício 2.2.7) e sugere o resultado geral: a derivada da composta é a composta das derivadas. Mais precisamente:

Teorema 2.2.3 (Derivação da Função Composta ou Regra da Cadeia)

Sejam $f : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ e $g : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$, funções diferenciáveis, $\mathbf{a} \in U$, $f(\mathbf{a}) \in V$ com U e V abertos. Então $f \circ g : U \cap f^{-1}(V) \rightarrow \mathbb{R}^m$ é diferenciável em \mathbf{a} e verifica-se:

$$D(f \circ g)(\mathbf{a}) = Df(g(\mathbf{a})) \circ Dg(\mathbf{a}).$$

Se f e g forem de classe C^1 então h é de classe C^1 .

De um ponto de vista de cálculo as derivadas parciais da composta são calculáveis em termos das derivadas parciais das funções que definem a composição usando o resultado anterior e o facto de à composição de aplicações lineares corresponder o produto de matrizes que as representam. Assim é importante compreender exemplos cujo protótipo mais simples é do tipo seguinte:

Exemplo 2.2.6 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $g = (g_1, g_2) : \mathbb{R} \rightarrow \mathbb{R}^2$. Se f e g forem diferenciáveis então

$$\frac{d(f \circ g)}{dt}(t) = \frac{\partial f}{\partial x_1}(g_1(t), g_2(t)) \frac{dg_1}{dt}(t) + \frac{\partial f}{\partial x_2}(g_1(t), g_2(t)) \frac{dg_2}{dt}(t).$$

Um outro exemplo do mesmo género é:

Exemplo 2.2.7 Seja $f(x, y) = (x + y, x - y)$ e $g(t_1, t_2, t_3) = (t_1 + 2t_2, t_2 + 2t_3)$. f e g são diferenciáveis. A derivada de $f \circ g$ é

$$\begin{aligned} D(f \circ g)(t_1, t_2, t_3) &= Df(g(t_1, t_2, t_3)) Dg(t_1, t_2, t_3) = \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & -2 \end{bmatrix}. \end{aligned}$$

Quando não há risco de confusão sobre os pontos em que se calculam as diversas derivadas parciais é comum abreviar uma fórmula como a do exemplo 2.2.6 como segue:

$$\frac{d}{dt}(f \circ g) = \frac{\partial f}{\partial x_1} \frac{dg_1}{dt} + \frac{\partial f}{\partial x_2} \frac{dg_2}{dt}$$

ou

$$\frac{d}{dt}(f \circ g) = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \frac{dx_2}{dt}.$$

Há risco de confusão em situações como a seguinte:

Exercício 2.2.8 Suponha que $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ é diferenciável, $f(0,1) = 0$ e $f(1,0) = 0$. Seja $g(x,y) = f(f(x,y), f(y,x))$. Calcule

$$\frac{\partial g}{\partial x}(0,1)$$

em termos de derivadas parciais de f em pontos convenientes. Convinha-lhe usar a notação $D_i f$ para evitar ambiguidades.

Exercício 2.2.9 Calcule a derivada da composição $h = f \circ g$ nos seguintes casos:

1. $f(x,y,z) = x^2 + y^2 + z^2$ e $g(t) = (t, 2t, 3t)$
2. $f(x,y) = (xy^5 + y \operatorname{ch} y^2, x \operatorname{tg}(\operatorname{sh} x^2) + 3y, x - y)$ e $g(t) = (3, 4)$.

Exercício 2.2.10 Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : [a,b] \rightarrow U$ diferenciáveis tais que f é constante no contradomínio de g . Mostre que $\nabla f(g(t)) \cdot g'(t) = 0$ para todo $t \in [a,b]$. Interprete este resultado como significando que, para funções diferenciáveis, o gradiente é ortogonal aos conjuntos de nível da função.

O teorema de derivação da função composta permite generalizar alguns resultados com facilidade à custa de resultados já conhecidos para funções reais de variável real. Por exemplo o teorema de Lagrange para funções escalares em que se relaciona a diferença entre os valores de uma função em dois pontos e a derivada no segmento de recta⁴ que os une.

Teorema 2.2.4 (do valor médio ou de Lagrange)

Sejam $U \subset \mathbb{R}^n$ um aberto e $f : U \rightarrow \mathbb{R}$ uma função diferenciável. Se $\mathbf{x}, \mathbf{y} \in U$ e $L(\mathbf{x}, \mathbf{y}) \subset U$ então existe $\theta \in]0, 1[$ tal que

$$f(\mathbf{y}) - f(\mathbf{x}) = \nabla f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x}).$$

Exercício 2.2.11 Prove o teorema do valor médio. Sugestão: considere a função de variável real $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$ e aplique o teorema do valor médio para funções a uma variável.

2.2.1 Exercícios suplementares

Exercício 2.2.12 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

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

- a) Determine justificadamente o maior subconjunto do domínio de f em que esta função é contínua.
- b) Uma função $H : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ verifica $H(0,1) = (1, -1)$ é diferenciável em $(0,1)$ sendo a matriz jacobiana de H nesse ponto dada por

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

Calcule a derivada dirigida $D_{(1,1)}(f \circ H)(0,1)$.

***Exercício 2.2.13** Se $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ está definida por

$$f(x,y) = \begin{cases} \frac{x^3-y^3}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0). \end{cases}$$

⁴Dados $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ define-se o segmento de recta unindo \mathbf{x} a \mathbf{y} como sendo o conjunto $L(\mathbf{x}, \mathbf{y}) = \{\mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}) : t \in [0, 1]\}$.

- a) Calcule o valor máximo de $D_{\mathbf{h}}f(1,2)$ quando \mathbf{h} é um vector unitário.
- b) Calcule a equação do plano tangente ao gráfico de f no ponto $(x, y, z) = (1, 2, -7/5)$.
- *c) Decida justificadamente se o gráfico de f constitui ou não uma variedade diferenciável. Se optar pela negativa determine o maior subconjunto do gráfico de f que efectivamente constitui uma variedade diferenciável. Em qualquer caso determine justificadamente a dimensão da variedade e o espaço normal no ponto $(1, 2, -7/5)$.

Exercício 2.2.14 Calcule as derivadas parciais de primeira ordem de

1. $f(x, y, z) = x^2 + y^2 + z^2$
2. $f(x, y) = \text{sen}(\text{sen}(\text{sen}(\text{sen}(x + y))))$
3. $f(x, y) = \int_0^{x+y} e^{-s^2} ds$

Exercício 2.2.15 Seja $f(x, y) = y \text{sen}(x^2 + \text{arctg}(y - \cos(x))) + 2$. Calcule $\frac{\partial f}{\partial x}(0, 0)$.

Exercício 2.2.16 Mostre que as seguintes funções são diferenciáveis e calcule as suas derivadas:

1. $f(x, y) = (x^2 + y, x - y)$
2. $f(x, y) = (x \int_0^y e^{\cos(s)} ds, y \int_0^x e^{\cos(s)} ds)$

Exercício 2.2.17 Calcule a derivada de $f \circ g$ nos seguintes casos:

1. $f(x, y, z) = x^2 + y^2 + z^2$ e $g(t) = (\text{sen}(t), \cos(t), 0)$;
2. $f(x, y) = (x + y, x - y)$ e $g(u, v) = (v, u)$;
3. $f(x, y, z, w) = \cos(e^{(x^2+y^2)} - z - w)$ e $g(p, q) = (0, 1, 2, 3)$.

2.2.2 Sugestões para os exercícios

2.2.14

- a) $\frac{\partial f}{\partial x} = 2x$, $\frac{\partial f}{\partial y} = 2y$ e $\frac{\partial f}{\partial z} = 2z$. Observe que o vector $(2x, 2y, 2z)$ é ortogonal à fronteira das bolas centradas em 0, isto é às esferas de equação da forma $x^2 + y^2 + z^2 = c$. Isto não é uma coincidência mas sim uma consequência do que foi aflorado no exercício 2.2.10 e que retomaremos!
- b) $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \cos(\text{sen}(\text{sen}(\text{sen}(x + y)))) \cos(\text{sen}(\text{sen}(x + y))) \cos(\text{sen}(x + y)) \cos(x + y)$;
- c) $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = e^{-(x+y)^2}$ (observe que não é necessário calcular o integral).

2.2.15 Observe que $f(x, 0) = 2$.

2.2.16 Ambas as funções são de classe C^1 , pois as derivadas parciais são contínuas. Portanto:

1. $Df = \begin{bmatrix} 2x & 1 \\ 1 & -1 \end{bmatrix}$.
2. $Df = \begin{bmatrix} \int_0^y e^{\cos(s)} ds & x e^{\cos(y)} \\ y e^{\cos(x)} & \int_0^x e^{\cos(s)} ds \end{bmatrix}$

2.2.17

1. Observe que $(f \circ g)(t) = 1$ para qualquer t .

2. Pela regra da cadeia temos:

$$D(f \circ g) = DfDg = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

3. Note que $Dg = 0$ pelo que $D(f \circ g) = 0$.



2.3 Derivadas parciais de ordem superior à primeira

Vamos considerar com *derivadas parciais de ordem superior à primeira* que, no essencial, se definem recursivamente.

Definição 2.3.1 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$. As derivadas parciais de segunda ordem, com respeito a x_i e x_j , $1 \leq i, j \leq n$, são definidas por*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j},$$

caso a expressão da direita esteja definida. Se $i = j$ escreve-se $\frac{\partial^2 f}{\partial x_i \partial x_i} = \frac{\partial^2 f}{\partial x_i^2}$. Proceda-se de modo análogo para derivadas parciais de ordem superior à segunda.

Exemplo 2.3.1 *Uma notação como*

$$\frac{\partial^4 u}{\partial x \partial y^2 \partial z}$$

indica que a função u foi derivada sucessivamente em ordem à variável z , duas vezes em ordem a y e finalmente em ordem a x .

Exemplo 2.3.2 *Seja $f(x, y) = x^2 + 2y^2 + xy$. Temos*

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial x} (4y + x) = 1.$$

Exemplo 2.3.3 *Seja $f(x, y, z) = \sin(x + y + z)$*

$$\begin{aligned} \frac{\partial^5 f}{\partial x^2 \partial y \partial z \partial y} &= \frac{\partial^4}{\partial x^2 \partial y \partial z} (\cos(x + y + z)) = -\frac{\partial^3}{\partial x^2 \partial y} (\sin(x + y + z)) = \\ &= -\frac{\partial^2}{\partial x^2} (\cos(x + y + z)) = \frac{\partial}{\partial x} (\sin(x + y + z)) = \cos(x + y + z). \end{aligned}$$

Exercício 2.3.1 *Seja $f(x, y) = x^2 + 2y^2 + xy$. Calcule $\frac{\partial^2 f}{\partial y \partial x}$; observe que o resultado é o mesmo do exemplo 2.3.2.*

O resultado deste último exercício ser o mesmo do exemplo 2.3.2 não é uma coincidência mas sim a consequência de um facto mais geral — o Teorema de Schwarz. Antes de o enunciarmos precisamos de uma definição:

Definição 2.3.2 *Considere uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$.*

- *Se U for aberto diz-se que f é de classe C^k em U , $k \in \mathbb{N}$, ou abreviadamente $f \in C^k(U)$, se todas as derivadas parciais de ordem k de f existirem e forem contínuas em U .*

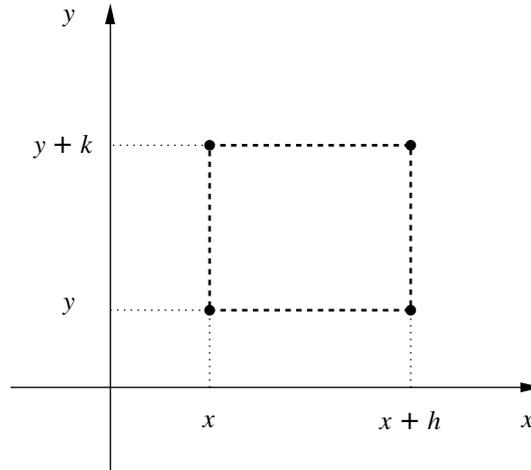


Figura 2.2: Convenções na demonstração da Proposição 2.2.2 e do Teorema 2.3.1.

- Se U não for aberto escrevemos $f \in C^k(U)$, $k \in \mathbb{N}$, se existir V aberto com $V \supset U$ e uma função $g \in C^k(V)$ tal que a restrição de g a U seja igual a f .
- f diz-se de classe $C^0(U)$ se for contínua em U .
- Adicionalmente, para U aberto, definimos $C^\infty(U) = \bigcap_{k \in \mathbb{N}} C^k(U)$ e para um conjunto não necessariamente aberto procedemos como anteriormente.

Na maior parte das aplicações do cálculo diferencial a hipótese de uma função ser de classe C^k para um certo k é natural. Certos resultados a citar a seguir serão válidos sob hipóteses mais gerais mas abstermo-nos de dar importância especial a tais hipóteses. Por vezes serão remetidas para problemas.

Exercício 2.3.2 Seja $p(x_1, \dots, x_n)$ um polinómio em n variáveis. Mostre que $\text{sen}(p(x_1, \dots, x_n))$ é uma função $C^\infty(\mathbb{R}^n)$.

Problema 2.3.1 Verifique que se $j < k$ então $C^k \subset C^j$.

O próximo teorema é um resultado muito importante que permite reduzir o número de cálculos necessários para determinar as derivadas parciais de ordem superior á primeira. Ele diz-nos que, sob certas condições, a ordem pela qual se deriva uma função é irrelevante.

Teorema 2.3.1 (Schwarz)

Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, \mathbf{a} um ponto interior a U , $f \in C^2(U)$. Então $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a})$ para quaisquer índices $1 \leq i, j \leq n$.

Ideia da demonstração. Basta considerar $n = 2$ e convencionamos $\mathbf{a} = (x, y)$. Notamos que

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{[f(x+h, y+k) - f(x+h, y)] - [f(x, y+k) - f(x, y)]}{hk} \quad (2.3)$$

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{[f(x+h, y+k) - f(x, y+k)] - [f(x+h, y) - f(x, y)]}{hk} \quad (2.4)$$

Designemos o numerador das fracções dos segundos membros de (2.3-2.4) por $D(h, k)$. Aplicando o teorema de Lagrange à função $g(t) = f(x+t, y+k) - f(x+t, y)$ no intervalo $[0, h]$ obtemos que

existe θ_1 , $0 < \theta_1 < 1$, tal que

$$D(h, k) = h \left[\frac{\partial f}{\partial x}(x + \theta_1 h, y + k) - \frac{\partial f}{\partial x}(x + \theta_1 h, y) \right].$$

Uma segunda aplicação do teorema de Lagrange permite obter que existe θ_2 , $0 < \theta_2 < 1$, tal que

$$D(h, k) = hk \frac{\partial^2 f}{\partial y \partial x}(x + \theta_1 h, y + \theta_2 k).$$

Substituição em (2.3) e justificação de que ambos os limites iterados igualam $\lim_{(h,k) \rightarrow (0,0)} D(h, k)$ permitem obter a igualdade pretendida.

Problema 2.3.2 *O último passo da demonstração da Proposição 2.3.1 merece alguns comentários. Por um lado θ_1 e θ_2 são funções de h e k . Por outro a relação entre um limite e um limite iterado é, em geral, mais complexa do que o leitor pode imaginar. Seja $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e (x_0, y_0) um ponto interior de U . Mostre que:*

- Pode existir $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$ sem que exista $\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y)$.*
- Se $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$ e $\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y)$ existirem então são iguais.*

Problema 2.3.3 *É óbvio da demonstração da Proposição 2.3.1 que a hipótese $f \in C^2$ pode ser aligeirada. Isto pode ser feito de várias formas. Formule e demonstre pelo menos dois resultados deste tipo com hipóteses “mínimas” não equivalentes.*

Exemplo 2.3.4 *Seja $f = 2xy$. f é de classe C^2 uma vez que é um polinómio, portanto temos a seguinte igualdade*

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

Exemplo 2.3.5 *Se f é de classe C^3 têm-se as seguintes igualdades:*

$$\frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x^2}$$

e

$$\frac{\partial^3 f}{\partial y^2 \partial x} = \frac{\partial^3 f}{\partial y \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y^2}.$$

Exercício 2.3.3 *Calcule as derivadas de todas as ordens de $f(x, y, z) = 2x^3z + xyz + x + z$ (observe que só há um número finito de derivadas não nulas. Porquê?).*

O conceito de derivada dirigida de ordem superior à primeira permite formalizar o enunciado da fórmula de Taylor de uma forma análoga ao resultado já conhecido para funções reais de variável real.

Definição 2.3.3 *Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. As derivadas dirigidas de ordem superior à primeira de f num ponto $\mathbf{x} \in U$ segundo \mathbf{h} definem-se recursivamente, se existirem, por $D_{\mathbf{h}}^{(1)} f(\mathbf{x}) = D_{\mathbf{h}} f(\mathbf{x})$ e*

$$D_{\mathbf{h}}^{(j)} f(\mathbf{x}) = D_{\mathbf{h}}(D_{\mathbf{h}}^{(j-1)} f(\mathbf{x})), \quad \text{se } j > 1.$$

Relembra-se que para funções diferenciáveis, e em particular de classe C^1 , temos $D_{\mathbf{h}} f(\mathbf{x}) = \mathbf{h} \cdot \nabla f(\mathbf{x})$.

Problema 2.3.4 Verifique que para funções de classe C^j num aberto o cálculo da derivada dirigida $D_{\mathbf{h}}^{(j)} f$ corresponde a aplicar à função f o operador diferencial $(\mathbf{h} \cdot \nabla)^j$ e consequentemente $D_{\mathbf{h}}^{(j)} f$ é um polinómio homogéneo⁵ de grau j nas componentes do vector \mathbf{h} . Se $\mathbf{h} = (h_1, h_2)$ verifique que para $n = 2$ e $j = 2$ temos

$$D_{\mathbf{h}}^{(2)} f = h_1^2 \frac{\partial^2 f}{\partial x_1^2} + 2h_1 h_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + h_2^2 \frac{\partial^2 f}{\partial x_2^2}.$$

Em geral obtenha

$$D_{\mathbf{h}}^{(j)} f = \sum_{i_1=1}^n \cdots \sum_{i_j=1}^n h_{i_1} \cdots h_{i_j} \frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}}.$$

Note que existem termos “repetidos” na fórmula anterior. Calcular o número de repetições é um problema de cálculo combinatório cuja solução no caso $n = 2$ é bem conhecida.

2.3.1 Exercícios suplementares

Exercício 2.3.4 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por:

$$f(x, y) = \begin{cases} xy, & \text{se } |y| > |x|, \\ 0, & \text{caso contrário.} \end{cases}$$

Mostre que:

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 0 \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = 1.$$

Explique porque é que isto não contradiz o teorema 2.3.1.

Exercício 2.3.5 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função limitada (não necessariamente contínua). Mostre que

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

é diferenciável na origem. Calcule a sua derivada. Dê um exemplo de uma função f tal que g não seja contínua no complementar da origem.

Exercício 2.3.6 Suponha $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, f bijectiva, diferenciável e f^{-1} também diferenciável. Mostre que $Df^{-1}(f(\mathbf{x})) = [Df(\mathbf{x})]^{-1}$. Use esta observação para, por exemplo, rededuzir a fórmula da derivada de arcsen.

2.3.2 Sugestões para os exercícios

2.3.4 O teorema 2.3.1 só se aplicaria se a função f fosse de classe C^2 . ◀

2.3.5 Use a definição de derivada para mostrar que g é diferenciável com derivada representada por $\nabla g(0, 0) = (1, 1)$. Para a segunda parte um exemplo possível é

$$f(x, y) = \begin{cases} 1, & \text{se } x \in \mathbb{Q}, \\ 0, & \text{caso contrário.} \end{cases}$$

2.3.6 Observe que $f(f^{-1}(x)) = x$. Diferencie esta expressão. $\frac{d}{dy}(\arcsen y) = \frac{1}{\sqrt{1-y^2}}$. ◀

⁵Um polinómio P de grau k diz-se homogéneo se $P(\lambda x) = \lambda^k P(x)$ para todo o $\lambda \in \mathbb{R}$.

2.4 Polinómio de Taylor

Tal como no caso de funções reais de variável real podemos construir aproximações polinomiais de funções de classe C^k .

Teorema 2.4.1 (Taylor)

Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe $C^k(U)$ com U um aberto e $\mathbf{x}_0 \in U$. Para cada $j \leq k$ existe um polinómio em n variáveis de grau j , único, $P_j : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - P_j(\mathbf{x})}{|\mathbf{x} - \mathbf{x}_0|^j} = 0. \quad (2.5)$$

O polinómio P_j é designado por polinómio de Taylor de ordem j de f relativo ao ponto \mathbf{x}_0 e é dado por

$$P_j(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{l=1}^j \frac{1}{l!} D_{\mathbf{x}-\mathbf{x}_0}^{(l)} f(\mathbf{x}_0). \quad (2.6)$$

O erro $E_j(x)$ da fórmula de Taylor é dado por

$$E_j(x) = f(x) - P_j(x).$$

Ideia da demonstração. Decorre do resultado já conhecido para $n = 1$ e do teorema de derivação da função composta por consideração da função auxiliar $g : [0, 1] \rightarrow \mathbb{R}$ definida por $g(t) = f(t(\mathbf{x} - \mathbf{x}_0) + \mathbf{x}_0)$ em que $\mathbf{x} \in B_r(\mathbf{x}_0) \subset U$.

Problema 2.4.1 Use o problema 2.3.4 para obter a fórmula de Taylor na forma:

$$f(\mathbf{x}) = \sum_{p=0}^k \sum_{i_1+\dots+i_n=p} \frac{1}{p!} \frac{\partial^p f}{\partial y_1^{i_1} \dots \partial y_n^{i_n}}(\mathbf{x}_0) (x_1 - x_{01})^{i_1} \dots (x_n - x_{0n})^{i_n} + E_k(\mathbf{x} - \mathbf{x}_0). \quad (2.7)$$

O leitor é aconselhado a pensar no polinómio de Taylor via a propriedade (2.5) e não simplesmente como um polinómio calculável via (2.6) ou (2.7).

Problema 2.4.2 Formule o Teorema de Taylor explicitando o resto da fórmula de Taylor numa forma análoga a uma das conhecidas para funções reais de variável real.

Poderá pensar-se que o cálculo do polinómio de Taylor para funções de várias variáveis e para uma ordem relativamente elevada é um pesadelo computacional. Nem sempre será assim se tirarmos partido, quando possível, de resultados já conhecidos para funções de uma variável.

Frequentemente em vez de escrevermos o termo de erro $E_k(\mathbf{x} - \mathbf{y})$, escrevemos $o(\|\mathbf{x} - \mathbf{y}\|^k)$, com o mesmo significado.

Exemplo 2.4.1 Se $f(x, y) = xy + \sin x$, a fórmula de Taylor de segunda ordem em torno de $(\pi, 0)$ é:

$$\begin{aligned} f(x, y) &= f(\pi, 0) + \frac{\partial f}{\partial x} \Big|_{(\pi, 0)} (x - \pi) + \frac{\partial f}{\partial y} \Big|_{(\pi, 0)} y + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \Big|_{(\pi, 0)} (x - \pi)^2 \\ &\quad + \frac{\partial^2 f}{\partial x \partial y} \Big|_{(\pi, 0)} (x - \pi)y + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} \Big|_{(\pi, 0)} y^2 + o(\|(x - \pi, y)\|^2), \end{aligned}$$

ou seja

$$f(x, y) = \pi - x + xy + o(\|(x - \pi, y)\|^2).$$

Exemplo 2.4.2 Se $f(x, y) = x^2 + 2xy + y^2$ então a sua expansão em fórmula de Taylor até à segunda ordem, em torno de qualquer ponto, é $x^2 + 2xy + y^2$. Com efeito, $f(x, y) - x^2 - 2xy - y^2 = 0$ pelo que (2.8) vale. Repare que isto evitou termos de calcular 5 derivadas!

Exercício 2.4.1 Calcule a fórmula de Taylor até à terceira ordem das seguintes funções:

1. $f(x, y, z) = x + y^2 + z$;
2. $f(x, y, z) = 1 + x + y + z + xy + xz + yz + xyz$;
3. $f(x, y) = e^x + xyz$.

Exercício 2.4.2 Mostre que a fórmula de Taylor de ordem k para um polinómio de grau k coincide com o polinómio.

Exercício 2.4.3 Demonstre a parte correspondente a unicidade do teorema de Taylor. [Suponha que existe um polinómio $p(x)$ para o qual (2.8) vale. Mostre que se existisse outro polinómio $q(x) \neq p(x)$, de grau menor ou igual ao grau de p obteríamos uma contradição.]

Em certos casos podemos utilizar o conhecimento da expansão em potências de uma função real de variável real para calcularmos a expansão em potências de expressões mais complicadas:

Exemplo 2.4.3 Queremos calcular a expansão de Taylor da função $\text{sen}(x^2 + y^4)$ até à ordem 6 em torno da origem. Sabemos que

$$\text{sen } t = t - \frac{t^3}{6} + o(|t|^3).$$

Deste modo temos

$$\text{sen}(x^2 + y^4) = x^2 + y^4 - \frac{(x^2 + y^4)^3}{6} + o((x^2 + y^4)^3)$$

pelo que

$$\text{sen}(x^2 + y^4) = x^2 + y^4 - \frac{x^6}{6} + o(\|(x, y)\|^6),$$

em que na última igualdade tivemos em atenção que $(x^2 + y^4)^3 = x^6 + 3x^4y^4 + 3x^2y^8 + y^{12} = x^6 + o(\|(x, y)\|^6)$ e $x^2 + y^4 \leq x^2 + y^2$ para $\|(x, y)\|$ suficientemente pequeno.

Exemplo 2.4.4 Seja

$$g(x, y) = \text{sen}(x^2 - y^2).$$

e suponhamos que pretendemos obter o polinómio de Taylor de sétima ordem de g relativo a $(0, 0)$.

Sabemos que o seno é uma função inteira cuja série de Taylor relativa a 0 (série de Mac Laurin) é

$$\text{sen } \lambda = \lambda - \frac{\lambda^3}{3!} + \frac{\lambda^5}{5!} - \dots + (-1)^{k+1} \frac{\lambda^{2k-1}}{(2k-1)!} + \dots$$

Tal permite-nos ter um palpite à cerca do polinómio de Taylor pretendido simplesmente por substituição formal de λ por $x^2 - y^2$ na igualdade anterior e só considerando os termos de grau menor ou igual a sete. Obtem-se um polinómio

$$Q(x, y) = (x^2 - y^2) - \frac{(x^2 - y^2)^3}{3!}$$

Resta provar que efectivamente se trata do polinómio de Taylor pretendido. Para tal usa-se a caracterização (2.5) do polinómio de Taylor. De facto

$$\lim_{\lambda \rightarrow 0} \frac{\text{sen } \lambda - \lambda + \frac{\lambda^3}{3!}}{\lambda^4} = 0$$

donde resulta

$$\lim_{(x,y) \rightarrow (0,0)} \frac{g(x,y) - Q(x,y)}{(x^2 - y^2)^4} = 0$$

e usando $|x^2 - y^2| \leq x^2 + y^2$ obtém-se

$$\lim_{(x,y) \rightarrow (0,0)} \frac{g(x,y) - Q(x,y)}{(x^2 + y^2)^4} = 0.$$

Assim Q é de facto o polinómio de Taylor pretendido e inclusivamente é idêntico ao polinómio de Taylor de oitava ordem. Note que obtivemos, por exemplo, que todas as derivadas parciais de ordens 1, 3, 4, 5, 7 e 8 de g em $(0,0)$ são nulas.

Exercício 2.4.4 Desenvolva em fórmula de Taylor $f(x,y) = e^{x^2+y^2}$ até à terceira ordem. Tente não calcular as derivadas directamente mas sim usar o facto de que o polinómio de Taylor de ordem k é o único polinómio de grau $\leq k$ tal que

$$\lim_{\|x-y\| \rightarrow 0} \frac{|f(x) - p(x)|}{\|x-y\|^k} = 0. \quad (2.8)$$

Exercício 2.4.5 Calcule a expansão em potências de $x-1$ e $y-2$ de

$$\text{sen}(x+y-3)$$

até à quarta ordem.

2.4.1 Exercícios suplementares

Exercício 2.4.6 Calcule a expansão de Taylor em torno do ponto $(1,1,1)$, até à quinta ordem de $xy + xyz + x^2 + y^2 + xyz$.

Exercício 2.4.7 Seja f uma função C^∞ . Desenvolva $\int_0^x f(s)ds$ em série de Taylor em torno de 0.

Exercício 2.4.8 Calcule a expansão em série de Taylor da função $e^{x^2 + \text{sen}((y-1)^2)}$ até à quarta ordem em torno de $x=0$ e $y=1$.

Exercício 2.4.9 Calcule a expansão em série de Taylor de

$$\text{sen}(x^{1000} + y^{1000} + z^{1000})$$

até à ordem 999 em torno da origem.

Exercício 2.4.10 Suponha que $f: \mathbb{R} \rightarrow \mathbb{R}$ e $v: \mathbb{R}^2 \rightarrow \mathbb{R}$ são de classe C^∞ e satisfazem

$$\begin{cases} \frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} \\ v(x,0) = f(x). \end{cases}$$

Desenvolva v em série de Taylor em torno da origem.

2.4.2 Sugestões para os exercícios

2.4.6 Neste caso a fórmula de Taylor coincide com o próprio polinómio $xy + xyz + x^2 + y^2 + xyz$ (veja o teorema 2.4.1). ◀

2.4.7 $\int_0^x f(s)ds = f(0)x + f'(0)\frac{x^2}{2} + \dots + f^{(n-1)}(0)\frac{x^n}{n!} + \dots$ ◀

2.4.8 Note que $\sin((y-1)^2) = (y-1)^2 + \frac{(y-1)^6}{6} + o(|y-1|^6)$ e que $e^t = 1 + t + \frac{t^2}{2} + o(t^3)$ pelo que $e^{x^2 + \sin((y-1)^2)} = 1 + x^2 + (y-1)^2 + (x^2 + (y-1)^2)^2 + o(\|(x, y-1)\|^4)$. ◀

2.4.9 Repare que $\sin(t) = t + o(t^2)$ para t numa vizinhança da origem. ◀

2.4.10 Note que, utilizando a equação, se tem $\frac{\partial v}{\partial t}(0,0) = \frac{\partial^2 f}{\partial x^2}(0)$, $\frac{\partial^2 v}{\partial t \partial x}(0,0) = \frac{\partial^3 f}{\partial x^3}(0)$. Use o método de indução. ◀

Capítulo 3

Extremos

Problemas envolvendo maximização ou minimização de funções envolvendo diversos parâmetros estão entre os mais importantes em Matemática. Aparecem frequentemente em física (por exemplo a mecânica lagrangeana), engenharia (maximizar a resistência de um mecanismo ou eficiência de um motor) ou economia (minimizar custos de produção ou otimizar investimentos). Neste capítulo vamos estudar métodos para determinar máximos e mínimos de funções definidas em subconjuntos de \mathbb{R}^n com valores em \mathbb{R} .

O leitor já deve conhecer que, para funções reais de variável real, os candidatos a pontos de extremo de entre os pontos interiores onde a função é diferenciável são exactamente aqueles onde a derivada se anula, chamados *pontos de estacionaridade*. A generalização deste facto para funções de mais de uma variável, a discutir mais à frente, são os pontos onde o gradiente da função se anula. Tal condição estabelece o chamado *sistema de estacionaridade* cujas soluções serão ainda conhecidas por pontos de estacionaridade.

O teorema de Taylor será utilizado para a classificação de pontos de estacionaridade de uma função de classe C^2 quanto a serem pontos de mínimo, máximo ou pontos de sela. Quanto a este último ponto é de notar que, num caso concreto, os critérios baseados na fórmula de Taylor poderão ser insuficientes por diversas razões e tal é abundantemente exemplificado nos exercícios¹

- Uma função pode ter um extremo num ponto onde não estão definidas algumas das derivadas parciais de primeira ordem.
- Uma função pode ter um extremo num ponto fronteiro do seu domínio.
- Uma função pode ter um extremo num ponto de estacionaridade não sendo de classe C^2 numa qualquer vizinhança desse ponto.
- Os critérios baseados na fórmula de Taylor podem ser inconclusivos.

Adicionalmente tais métodos pressupõem que o sistema de estacionaridade da função é explicitamente resolúvel o que, dado a sua não linearidade, é algo que em geral não se verificará.

Em tais casos uma sistematização de todos os possíveis métodos de ataque ao problema de determinação dos pontos de extremo local de uma função é impossível. Cremos no entanto que os raciocínios mais interessantes estão bem exemplificados a seguir.

Alguns dos métodos a utilizar pressupõem alguns conhecimentos de Álgebra Linear. Como  ência sugere-se [4].

¹Exemplos típicos para funções reais de variável real com o domínio da função o intervalo $[-1, 1]$: $x \mapsto |x|$, $x \mapsto x$, $x \mapsto |x|^{3/2}$, $x \mapsto \begin{cases} e^{-1/x^2} & \text{se } x \neq 0, \\ 0 & \text{caso contrário.} \end{cases}$

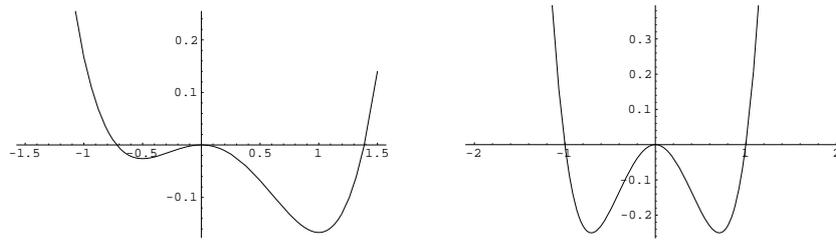


Figura 3.1: Os gráficos de $f(x) = -\frac{x^2+x^4}{4} - \frac{x^3}{6}$ e $g(x) = x^4 - x^2$.

3.1 Extremos

Provavelmente o leitor terá uma ideia intuitiva do que é um ponto de extremo de uma função, ou seja, um ponto de máximo ou de mínimo. Começaremos portanto por formalizar estas ideias do ponto de vista matemático. A primeira definição é a de máximo e mínimo local de uma função real.

Definição 3.1.1 *Seja $f : A \rightarrow \mathbb{R}$, com $A \subset \mathbb{R}^n$. Um ponto $\mathbf{x}_0 \in A$ é um ponto de máximo (resp. mínimo) local e $f(\mathbf{x}_0)$ máximo (resp. mínimo) local de f se existir uma vizinhança² V de \mathbf{x}_0 tal que, $\forall \mathbf{x} \in V \cap A$,*

$$f(\mathbf{x}) \leq f(\mathbf{x}_0), \quad (\text{resp. } f(\mathbf{x}) \geq f(\mathbf{x}_0)).$$

Note que, de acordo com a definição anterior, uma função pode ter vários extremos locais cada um deles ocorrendo em vários pontos de extremo local.

Exemplo 3.1.1 *Seja f a função definida em \mathbb{R} , constante igual a 1. Então qualquer número real é um ponto de máximo (e também mínimo) de f .*

O último exemplo ilustra a necessidade de distinguir estes casos degenerados de outros mais interessantes. Assim temos a seguinte definição.

Definição 3.1.2 *O máximo (resp. mínimo) é estrito se a igualdade na definição anterior só se verificar para $\mathbf{x} = \mathbf{x}_0$. O máximo (resp. mínimo) é global (ou absoluto) se, $\forall \mathbf{x} \in A$*

$$f(\mathbf{x}) \leq f(\mathbf{x}_0), \quad (\text{resp. } f(\mathbf{x}) \geq f(\mathbf{x}_0)).$$

Exemplo 3.1.2 *A função $f(x) = -\frac{x^2+x^4}{4} - \frac{x^3}{6}$ tem um máximo local em $x = 0$, um mínimo local em $x = -\frac{1}{2}$ e um mínimo absoluto em $x = 1$, como se pode observar na figura 3.1. A função $g(x) = x^4 - x^2$ tem um mínimo absoluto para $x = 1$. No entanto, este mínimo não é único pois $x = -1$ é outro ponto de mínimo absoluto tendo-se $g(1) = g(-1)$. Veja a figura 3.1.*

Exemplo 3.1.3 *Provemos que a função $f(x) = x^2$ tem um mínimo absoluto estrito na origem. Tal decorre de $f(0) = 0 < x^2 = f(x)$ para $x \neq 0$.*

Exercício 3.1.1 *Seja $f : A \rightarrow \mathbb{R}$, com $A = \{a\}$, o conjunto só com um ponto. Justifique que $x = a$ é ponto de mínimo e ponto de máximo estrito simultaneamente.*

Nem sempre dada uma função podemos garantir a existência de máximos ou mínimos, como se pode ver pelos exemplos seguintes:

²Por exemplo, uma bola de raio ϵ centrada em \mathbf{x}_0 .

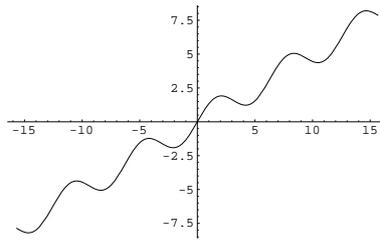


Figura 3.2: O gráfico de $f(x) = \frac{x}{2} + \text{sen } x$

Exemplo 3.1.4 Seja $f :]0, 1[\rightarrow \mathbb{R}$ definida por $f(x) = x$. Note que f não tem mínimo nem máximo pois não fazem parte do domínio os pontos 0 e 1 onde a função definida pela mesma fórmula mas cujo domínio fosse o intervalo fechado $[0, 1]$ atinge os seus valores extremos.

Exemplo 3.1.5 Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = \frac{x}{2} + \text{sen } x$. Embora f tenha máximos e mínimos locais (ver figura 3.2) f não tem nenhum máximo ou mínimo global pois $\lim_{x \rightarrow +\infty} f(x) = +\infty$ e $\lim_{x \rightarrow -\infty} f(x) = -\infty$.

Exemplo 3.1.6 Seja $f(x) = x^2$ se $x \in \mathbb{R} \setminus \{0\}$, $f(0) = 1$. Esta função não tem nenhum mínimo pois f nunca se anula embora f tome valores positivos arbitrariamente pequenos.

Exercício 3.1.2 Seja f a função do exemplo 3.1.6. Mostre que $f(0)$ é um máximo local mas não global.

Antes de prosseguirmos convém sumarizar informalmente o que aprendemos nos 3 últimos exemplos. A função do exemplo 3.1.4 não tem máximo nem mínimo porque retirámos os extremos a um intervalo limitado e fechado fazendo com que os valores extremos da função não sejam atingidos nesses pontos. No exemplo seguinte não encontramos extremos absolutos pois a função é ilimitada o que é possível graças para uma função contínua se o domínio não é compacto (neste caso não é limitado). Finalmente no último destes exemplos a função não tem mínimo porque ocorre uma descontinuidade no ponto onde o mínimo deveria ocorrer.

Estes exemplos sugerem que, para garantir a existência de extremos, seja usual tentar lidar com funções contínuas definidas em conjuntos limitados e fechados (compactos). O próximo teorema mostra que estas condições são efectivamente suficientes para garantir a existência de extremos:

Teorema 3.1.1 (Weierstrass)

Seja $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ contínua com A compacto. Então f tem máximo e mínimo (globais) em A .

Ideia da demonstração. Veja o exercício 2.1.16.

Ficamos assim com um critério abstracto para garantir a existência de máximos e mínimos, independentemente da aparência mais ou menos complicada da definição da função:

Exemplo 3.1.7 A função $f : [0, 1] \rightarrow \mathbb{R}$ dada por $f(x) = \frac{e^{\text{sen}(x+\log(x+1))}}{1+100x^2}$ é contínua e $[0, 1]$. Portanto tem pelo menos um ponto de máximo e um ponto de mínimo globais em $[0, 1]$.

Exemplo 3.1.8 Consideremos o subconjunto $K \subset \mathbb{R}^2$ definido pela condição $|x| + |y| \leq 1$. Seja f a função aí definida por $f(x, y) = x^2 + y^2$. Como K é compacto (porque é limitado e fechado), f tem de ter máximo e mínimo. Reparando que f é o quadrado da distância à origem concluímos que ocorre um mínimo (global) na origem. Os pontos de máximo serão os pontos do conjunto mais afastados da origem, que neste caso são $(\pm 1, 0)$ e $(0, \pm 1)$.

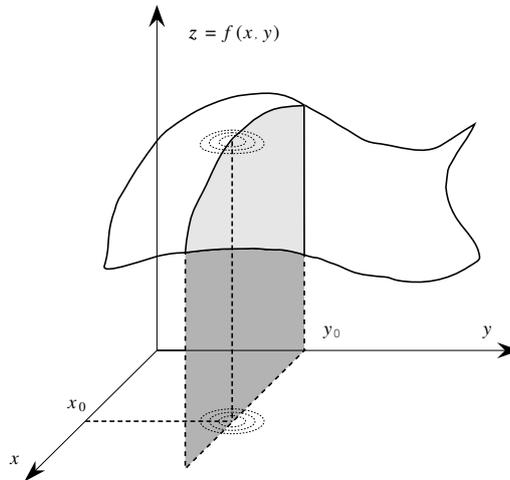


Figura 3.3: Fixar todas as variáveis excepto uma define uma função de uma variável. Se f tiver um máximo local em (x_0, y_0) e fixarmos a segunda variável em y_0 então tal função tem um máximo em x_0 .

Exercício 3.1.3 Diga em quais dos seguintes subconjuntos de \mathbb{R}^2 pode garantir a existência de mínimos para qualquer função contínua f . No caso de a resposta ser negativa apresente um exemplo.

1. $\text{máx}\{|x|, |y|\} = 1$
2. $\text{máx}\{|x|, |y|\} \leq 1$
3. $\text{máx}\{|x|, |y|\} \geq 1$
4. $\text{máx}\{|x|, |y|\} > 1$
5. $\text{máx}\{|x|, |y|\} < 1$

Exercício 3.1.4 Mostre que a função $f(x) = x^4$ tem mínimo e não tem máximo no intervalo $] -1, 1[$. Porque é que isto não contradiz o teorema de Weierstrass?

Em casos simples é possível seleccionar os candidatos a extremos utilizando raciocínios *ad hoc*. No exemplo 3.1.8, a função em questão é a distância à origem e por isso tem um mínimo em 0. No entanto, convém ter um critério, de aplicação fácil, que permita reduzir o número de candidatos a pontos de máximo ou mínimo a serem analisados. O resultado do próximo teorema permite fazer isto, daí a sua importância.

Definição 3.1.3 Seja $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável num ponto $\mathbf{a} \in \text{int } A$. Diz-se que \mathbf{a} é um ponto de estacionaridade (ou ponto crítico) de f se $\nabla f(\mathbf{a}) = 0$.

Teorema 3.1.2

Seja $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável num ponto $\mathbf{x} \in \text{int } A$. Se \mathbf{x} é ponto de extremo de f então é ponto de estacionaridade, ou seja $\nabla f(\mathbf{x}) = 0$.

Ideia da demonstração. Seja (x_1, \dots, x_n) um ponto de extremo duma função f e considere

$$g_i(t) = f(x_1, \dots, t, \dots, x_n).$$

g_i tem um extremo em $t = x_i$. Aplique o resultado conhecido em dimensão 1 a g_i no ponto x_i .

Exemplo 3.1.9 Suponhamos que pretendemos encontrar os extremos da função $f(x, y) = x^2 + y^2$ no conjunto $x^2 + y^2 < 1$. Como o conjunto é aberto todos os pontos de extremo de f (se existirem) serão interiores, pelo que nestes pontos o gradiente de f será nulo, isto é

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (0, 0).$$

Deste modo, resolvendo a equação

$$\nabla f = (2x, 2y) = (0, 0),$$

podemos determinar todos os possíveis extremos de f . Concluimos portanto, que o único ponto em que pode ocorrer um extremo é $(x, y) = (0, 0)$. Como $f(0, 0) = 0$ e a função é sempre positiva em todos os outros pontos este será necessariamente um mínimo (absoluto) de f .

O teorema anterior e o teorema de Weierstrass implicam um critério de detecção de pontos de extremo que resumizamos no seguinte corolário:

Corolário 3.1.3

Seja $f : A \rightarrow \mathbb{R}$, A compacto (limitado e fechado) e f contínua. Então f tem pelo menos um ponto de máximo e um ponto de mínimo global. Para além disso, os únicos pontos que podem ser extremos de f são

1. pontos na fronteira de A ;
2. pontos onde $\nabla f = 0$;
3. pontos onde f não é diferenciável.

Exercício 3.1.5 Determine (se existirem) os máximos e mínimos das seguintes funções:

1. $f(x, y) = x^4 + y^4$ em $|x| + y^2 < 1$.
2. $f(x, y) = x^2 - y^2$ no conjunto $x^2 + y^2 < 1$.
3. $f(x, y) = xy$ em $|x| + |y| < 1$.
4. $f(x, y) = \sqrt{x^2 + y^2}$ em $x^2 + y^2 < 1$.

Porém nem todos os pontos críticos de uma função são máximos ou mínimos. Isto motiva a seguinte definição:

Definição 3.1.4 Diz-se que um ponto de estacionaridade \mathbf{a} é um ponto de sela de uma função f se qualquer que seja a vizinhança de \mathbf{a} existirem pontos nessa vizinhança onde a função toma valores inferiores e superiores a $f(\mathbf{a})$.

Exemplo 3.1.10 Seja $f(x) = x^3$ então 0 é um ponto de sela de f pois embora seja um ponto crítico de f ($f'(x) = 3x^2$ anula-se na origem) não se trata de um ponto de máximo ou mínimo (porque $f(x) < f(0)$ para $x < 0$ e $f(x) > f(0)$ para $x > 0$).

Exercício 3.1.6 Verifique que $(0, 0)$ é um ponto de sela³ de $x^2 - y^2$.

No exemplo 3.1.9 e no exercício 3.1.5 os conjuntos onde as funções estavam definidas eram abertos. Consequentemente todos os pontos de extremo eram pontos de estacionaridade. Não é este o caso do próximo exemplo, onde nos temos de preocupar com a possibilidade de haver máximos ou mínimos que, por estarem na fronteira do domínio, não sejam pontos de estacionaridade.

³A expressão *ponto de sela* é motivada pelos gráficos de funções em exemplos como este. Claro que acabamos por usar a expressão em situações mais gerais.

Exemplo 3.1.11 Suponhamos que queremos determinar os extremos da função

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

no quadrado $[-1, 1] \times [-1, 1]$.

O gradiente de f é dado por

$$\nabla f = (y(1 - x^2 - y^2) - 2x^2y, x(1 - x^2 - y^2) - 2xy^2).$$

Os pontos de estacionaridade estarão entre as soluções de

$$\begin{cases} y - 3x^2y - y^3 = 0 \\ x - 3xy^2 - x^3 = 0 \end{cases} \quad (3.1)$$

no interior do quadrado, isto é, verificando simultaneamente $-1 < x < 1$ e $-1 < y < 1$. O sistema (3.1) admite como soluções:

1. $(x, y) = (0, 0)$;
2. $x = 0, y \neq 0$ e portanto $1 - y^2 = 0$, ou seja $(x, y) = (0, \pm 1)$;
3. $x \neq 0, y = 0$ e portanto $1 - x^2 = 0$, ou seja $(x, y) = (\pm 1, 0)$;
4. pontos que verifiquem $x \neq 0, y \neq 0$ e

$$\begin{cases} 3x^2 + y^2 = 1 \\ x^2 + 3y^2 = 1. \end{cases} \quad (3.2)$$

O sistema 3.2 não é linear em (x, y) mas é linear em (x^2, y^2) e tem como solução

$$x^2 = \frac{1}{4} \quad y^2 = \frac{1}{4}.$$

Deste modo $(1/2, 1/2)$, $(-1/2, 1/2)$, $(1/2, -1/2)$ e $(-1/2, -1/2)$ satisfazem o sistema de estacionaridade.

De entre as soluções de (3.1) as que são pontos interiores do domínio fornecem a lista de possíveis candidatos a extremos locais em pontos interiores: $(0, 0)$, $(1/2, 1/2)$, $(-1/2, 1/2)$, $(1/2, -1/2)$ e $(-1/2, -1/2)$. Avaliando a função f nestes pontos obtemos $f(0, 0) = f(\pm 1, 0) = f(0, \pm 1) = 0$, $f(\pm 1/2, \pm 1/2) = 1/8$ e $f(\pm 1/2, \mp 1/2) = -1/8$.

Para avaliar o que se passa sobre a fronteira do domínio consideramos $f(-1, y) = y^3$ para $y \in [-1, 1]$, $f(1, y) = -y^3$ para $y \in [-1, 1]$, $f(x, 1) = -x^3$ para $x \in [-1, 1]$, $f(x, -1) = x^3$ para $x \in [-1, 1]$. Todas estas funções de uma variável real são estritamente monótonas de maneira que basta considerar os valores da função nos vértices do quadrado: $f(1, 1) = f(-1, -1) = -1$ e $f(-1, 1) = f(1, -1) = 1$.

Portanto $(1, 1)$ e $(-1, -1)$ são pontos de mínimo global e $(1, -1)$ e $(-1, 1)$ são pontos de máximo global.

Temos agora de estudar o que acontece nos outros pontos pois podem ser máximos ou mínimos locais ou apenas pontos de sela. Quanto ao ponto $(0, 0)$ é fácil de verificar que xy assume valores positivos e negativos numa vizinhança da origem. Por outro lado se (x, y) estiver suficientemente próximo de $(0, 0)$ a função $1 - x^2 - y^2$ é positiva. Portanto f numa vizinhança da origem assume valores positivos e negativos. Logo $(0, 0)$ é um ponto de sela.

Quanto ao ponto $(1/2, 1/2)$ classificamo-lo usando um raciocínio ad hoc baseado na utilização do teorema de Weierstrass. Note-se que $(1/2, 1/2)$ é um ponto interior do conjunto compacto $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1, x \geq 0, y \geq 0\}$, que f vale 0 sobre ∂A e $f > 0$ no interior de A . O teorema de Weierstrass garante que f terá um máximo em A (global relativamente a A) que ocorrerá necessariamente num ponto interior. Tal ponto é então um ponto de estacionaridade. O

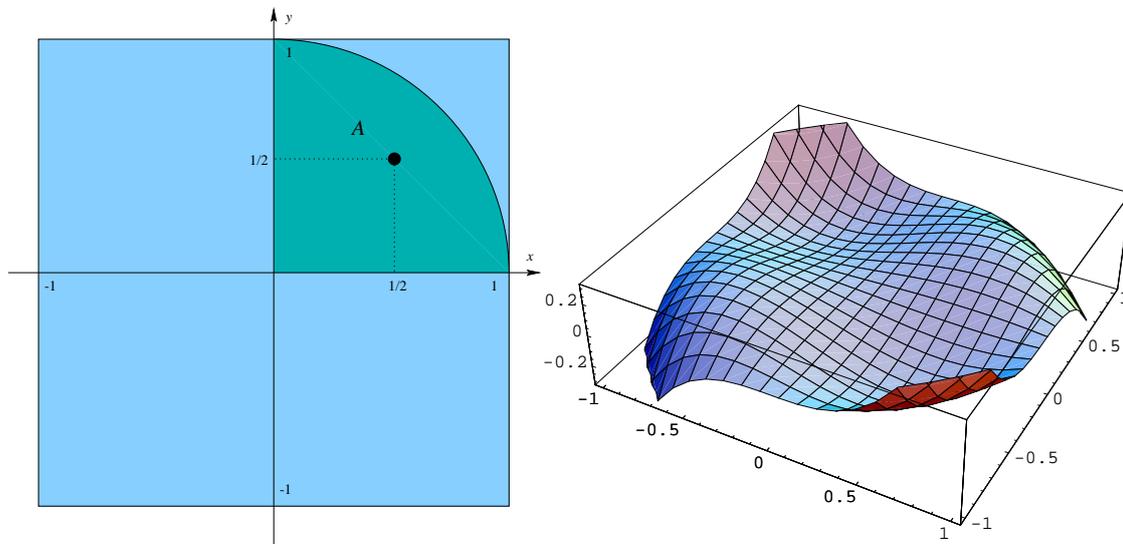


Figura 3.4: Estudo de $f(x, y) = xy(1 - x^2 - y^2)$ quanto a existência de pontos de extremo em $[-1, 1] \times [-1, 1]$. Tente identificar as propriedades deduzidas para a função com o que é evidenciado no gráfico gerado numericamente à direita.

único ponto de estacionaridade em $\text{int } A$ é $(1/2, 1/2)$ logo este ponto é um ponto de máximo local de f (relativamente ao quadrado $[-1, 1] \times [-1, 1]$). Este raciocínio vale para $(1/2, -1/2)$, $(-1/2, 1/2)$ e $(-1/2, -1/2)$ chegando-se de maneira análoga à conclusão que $(1/2, -1/2)$, $(-1/2, 1/2)$ são pontos de mínimo local e $(-1/2, -1/2)$ um ponto de máximo local (ou use o facto de a função ser ímpar em cada uma das variáveis).

Exercício 3.1.7 Determine, se existirem, os pontos de máximo e mínimo local da função $(x, y) \mapsto xy$ no quadrado $\text{máx}\{|x|, |y|\} \leq 1$.

Para terminar esta secção vamos apresentar um exemplo em que usamos propriedades de simetria e uma mudança de variável para determinar extremos

Exemplo 3.1.12 Seja $f(x, y, z, w) = x^2 + y^2 - z^2 - w^2 + (x^2 + y^2)^2$. Definindo $r_1^2 = x^2 + y^2$ e $r_2^2 = z^2 + w^2$ temos $f(x, y, z, w) = r_1^2 - r_2^2 + r_1^4$. Portanto, determinando os máximos e mínimos de $g(r_1, r_2) = r_1^2 - r_2^2 + r_1^4$, podemos recuperar os máximos e mínimos de f .

Exercício 3.1.8 Determine os extremos de $g(r_1, r_2) = r_1^2 - r_2^2 + r_1^4$. Utilize este resultado para calcular os extremos de $f(x, y, z, w) = x^2 + y^2 - z^2 - w^2 + (x^2 + y^2)^2$.

3.1.1 Exercícios suplementares

Exercício 3.1.9 Determine os pontos de extremo de:

1. $f(x, y, z) = x^2 + y^2 + z^2$ com $|x| + |y| + |z| \leq 1$.
2. $f(x, y) = x + y$ com $x^2 + y^2 \leq 1$.
3. $f(x, y) = x^2 + y^2 - (x^2 + y^2)^2$.
4. $f(x, y) = x^3 y^3 (1 - x^6 - y^6)$ para $(x, y) \in [-1, 1] \times [-1, 1]$.

Exercício 3.1.10 Seja $f : \mathbb{R} \rightarrow \mathbb{R}$, contínua, satisfazendo

$$\lim_{x \rightarrow \pm\infty} f(x) = +\infty.$$

Prove que f tem pelo menos um mínimo.

Exercício 3.1.11 (Mínimos quadrados) O método dos mínimos quadrados tem como objectivo determinar a recta $y = ax + b$ que “melhor aproxima” certos dados experimentais (x_i, y_i) , com $1 \leq i \leq n$. Uma função que permite medir quanto é que uma dada recta na forma $y = ax + b$ aproxima os pontos experimentais é

$$g(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2.$$

Calcule os pontos de estacionariade de g para determinar que equações é que a e b satisfazem (a prova de que o ponto de estacionaridade é mesmo um mínimo é deixada para um exercício posterior).

3.1.2 Sugestões para os exercícios

3.1.9

1. Note que f é o quadrado da distância à origem.
2. Como f não tem pontos de estacionaridade em $x^2 + y^2 < 1$ os seus extremos (que existem pelo teorema de Weirstrass) têm de se encontrar na fronteira. Escreva os pontos da fronteira com $x = \cos(\theta)$ e $y = \sin(\theta)$. Determine os extremos de $\cos(\theta) + \sin(\theta)$ com $\theta \in [0, 2\pi]$.
3. Determine os extremos de $r^2 - r^4$ com $r \geq 0$. Faça $r^2 = x^2 + y^2$.
4. Recorde o exemplo 3.1.11 substituindo $x \leftrightarrow x^3$ e $y \leftrightarrow y^3$.

3.1.10 Utilize o teorema do valor médio. ◀

3.1.11 Se g tiver mínimo em (a, b) verifica-se $\nabla g = 0$. Portanto a e b satisfazem as equações ▶

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}.$$

3.2 Testes de Segunda Ordem

Nesta secção vamos estudar um método que permite classificar os pontos de estacionaridade de funções. No caso unidimensional, quando a segunda derivada não se anula, um ponto de estacionaridade de uma função é de máximo ou de mínimo dependendo do sinal da segunda derivada. Para funções f de \mathbb{R}^n em \mathbb{R} a segunda derivada de f é representada por uma forma bilinear definida por uma matriz chamada hessiana. Classificando a forma quadrática definida pela hessiana quanto a ser definida positiva, negativa, indefinida, semidefinida, . . . , ou de forma equivalente determinando o sinal dos seus valores próprios, é possível estudar a classificação de pontos de estacionaridade quanto a serem pontos de máximo ou mínimo. À semelhança do caso unidimensional quando a derivada é nula, este teste pode não ser conclusivo se a forma quadrática for semidefinida, isto é todos os valores próprios tiverem o mesmo sinal excepto alguns nulos.

Comecemos por precisar alguns dos termos usados no parágrafo anterior.

Definição 3.2.1 Seja A uma matriz simétrica, ou seja $A = A^T$ e considere-se a forma quadrática Q_A definida por A via $Q_A(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x}$ para $\mathbf{x} \in \mathbb{R}^n$.

1. Diz-se que A é definida positiva (resp. negativa) se a forma quadrática Q_A for definida positiva (resp. negativa), isto é, $Q_A(\mathbf{x}) > 0$ (resp. $Q_A(\mathbf{x}) < 0$) para todo o $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$.
2. Diz-se que A é semi-definida positiva⁴ (resp. negativa) se a forma quadrática Q_A for semi-definida positiva (resp. negativa), isto é, $Q_A(\mathbf{x}) \geq 0$ (resp. $Q_A(\mathbf{x}) \leq 0$) para todo o $\mathbf{x} \in \mathbb{R}^n$ e existe algum $\mathbf{y} \neq 0$ tal que $Q_A(\mathbf{y}) = 0$.
3. Caso nenhuma destas situações se verifique diz-se que a matriz é indefinida esta situação corresponde a Q_A ser indefinida, isto é, existirem $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ tais que $Q_A(\mathbf{y}) < 0$ e $Q_A(\mathbf{z}) > 0$.

A definição anterior poderia ter sido feita em termos de valores próprios (consultar por exemplo [4] ou resolver o exercício 3.2.2) graças ao seguinte resultado básico de Álgebra Linear.

Proposição 3.2.1

Seja Q_A uma forma quadrática definida por uma matriz simétrica A via $Q_A(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x}$ para $\mathbf{x} \in \mathbb{R}^n$. Então:

1. Q_A é definida positiva (resp. negativa) se e só se todos os valores próprios de A forem positivos (resp. negativos).
2. Q_A é semi-definida positiva (resp. negativa) se e só se todos os valores próprios de A forem não negativos (resp. positivos) e pelo menos um nulo.
3. Q_A é indefinida se existir um valor próprio positivo e um valor próprio negativo.

Exemplo 3.2.1 Seja

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Os valores próprios de A são definidos pela equação

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 2 & 0 \\ 2 & 4 - \lambda & 0 \\ 0 & 0 & 1 - \lambda \end{vmatrix} = \lambda(1 - \lambda)(5 - \lambda) = 0,$$

que tem como soluções $\lambda = 0, 1, 5$. Portanto concluímos que A é semi-definida positiva.

Exercício 3.2.1 Mostre que a única matriz simultaneamente semidefinida positiva e semidefinida negativa é a matriz nula.

Que basta considerar matrizes simétricas ao lidar com formas quadráticas é uma das conclusões do exercício seguinte.

Exercício 3.2.2 Em geral podemos definir forma quadrática Q_A associada a uma matriz A via $Q_A(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x}$.

1. Mostre que $Q_A = Q_{A^\sharp}$, onde $A^\sharp = \frac{A+A^T}{2}$ em que A^\sharp é chamada a simetrização de A . Portanto substituir A pela sua simetrização não altera Q_A . Sugere-se que antes de provar o caso geral, convença-se que este facto é verdadeiro com o exemplo

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

⁴Esta definição de forma semidefinida não é a mesma de, por exemplo, [4] aonde uma forma ou matriz definida é necessariamente semidefinida. Assim definida, indefinida e semidefinida são termos mutuamente exclusivos.

2. Demonstre a proposição 3.2.1.

Calcular valores próprios não é uma tarefa trivial e é conveniente dispor de critérios mais fáceis de aplicar.

Proposição 3.2.2

Seja

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

uma matriz $n \times n$. Consideremos as submatrizes A_k que consistem nos elementos das primeiras k linhas e k colunas de A , isto é,

$$A_1 = [a_{11}] \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \cdots$$

Então,

1. A é definida positiva se e só se $\det A_i > 0$ para todo o i .
2. A é definida negativa se e só se $\det A_i < 0$ para i ímpar e $\det A_i > 0$ para i par.

Exemplo 3.2.2 *Seja*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 4 \end{bmatrix}.$$

Portanto

$$A_1 = [1] \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad A_3 = A$$

e temos

$$\det A_1 = 1 \quad \det A_2 = 2 \quad \det A_3 = 6.$$

Como todos estes valores são positivos concluímos que A é definida positiva.

Exercício 3.2.3 *Prove a proposição para matrizes diagonais.*

Para o caso de matrizes semi-definidas o critério é ligeiramente mais complexo. Dada uma matriz A uma submatriz principal de A é qualquer matriz que se obtém de A suprimindo linhas e colunas em pares correspondentes (e.g. a primeira e a terceira linhas e colunas).

Exemplo 3.2.3 *Seja*

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{bmatrix}$$

Suprimindo a primeira linha e primeira coluna obtemos a submatriz principal

$$\begin{bmatrix} 7 & 8 & 9 & 10 \\ 12 & 13 & 14 & 15 \\ 17 & 18 & 19 & 20 \\ 22 & 23 & 24 & 25 \end{bmatrix}$$

Suprimindo a segunda e terceira linhas e colunas obtemos a submatriz principal

$$\begin{bmatrix} 1 & 4 & 5 \\ 16 & 19 & 20 \\ 21 & 24 & 25 \end{bmatrix}.$$

Proposição 3.2.3

Uma matriz A é semi-definida positiva se e só se todas as submatrizes principais de A têm determinantes não negativos e pelo menos um é nulo. Uma matriz A é semi-definida negativa se e só se todas as submatrizes principais de A têm determinantes não negativos ou não positivos conforme o número de linhas ou colunas da submatriz é par ou ímpar e pelo menos um é nulo.

Exemplo 3.2.4 *Seja*

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & -5 \end{bmatrix}.$$

Retirando a primeira e terceira linhas e colunas obtemos a submatriz $[2]$ cujo determinante é positivo. Retirando a primeira e segunda linhas e colunas obtemos a submatriz $[-5]$ cujo determinante é negativo. Portanto concluímos que a matriz não pode ser nem semidefinida positiva nem semidefinida negativa pelo que é indefinida.

Exemplo 3.2.5 *Seja*

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 5 \end{bmatrix}.$$

O determinante de A é zero pelo que a matriz não pode ser nem definida positiva nem definida negativa. O mesmo acontece ao determinante de qualquer submatriz obtida de A não retirando a primeira linha e coluna. Portanto basta analisar 3 submatrizes; retirando a primeira e segunda linhas e colunas obtemos a submatriz $[5]$ cujo determinante é positivo; retirando a primeira e terceira linhas e colunas obtemos a submatriz $[2]$ cujo determinante é positivo; retirando a primeira linha e coluna obtemos a submatriz

$$\begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$$

cujo determinante é 9 e portanto também positivo. Portanto concluímos que a matriz é semidefinida positiva.

Exercício 3.2.4 *Classifique a matriz A dada por*

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 5 \end{bmatrix}$$

quanto a ser definida ou semidefinida positiva, negativa ou indefinida

Exercício 3.2.5 *Classifique a matriz A dada por*

$$A = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 5 \end{bmatrix}$$

quanto a ser definida ou semidefinida positiva, negativa ou indefinida

Depois destas definições preliminares vamos definir a matriz hessiana⁵.

Definição 3.2.2 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 . A matriz hessiana de f , $H(f)$, é dada por*

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

⁵A matriz hessiana H define uma forma bilinear $(x, y) \mapsto x \cdot Hy$ que desempenha o papel de segunda derivada de uma função de \mathbb{R}^n em \mathbb{R} . Não desenvolveremos este assunto neste texto.

Exemplo 3.2.6 Seja $f(x, y) = x^2 + y^2$. A sua matriz hessiana é

$$H(f) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Exercício 3.2.6 Calcule a matriz hessiana de $f(x, y, z) = xyz$.

Exercício 3.2.7 1. Defina uma função cuja matriz hessiana seja, em qualquer ponto

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

2. Será que a função que encontrou na alínea anterior é única? Se não for tente encontrar uma fórmula geral para esta família de funções.

3. Em que condições é que a matriz

$$\begin{bmatrix} a & b \\ d & c \end{bmatrix}$$

é a hessiana de alguma função de classe C^2 ?

O resultado básico para classificar pontos de estacionaridade usando o termo de segunda ordem da fórmula de Taylor é

Teorema 3.2.4

Sejam $U \subset \mathbb{R}^n$ um aberto, $f : U \rightarrow \mathbb{R}$ uma função de classe $C^2(U)$ e $\mathbf{x}_0 \in U$ um ponto de estacionaridade de f .

- i) Se $D_{\mathbf{h}}^{(2)} f(\mathbf{x}_0) > 0$ para todo o $\mathbf{h} \neq 0$ então \mathbf{x}_0 é um ponto de mínimo local;
- ii) Se $D_{\mathbf{h}}^{(2)} f(\mathbf{x}_0) \geq 0$ para todo o vector \mathbf{h} e existe um vector $\mathbf{k} \neq 0$ tal que $D_{\mathbf{k}}^{(2)} f(\mathbf{x}_0) = 0$ então \mathbf{x}_0 não é um ponto de máximo local;
- iii) Se $D_{\mathbf{h}}^{(2)} f(\mathbf{x}_0) < 0$ para todo o $\mathbf{h} \neq 0$ então \mathbf{x}_0 é um ponto de máximo local;
- iv) Se $D_{\mathbf{h}}^{(2)} f(\mathbf{x}_0) \leq 0$ para todo o vector \mathbf{h} e existe um vector $\mathbf{k} \neq 0$ tal que $D_{\mathbf{k}}^{(2)} f(\mathbf{x}_0) = 0$ então \mathbf{x}_0 não é um ponto de mínimo local;
- v) Se existem $\mathbf{h}, \mathbf{k} \in \mathbb{R}^n$ tais que $D_{\mathbf{h}}^{(2)} f(\mathbf{x}_0) < 0$ e $D_{\mathbf{k}}^{(2)} f(\mathbf{x}_0) > 0$ então \mathbf{x}_0 é um ponto de sela.

Ideia da demonstração. Para provar (ii), (iv) e (v) basta considerar as restrições de f às rectas passando por \mathbf{x}_0 e nas direcções de \mathbf{h} ou \mathbf{k} e usar os resultados conhecidos⁶ para dimensão 1. Para provar (i) ou (iii) devemos estudar o sinal de $f(\mathbf{x}) - f(\mathbf{x}_0)$ provando que se mantém constante numa bola de raio suficientemente pequeno centrada em \mathbf{x}_0 . Isto é equivalente a estudar o sinal de

$$\frac{f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)}{|\mathbf{h}|^2} = \frac{1}{2} D_{\mathbf{h}/|\mathbf{h}|}^{(2)} f(\mathbf{x}_0) + \frac{E_f(\mathbf{x}_0, \mathbf{h})}{|\mathbf{h}|^2}$$

em que a última parcela do segundo membro tende para 0 quando $\mathbf{h} \rightarrow 0$ de acordo com o teorema de Taylor. Para completar a demonstração, por exemplo no caso (i), basta mostrar que para $\mathbf{h} \neq 0$ temos $D_{\mathbf{h}/|\mathbf{h}|}^{(2)} f(\mathbf{x}_0)$ minorado por um número $m > 0$ e que existe uma bola centrada em \mathbf{x}_0 tal que aí $\frac{E_f(\mathbf{x}_0, \mathbf{h})}{|\mathbf{h}|^2} > -m$. O último destes dois factos segue da definição de limite e o primeiro pode ser justificado usando resultados de álgebra linear sobre formas quadráticas ou o teorema de Weierstrass aplicado à função⁷ $S^{n-1} \ni \eta \mapsto D_{\eta}^{(2)} f(\mathbf{x}_0)$.

⁶Obviamente pode refazer-se a demonstração mas queremos acentuar que não existe nenhuma ideia essencialmente nova em jogo.

⁷ $S^{n-1} \equiv \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| = 1\}$.

O teorema anterior pode ser enunciado usando a terminologia de álgebra linear referente a formas quadráticas. Com efeito $D_{\mathbf{h}}^{(2)}f(\mathbf{x}_0)$ é a forma quadrática definida pela *matriz hessiana* de f no ponto \mathbf{x}_0 , $H_f(\mathbf{x}_0) \equiv \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) \right]_{i,j=1,\dots,n}$, isto é, $D_{\mathbf{h}}^{(2)}f(\mathbf{x}_0) = \mathbf{h} \cdot H_f(\mathbf{x}_0)\mathbf{h}$. As situações (i-v) no enunciado do teorema correspondem respectivamente a esta forma quadrática⁸ ser *definida positiva*, *semidefinida positiva não nula*, *definida negativa*, *semidefinida negativa não nula* e *indefinida*.

Corolário 3.2.5

Seja $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^2 numa vizinhança um ponto de estacionaridade em \mathbf{x}_0 . Então:

1. Se $H(f)(\mathbf{x}_0) = 0$ o teste é inconclusivo.
2. Se $H(f)(\mathbf{x}_0)$ for definida positiva (resp. negativa) então \mathbf{x}_0 é um ponto de mínimo (resp. máximo) local.
3. Se $H(f)(\mathbf{x}_0)$ for semi-definida positiva (resp. negativa) mas não nula então \mathbf{x}_0 não é um ponto de máximo (resp. mínimo) local, isto é, pode ser ponto de mínimo (resp. máximo) local ou ponto de sela.
4. Se $H(f)(\mathbf{x}_0)$ for indefinida então \mathbf{x}_0 é um ponto de sela.

O teorema e o corolário não podem ser melhorados, através de informação só relativa a derivadas de segunda ordem e de maneira a fornecer informação adicional para os casos em que a forma quadrática é semidefinida, devido aos exemplos triviais que se seguem (3.2.8, 3.2.9).

Exemplo 3.2.7 Seja $f(x, y) = x^2 + y^2$. O ponto $(0, 0)$ é um ponto de estacionaridade (verifique!). A matriz hessiana de f no ponto $(0, 0)$ é

$$H(f) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

que é definida positiva (os valores próprios são positivos). Portanto $(0, 0)$ é um ponto de mínimo local.

Exemplo 3.2.8 Seja $f(x, y) = x^2 + y^4$. O ponto $(0, 0)$ é um ponto de estacionaridade (verifique!). A matriz hessiana de f no ponto $(0, 0)$ é

$$H(f) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$

que é semi-definida positiva (os valores próprios são não negativos). Portanto $(0, 0)$ não é um ponto de máximo local. É fácil verificar que $(0, 0)$ é um ponto de mínimo local e não um ponto de sela. Com efeito, basta observar que, se $(x, y) \neq (0, 0)$, se tem $f(x, y) > f(0, 0) = 0$.

Exemplo 3.2.9 Seja $f(x, y) = x^2 - y^4$. O ponto $(0, 0)$ é um ponto de estacionaridade (verifique!). A matriz hessiana de f no ponto $(0, 0)$ é

$$H(f) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$

que é semi-definida positiva (os valores próprios são não negativos). Portanto $(0, 0)$ não é um ponto de máximo local. No entanto $(0, 0)$ não é um ponto de mínimo local; com efeito, temos $f(0, 0) = 0$ mas $f(0, y) = -y^4 < 0$ para $y \neq 0$ pelo que concluímos que $(0, 0)$ é um ponto de sela.

⁸Esta terminologia relativa a formas quadráticas usa-se também para as matrizes que as definem.

Exemplo 3.2.10 Seja $f(x, y) = x^2 - y^2$. O ponto $(0, 0)$ é um ponto de estacionaridade (verifique!). A matriz hessiana de f no ponto $(0, 0)$ é

$$H(f) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$

que é indefinida (um dos valores próprios é positivo e outro é negativo). Portanto $(0, 0)$ é um ponto de sela.

Exercício 3.2.8 Prove que $(0, 0)$ é um ponto de estacionaridade de f e classifique-o quanto a ser ponto de máximo, ponto de mínimo ou ponto de sela quando f é definida em \mathbb{R}^2 por:

1. $f(x, y) = 2x^2 + y^2$;
2. $f(x, y) = xy$;
3. $f(x, y) = x^2 + 2xy + y^2$;
4. $f(x, y) = y^4 - x^4$;
5. $f(x, y) = x^3$;
6. $f(x, y) = y^2$.

Problema 3.2.1 Elabore um critério para classificar formas quadráticas definidas por uma matriz 2×2 da forma $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ em função do sinal de $d = ac - b^2$ e do sinal de a .

Os exemplos de aplicação do critério de segunda ordem até agora apresentados são no essencial triviais e poderiam ser analisados por outros processos. Destinavam-se a definir situações típicas e balizar as limitações do resultado. O exemplo seguinte já tem um carácter menos trivial.

Exemplo 3.2.11 Considere-se a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = xy + x^2y^3 - x^3y^2$. Tentemos estudá-la quanto à existência de extremos.

Começamos por notar que graças a f ser um polinómio reconhecemos imediatamente que f coincide com o seu desenvolvimento de Taylor de ordem igual ou superior ao seu grau. Tal é verdadeiro em particular relativamente a $(0, 0)$ que reconhecemos como um ponto de estacionaridade (ausência de termos de primeira ordem) que é um ponto de sela (termo de segunda ordem xy).

Para determinar outros pontos de estacionaridade consideramos o sistema de estacionaridade

$$\begin{cases} \frac{\partial f}{\partial x} \equiv y + 2xy^3 - 3x^2y^2 = 0 \\ \frac{\partial f}{\partial y} \equiv x + 3x^2y^2 - 2x^3y = 0 \end{cases}$$

que pode ser escrito de forma equivalente como

$$\begin{cases} y(1 + 2xy^2 - 3x^2y) = 0 \\ x(1 + 3xy^2 - 2x^2y) = 0. \end{cases}$$

Daí decorre que a única solução sobre os eixos coordenados é $(0, 0)$ que já foi estudada. Podemos então limitarmo-nos a analisar

$$\begin{cases} 1 + 2xy^2 - 3x^2y = 0 \\ 1 + 3xy^2 - 2x^2y = 0. \end{cases}$$

Subtraindo termo a termo obtemos $xy^2 + x^2y = 0$ ou seja $xy(y + x) = 0$. Assim eventuais soluções adicionais do sistema de estacionaridade encontrar-se-iam ou sobre os eixos coordenados (hipótese já estudada) ou sobre a recta $y = -x$. Substituindo y por $-x$ na primeira equação obtemos

$1 + 5x^3 = 0$ o que fornece um segundo e último ponto de estacionaridade: $(-5^{-1/3}, 5^{-1/3})$. Para classificá-lo calculamos

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= 2y^3 - 6xy^2 \\ \frac{\partial^2 f}{\partial y^2} &= 6x^2y - 2x^3 \\ \frac{\partial^2 f}{\partial x \partial y} &= 1 + 6xy^2 - 6x^2y\end{aligned}$$

pelo que

$$H_f(-5^{-1/3}, 5^{-1/3}) = \begin{bmatrix} 8/5 & 7/5 \\ 7/5 & 8/5 \end{bmatrix}$$

uma matriz definida positiva pelo que este ponto de estacionaridade é um ponto de mínimo local sendo o mínimo local $f(-5^{-1/3}, 5^{-1/3}) = -\frac{3}{5}5^{-2/3}$.

Considerando, por exemplo, $\lim_{\lambda \rightarrow +\infty} f(1, \lambda) = +\infty$, $\lim_{\lambda \rightarrow +\infty} f(\lambda, 1) = -\infty$ verifica-se que esta função não tem extremos absolutos.

O teorema 3.2.4 é passível de várias generalizações. Aconselha-se no entanto o aluno a começar por dominar o critério de segunda ordem e as ideias na sua demonstração pois são a base de qualquer uma dessas generalizações. Mais geralmente um polinómio homogéneo de grau k designa-se por *forma de grau k* . Uma generalização imediata do resultado anterior é

Problema 3.2.2 *Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe $C^k(D)$ e \mathbf{x}_0 um ponto interior a D tal que $D_{\mathbf{h}}^{(j)} f(\mathbf{x}_0) = 0$ para $j < k$ e $\mathbf{h} \in \mathbb{R}^n$ e a forma de grau k Q definida por $Q(\mathbf{h}) = D_{\mathbf{h}}^{(k)} f(\mathbf{x}_0)$ é definida positiva. Prove que \mathbf{x}_0 é um ponto de mínimo local de f . Formule e demonstre outras generalizações do mesmo tipo do teorema 3.2.4.*

Generalizações deste tipo poderão ser encontradas por exemplo em [2] (ver também o exercício 3.2.12 e o problema 3.2.4). Factos triviais mas muito úteis são

Problema 3.2.3

- Seja Q uma forma não nula de grau ímpar. Prove que Q é uma forma indefinida.
- Seja P um polinómio de grau ímpar. Prove que P não é limitado superior ou inferiormente.

Exemplo 3.2.12 *Considere-se a função $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por*

$$g(x, y) = e^{x^2 - y^2} + y^2.$$

e tentemos classificar o ponto de estacionaridade $(0, 0)$.

De maneira análoga ao exemplo 2.4.4 obtemos a partir da série de Taylor da exponencial

$$g(x, y) = 1 + x^2 + \sum_{j=2}^{\infty} \frac{(x^2 - y^2)^j}{j!}$$

para todo o $(x, y) \in \mathbb{R}^2$. Note-se que a análise através do termo de segunda ordem da fórmula de Taylor só nos permite afirmar que $(0, 0)$ não é um ponto de máximo devido à forma quadrática se anular na direcção do eixo dos y 's. Podemos tentar compreender o que se passa usando os termos de ordem superior da fórmula de Taylor naquela direcção. O primeiro desses termos que não se anula é de ordem 4, mais precisamente,

$$g(x, y) = 1 + x^2 + \frac{(x^2 - y^2)^2}{2} + E(x, y)$$

em que $\frac{E(x,y)}{(x^2+y^2)^2} \rightarrow 0$ quando $(x,y) \rightarrow 0$. É de suspeitar que $(0,0)$ é um ponto de mínimo e tentaremos prová-lo usando o mesmo raciocínio da demonstração do teorema 3.2.4 em que a minimização do termo de segunda ordem por um número positivo é substituída pela minimização simultânea dos termos de segunda e quarta ordem. A ideia natural é usar o termo de quarta ordem para direcções “próximas” da do eixo dos y 's e o termo de segunda ordem para as restantes. Como o termo de quarta ordem se anula para $|x| = |y|$ e o de segunda ordem para $x = 0$ tentamos caracterizar tais direcções respectivamente por $|x| < \frac{1}{2}|y|$ e $|x| \geq \frac{1}{2}|y|$.

Seja então $|x| < \frac{1}{2}|y|$. Obtemos

$$\begin{aligned} & \frac{1}{(x^2 + y^2)^2} \left(\frac{1}{2} D_{(x,y)}^{(2)} g(0,0) + \frac{1}{4!} D_{(x,y)}^{(4)} g(0,0) \right) \\ &= \frac{1}{(x^2 + y^2)^2} \left(x^2 + \frac{(x^2 - y^2)^2}{2} \right) > \frac{1}{(\frac{5}{4}y^2)^2} \frac{(x^4 - 2x^2y^2 + y^4)}{2} > \frac{8}{25} \frac{x^4 - \frac{1}{2}y^4 + y^4}{y^4} > \frac{4}{25}. \end{aligned}$$

Por outro lado para $|x| \geq \frac{1}{2}|y|$ obtém-se

$$\begin{aligned} & \frac{1}{(x^2 + y^2)^2} \left(\frac{1}{2} D_{(x,y)}^{(2)} g(0,0) + \frac{1}{4!} D_{(x,y)}^{(4)} g(0,0) \right) \\ &= \frac{1}{(x^2 + y^2)^2} \left(x^2 + \frac{(x^2 - y^2)^2}{2} \right) > \frac{x^2}{(x^2 + y^2)^2} \geq \frac{x^2 + y^2}{4(x^2 + y^2)^2} = \frac{1}{4(x^2 + y^2)}. \end{aligned}$$

Agora já é possível aplicar um raciocínio idêntico ao do teorema 3.2.4 para concluir que $(0,0)$ é efectivamente um ponto de mínimo.

O leitor poderá ter considerado a resolução do exercício 3.2.12 algo *ad hoc* e suspeitado que existe um resultado abstracto que poderia ter sido usado. De facto assim é embora a maior parte das ideias relevantes já conste da resolução do exercício.

Problema 3.2.4 *Sejam $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^k(D)$, \mathbf{x}_0 um ponto interior a D . Suponha-se que existe $l < k$ tal que $D_{\mathbf{h}}^{(j)} f(\mathbf{x}_0) = 0$ para todo o $j < l$ e todo o $\mathbf{h} \in \mathbb{R}^n$, e que $\mathbf{h} \mapsto Q_l(\mathbf{h}) \equiv D_{\mathbf{h}}^{(l)} f(\mathbf{x}_0)$ é semidefinida positiva. Designamos os vectores unitários que anulam Q_l como direcções singulares. Suponha-se ainda que $D_{\eta}^{(j)} f(\mathbf{x}_0) = 0$ para toda a direcção singular η e $l < j < k$ e que $Q_k(\eta) \equiv D_{\eta}^{(k)} f(\mathbf{x}_0) > 0$ para toda a direcção singular η . Mostre que:*

- O conjunto formado por todas as direcções singulares é um subconjunto fechado de S^{n-1} que designamos por F .
- Q_k tem um mínimo $m_1 > 0$ sobre F e um mínimo m_2 sobre S^{n-1} .
- Existe um aberto $A \supset F$ tal que $Q_k(\eta) > \frac{m_1}{2}$ para todo o $\eta \in S^{n-1} \cap A$.
- Q_l tem um mínimo $m_3 > 0$ sobre $S^{n-1} \setminus A$.
- Valem as estimativas

$$\frac{f(\mathbf{x}) - f(\mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^k} \geq \begin{cases} \frac{m_3}{l!|\mathbf{x} - \mathbf{x}_0|^{k-l}} + \frac{m_2}{k!} + \frac{E_f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^k} & \text{se } \frac{\mathbf{x} - \mathbf{x}_0}{|\mathbf{x} - \mathbf{x}_0|} \notin A, \\ \frac{m_1}{2k!} + \frac{E_f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^k} & \text{se } \frac{\mathbf{x} - \mathbf{x}_0}{|\mathbf{x} - \mathbf{x}_0|} \in A, \end{cases}$$

em que $\frac{E_f(\mathbf{x}_0, \mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^k} \rightarrow 0$ quando $\mathbf{x} \rightarrow \mathbf{x}_0$.

- \mathbf{x}_0 é um ponto de mínimo local de f .

Para terminar convém referir mais uma vez que os testes baseados na fórmula de Taylor podem ser inconclusivos devido às razões apontadas na introdução a este capítulo e aí exemplificadas com funções reais de variável real.

3.2.1 Exercícios suplementares

Exercício 3.2.9 Considere a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por

$$f(x, y, z) = 2 - \left(z - \sqrt{x^2 + y^2} \right)^2 + \left(z - \sqrt{x^2 + y^2} \right)^3.$$

- a) Determine os respectivos pontos de extremo local e absoluto e, se tais pontos existirem, classifique-os quanto a serem pontos de máximo ou de mínimo.
- b) Determine um polinómio de grau menor ou igual a dois, $P(x, y, z)$, tal que

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

ou justifique que tal polinómio não existe.

Exercício 3.2.10 Considere a função $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por

$$g(x, y, z) = x^3(y^2 + z^2) \left(1 - x - \sqrt{y^2 + z^2} \right).$$

Estude g quanto à existência de extremos relativos e absolutos. Determine tais extremos se existirem e os pontos onde ocorrem. Sugestão: Considere primeiro $h(x, \rho) = x^3\rho^2(1 - x - \rho)$.

Exercício 3.2.11 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} \frac{xy^5}{x^2+y^4}, & \text{se } (x, y) \neq (0, 0) \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$$

- a) Determine justificadamente o maior subconjunto do domínio de f em que existem e são iguais as derivadas parciais $\frac{\partial^2 f}{\partial x \partial y}$ e $\frac{\partial^2 f}{\partial y \partial x}$.
- b) Determine e classifique os pontos de estacionaridade de f quanto a serem pontos de extremo ou pontos de sela.
- c) Determine o máximo e o mínimo da restrição de f ao conjunto $A = \{(x, y) \in \mathbb{R}^2 : x \geq y^2, y \leq -x^2\}$ e os pontos em que ocorrem esses extremos.

Exercício 3.2.12 Considere a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

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

Determine, se existirem, os pontos de estacionaridade de f e classifique-os quanto a serem pontos de extremo relativo ou pontos de extremo absoluto.

3.2.2 Sugestões para os exercícios

3.2.9

- a) A função é constante sobre cada uma das superfícies de equação $z - \sqrt{x^2 + y^2} = \alpha$, $\alpha \in \mathbb{R}$ pelo que basta estudar a função $\mathbb{R} \ni \alpha \mapsto 2 - \alpha^2 + \alpha^3$. Conclui-se facilmente que f tem um máximo para $z - \sqrt{x^2 + y^2} = 0$ e um mínimo para $z - \sqrt{x^2 + y^2} = 2/3$. Tais extremos não são absolutos.
- b) Tal polinómio existe e é obviamente o polinómio de Taylor de segunda ordem de f relativo ao ponto $(1, 1, 0)$.

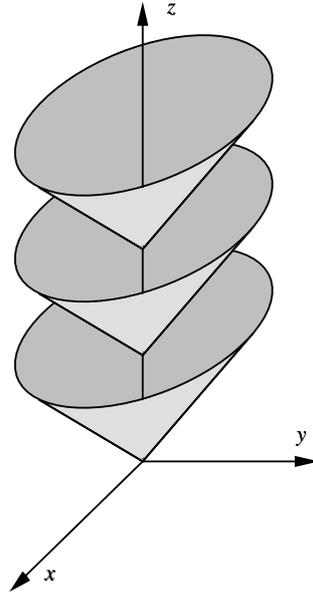


Figura 3.5: Esta figura acompanha a sugestão de solução do exercício 3.2.9. A função f é constante sobre cada uma das folhas de cone $z - \sqrt{x^2 + y^2} = \alpha$.

3.2.10 Já vimos no exercício 3.2.9 as vantagens em, quando possível, usar simetrias da função a estudar para estudar um problema equivalente em dimensão inferior. Naquele caso acabámos estudando um problema unidimensional. No caso presente podemos estudar, usando a sugestão, um problema bidimensional do qual recuperaremos o problema original por rotação em torno do eixo dos x 's.

Vamos então estudar quanto à existência de extremos a função $g : \{(x, \rho) \in \mathbb{R}^2 : \rho \geq 0\} \rightarrow \mathbb{R}$ definida por $g(x, \rho) = x^3 \rho^2 (1 - x - \rho)$. Algo que convém fazer *antes* de iniciar qualquer tipo de cálculo é tentar identificar linhas de nível da função. A intersecção num ponto interior de tais linhas de nível formando um ângulo não nulo fornece-nos imediatamente a localização de um ponto de estacionaridade⁹. Tal é particularmente fácil para g pois esta função anula-se sobre o eixo dos x 's, sobre o eixo dos ρ 's e sobre a recta $1 - x - \rho = 0$. Isto identifica como ponto de estacionaridade $(x, \rho) = (1, 0)$ e se considerássemos a função estendida para $\rho < 0$ usando a mesma fórmula o mesmo se poderia dizer dos pontos $(0, 1)$ e $(0, 0)$. É fácil de verificar por análise do sinal de g que todos estes pontos são pontos de sela. Uma observação adicional que se obtém dessa análise é o facto de g ser positiva no interior do triângulo T limitado pelas rectas atrás referidas, isto é,

$$T = \{(x, \rho) \in \mathbb{R}^2 : x \geq 0, \rho \geq 0, x + \rho \leq 1\}.$$

Como T é limitado e fechado há-de existir no interior de T pelo menos mais um ponto de máximo de g que será portanto mais um ponto de estacionaridade de g . Eventualmente existirão outros pontos de estacionaridade. Todos estes factos servirão para verificar a resolução do sistema de estacionaridade de g

$$\begin{cases} \frac{\partial g}{\partial x} \equiv x^2 \rho^2 (3(1 - x - \rho) - x) = 0 \\ \frac{\partial g}{\partial \rho} \equiv x^3 \rho (2(1 - x - \rho) - \rho) = 0. \end{cases}$$

Verificamos imediatamente que todos os pontos sobre os eixos são pontos de estacionaridade. Todos os pontos sobre o eixo dos ρ 's são pontos de sela por análise do sinal de g . Sobre o eixo dos

⁹Enunciado e justificação rigorosa desta afirmação são algo que não pretendemos apresentar neste momento. Veja mais à frente o problema ??.

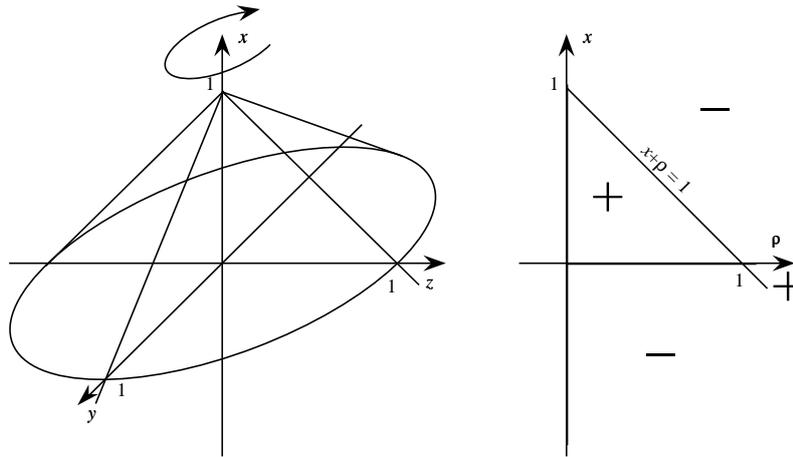


Figura 3.6: Esta figura acompanha o exercício 3.2.10. A função f exibe simetria radial relativamente ao eixo dos x 's. No gráfico da direita indicam-se os zeros e sinais de g .

x 's a situação é mais complexa: $(x, 0)$ é um ponto de mínimo se $0 < x < 1$, um ponto de máximo se $x < 0$ ou $1 < x$, e um ponto de sela se $x = 0$ ou $x = 1$. Pontos de estacionaridade que não se encontrem sobre os eixos deverão satisfazer

$$\begin{cases} 3(1 - x - \rho) - x = 0 \\ 2(1 - x - \rho) - \rho = 0. \end{cases}$$

Este sistema linear tem uma única solução: $(1/2, 1/3)$, a solução no interior de T cuja existência já tinha sido garantida e que sabemos tratar-se de um ponto de máximo.

É fácil verificar que g e consequentemente f não têm extremos absolutos.

Podemos concluir que f possui pontos de máximo local nos pontos da circunferência definida por $x = 1/2, y^2 + z^2 = 1/9$ onde f vale $1/432$, outros pontos de máximo local nos pontos $(x, 0, 0)$ com $x < 0$ ou $x > 1$ onde f vale 0 , e pontos de mínimo local nos pontos $(x, 0, 0)$ com $0 < x < 1$ onde f vale 0 . ◀

3.2.11

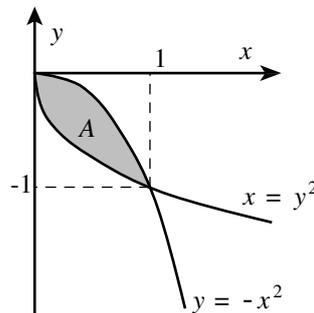


Figura 3.7: Esta figura acompanha os esboços de resolução dos Exercícios 3.2.11 e 3.2.12.

- a) No complementar da origem f é uma função de classe C^∞ pelo que aí verifica-se a igualdade $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. Resta-nos investigar o que se passa na origem. Como a função se anula sobre

os eixos coordenados decorre da definição de derivada parcial que

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0.$$

Além disso se $(x, y) \neq (0, 0)$ temos

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{y^5(x^2 + y^4) - 2x^2y^5}{(x^2 + y^4)^2} = \frac{y^9 - x^2y^5}{(x^2 + y^4)^2} \\ \frac{\partial f}{\partial y}(x, y) &= \frac{5xy^4(x^2 + y^4) - 4xy^8}{(x^2 + y^4)^2} = \frac{xy^8 + 5x^3y^4}{(x^2 + y^4)^2} \end{aligned}$$

donde decorre usando a definição de derivada parcial

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = 1 \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 0$$

pelo que o conjunto pretendido é $\mathbb{R}^2 \setminus \{(0, 0)\}$.

- b) Do cálculo das derivadas parciais de primeira ordem sabemos que $(0, 0)$ é um ponto de estacionaridade e outros pontos de estacionaridade serão soluções de

$$\begin{cases} y^9 - x^2y^5 & = 0 \\ xy^8 + 5x^3y^4 & = 0 \end{cases}$$

donde todos os pontos sobre o eixo dos x 's são pontos de estacionaridade. Outros pontos de estacionaridade deverão satisfazer

$$\begin{cases} y^4 - x^2 & = 0 \\ xy^4 + 5x^3 & = 0. \end{cases}$$

Da primeira equação deste sistema eventuais soluções adicionais devem satisfazer $y^4 = x^2$. Por substituição na segunda equação obtém-se unicamente a solução $(x, y) = (0, 0)$. Estabelecemos então que o conjunto dos pontos de estacionaridade é o eixo dos x 's. Por análise do sinal da função na sua vizinhança verificamos que todos são pontos de sela.

- c) Os extremos absolutos de f restringida a A ocorrem nalgum ponto de A pois trata-se de um conjunto limitado e fechado. Se ocorressem em pontos interiores tais pontos seriam pontos de extremo local o que da alínea anterior não acontece. Assim estudamos a restrição de f à fronteira de A (veja a figura 3.7). Definimos $g(y) = f(y^2, y) = y^3/2$ para $-1 \leq y \leq 0$. Temos $-1/2 = g(-1) < g(y) < g(0) = 0$ sempre que $-1 < y < 0$. Definimos $h(x) = f(x, -x^2) = -\frac{x^9}{1+x^6}$ para $0 \leq x \leq 1$. Como $h'(x) = -\frac{9x^8(1+x^6) - 6x^{14}}{(1+x^6)^2} = -\frac{3x^{14} + 9x^8}{(1+x^6)^2} < 0$ para $0 < x < 1$ temos $-1/2 = h(1) < h(x) < h(0) = 0$ para $0 < x < 1$. As funções g e h dão-nos os valores de f sobre a fronteira de A . Podemos concluir que $-1 = f(1, -1) < f(x, y) < f(0, 0) = 0$ para todo os $(x, y) \in A \setminus \{(0, 0), (1, -1)\}$.

◀

3.2.12 O conjunto de zeros de f está esboçado na figura 3.2.11. Observe que $(0, 0)$ e $(1, -1)$ são necessariamente pontos de sela e que existirá pelo menos um ponto de extremo local na região definida por $-\sqrt{x} \leq y \leq -x^2$. A solução do sistema de estacionaridade permite obter com efeito que os únicos pontos de estacionaridade são $(0, 0)$, $(1/2, -1/2)$ e $(1, -1)$. Por análise do sinal de f conclui-se que $(1/2, -1/2)$ é um ponto de mínimo local. A função não tem extremos absolutos. ◀

Capítulo 4

Teoremas da Função Inversa e da Função Implícita

Neste capítulo vamos estudar condições que permitem assegurar a existência da inversa de funções de $\mathbb{R}^n \rightarrow \mathbb{R}^n$, bem como condições que garantam a resolubilidade de equações da forma $f(x, y) = 0$ de modo a obtermos uma das variáveis em função da outra. Em casos simples conseguimos inverter as funções ou resolver as equações explicitamente; no entanto, na maioria dos casos, tal tarefa é complexa se não impossível. Os resultados gerais que obteremos (teoremas 4.3.1 e 4.4.1) asseguram a resolução destas questões num sentido *local* a precisar.

Exercício 4.0.13 *Convença-se da dificuldade de resolver problemas do tipo mencionado tentando inverter a função $f : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^2$ definida por*

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

Note que a análise deste problema pode ser feita de uma forma simples!

4.1 Invertibilidade de funções

Começemos por recordar a definição de função injectiva

Definição 4.1.1 *Diz-se que uma função $f : A \rightarrow B$, onde A e B são conjuntos arbitrários, é injectiva se, sempre que $x \neq y$ ($x, y \in A$), se tenha $f(x) \neq f(y)$.*

Observe que a definição anterior é equivalente a dizer que se $f(x) = f(y)$ então necessariamente se verifica $x = y$. É também equivalente a mostrar que a equação $f(x) = a$, para $a \in B$, tem, quando muito, uma solução.

Exercício 4.1.1 *Prove estas duas últimas afirmações.*

Consideremos agora o seguinte exemplo:

Exemplo 4.1.1 *Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}^+ \times \mathbb{R}^+$ a função definida por*

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

Provemos que ela é injectiva. Suponhamos que $f(x_1, y_1) = f(x_2, y_2)$. Então

$$e^{x_1} = e^{x_2} \quad e^{x_1+y_1} = e^{x_2+y_2}.$$

A primeira equação implica $x_1 = x_2$. Utilizando este resultado na segunda equação obtemos $y_1 = y_2$ pelo que f é injectiva.

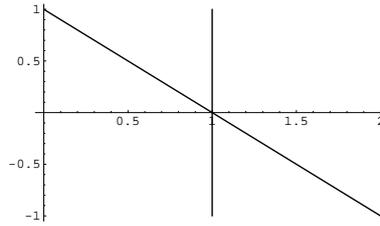


Figura 4.1: Rectas $x = 1$ e $x + y = 1$

Exercício 4.1.2 Prove que a função identidade de \mathbb{R}^n em \mathbb{R}^n , isto é, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $f(x) = x$, é injectiva.

Poderíamos ter resolvido o exemplo anterior utilizando o método gráfico que veremos de seguida:

Exemplo 4.1.2 Seja (a, b) com $a, b > 0$ um ponto no contradomínio de f . Queremos mostrar que o sistema

$$e^x = a \quad e \quad e^{x+y} = b$$

só tem uma solução. Graficamente, as soluções vão ser a intersecção das rectas da forma $x = \log a \equiv c$ e $x + y = \log b \equiv d$. Como se pode ver na figura 4.1 (para $c = d = 1$) estas rectas intersectam-se num único ponto uma vez que não são paralelas. Assim, como para cada par (a, b) existe no máximo uma pré-imagem,¹ concluímos que a função é injectiva.

Este exemplo sugere que é possível, utilizando apenas argumentos de natureza geométrica, verificar a injectividade de uma função. Sistematizemos este processo. Seja f uma função contínua, $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, com $f = (f_1, f_2)$. Suponhamos que queremos estudar a injectividade de f bem como caracterizar o seu contradomínio.

Consideremos C_a^1 , conjunto de nível de f_1 , definido por $f_1(x, y) = a$ e C_b^2 , definido por $f_2(x, y) = b$, conjunto de nível de f_2 , sendo a e b reais fixos. Podemos (em princípio), para cada par (a, b) , desenhar estes dois conjuntos; estudando o número de pontos de intersecção destas curvas para valores de a e b arbitrários podemos tirar conclusões importantes sobre a injectividade e contradomínio de f , tal como afirma a próxima proposição (observe a figura).

Proposição 4.1.1

Seja $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ($f = (f_1, f_2)$) uma função contínua. Defina-se

$$C_a^1 = \{(x, y) \in A : f_1(x, y) = a\} \quad e \quad C_b^2 = \{(x, y) \in A : f_2(x, y) = b\}.$$

Então:

1. o contradomínio de f é o conjunto de pontos $(a, b) \in \mathbb{R}^2$ tais que $C_a^1 \cap C_b^2 \neq \emptyset$;
2. a função é injectiva sse para qualquer par $(a, b) \in \mathbb{R}^2$, o conjunto $C_a^1 \cap C_b^2$ tiver no máximo um elemento.

Exercício 4.1.3 Demonstre a proposição anterior.

Exercício 4.1.4 Decida se a função $f(x, y) = (x + y, x^2 + y^2)$ é ou não injectiva.

Este método, sendo bastante geral para o caso de funções de $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, não é fácil de aplicar, pelo menos directamente, no caso mais geral de funções com mais de 2 variáveis, visto que o desenho de superfícies em \mathbb{R}^3 é bastante difícil e em \mathbb{R}^n , $n \geq 4$, praticamente impossível. No entanto, nalguns casos particulares ainda é possível utilizar ideias semelhantes, como podemos verificar no exemplo seguinte.

¹A pré-imagem de (a, b) é o conjunto de todos os pontos \mathbf{x} do domínio de f tais que $f(\mathbf{x}) = (a, b)$.

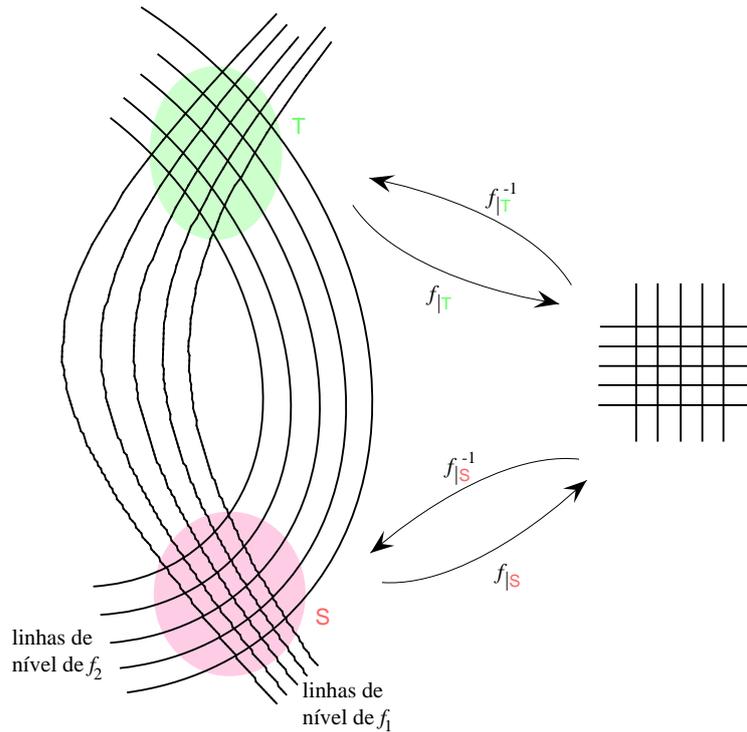


Figura 4.2: O método gráfico para analisar invertibilidade de aplicações de \mathbb{R}^2 em \mathbb{R}^2 e invertibilidade local versus invertibilidade global. As curvas de nível de f_1 e f_2 intersectam-se em dois pontos pelo que (f_1, f_2) não é injectiva. No entanto a restrição a S ou a T é injectiva.

Exemplo 4.1.3 Consideremos a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por

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

Mostremos que ela não é injectiva. Seja (a, b, c) um ponto no contradomínio de f . Podemos tomar, por exemplo, $b = c = 0$ e $a > 0$. A equação

$$f(x, y, z) = (a, 0, 0)$$

tem como solução os pontos que estão na intersecção da esfera centrada na origem definida por

$$x^2 + y^2 + z^2 = a$$

(note que esta equação define uma esfera pois $a > 0$) com a recta definida por

$$x + y + z = 0 \quad x - y = 0 \tag{4.1}$$

(a equação $x + y + z = 0$ define um plano que intersecta o plano $x - y = 0$ numa recta). Esta recta, que passa pela origem $((x, y, z) = (0, 0, 0))$ satisfaz o sistema 4.1), intersecta qualquer esfera centrada na origem em dois pontos distintos. Desta observação concluímos imediatamente que f não pode ser injectiva.

A complexidade de exemplos como os anteriores não ocorre para transformações lineares. Nesse caso a injectividade local garante invertibilidade global.

Exercício 4.1.5 Seja T uma transformação linear de \mathbb{R}^n em \mathbb{R}^n . Justifique que o contradomínio de T é \mathbb{R}^n sse T é injectiva numa vizinhança de 0 sse T é invertível.

A próxima proposição relaciona a injectividade com a possibilidade de invertermos uma função.

Proposição 4.1.2

Seja f uma função de $A \subset \mathbb{R}^n$ em $B \subset \mathbb{R}^n$. Se f for injectiva, existe uma função $g : f(A) \subset B \rightarrow A$ tal que $(g \circ f)(x) = x$ para todo o $x \in A$. A esta função g chama-se inversa de f e designa-se por f^{-1} .

NOTA: A função inversa terá como domínio a imagem por f de A , ou seja o conjunto $f(A) = \{y \in B : y = f(x), x \in A\}$ e não o conjunto B a não ser que f seja sobrejectiva (isto é $f(A) = B$).

Exemplo 4.1.4 Vamos calcular a inversa da função $f : [\pi, 2\pi] \rightarrow \mathbb{R}$ definida por $f(x) = \cos x$. Sabemos que neste intervalo a função \cos é injectiva (desenhe o gráfico do coseno!). Também sabemos que o contradomínio de f é o intervalo $[-1, 1]$ pelo que a inversa será uma função $f^{-1} : A \subset [-1, 1] \rightarrow [\pi, 2\pi]$. A função $\arccos x$ é a inversa do coseno mas no intervalo $[0, \pi]$. É fácil verificar que a inversa de f é dada por $f^{-1}(y) = 2\pi - \arccos y$.

Exercício 4.1.6 Calcule a inversa da função $f_n(x) = \sin x$, onde $f_n : [(n - 1/2)\pi, (n + 1/2)\pi] \rightarrow \mathbb{R}$.

4.1.1 Exercícios Suplementares

Exercício 4.1.7 Diga se as seguintes funções são ou não injectivas:

1. $f : \{(x, y) \in \mathbb{R}^2 : y \neq 0\} \rightarrow \mathbb{R}^+ \times \mathbb{R}^+$ definida por $f(x, y) = (e^{x/y}, x^2 + y^2)$.
2. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $f(x, y) = (xy, x^2 - y^2)$.
3. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $f(x, y) = (x^2 + 2y^2, 2x^2 + y^2)$.
4. $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$.

Exercício 4.1.8 Mostre que a composição de funções injectivas é uma função injectiva.

Exercício 4.1.9 Dê uma condição para que uma transformação linear de $\mathbb{R}^n \rightarrow \mathbb{R}^n$ seja injectiva.

Exercício 4.1.10

1. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função estritamente monótona. Justifique que f é injectiva.
2. Dê um exemplo de uma função monótona não injectiva.
3. Dê um exemplo de uma função não monótona injectiva.

Exercício 4.1.11 Prove que uma função real de variável real monótona mas não estritamente monótona não é injectiva.

Exercício 4.1.12 Seja $f : \mathbb{R} \rightarrow \mathbb{R}$, contínua. Prove que f é estritamente monótona sse for injectiva. Dê um exemplo de um conjunto $A \subset \mathbb{R}$ e de uma função $f : A \rightarrow \mathbb{R}$ contínua tal que f não seja monótona mas seja injectiva.

Exercício 4.1.13 Mostre que a função $f(v) = \frac{v}{\sqrt{1-v^2}}$ com $v \in]-1, 1[$ é injectiva e determine o seu contradomínio.

Exercício 4.1.14 Mostre que uma função real de variável real par nunca é injectiva.

Exercício 4.1.15 Mostre que uma função real de variável real diferenciável é injectiva se a sua derivada for sempre positiva ou sempre negativa.

Exercício 4.1.16 Consideremos a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}^+ \times \mathbb{R}^2$ definida por

$$f(x, y, z) = (e^{x+z}, (x+y)^3, (x-y)^5).$$

Mostre que ela é injectiva.

Exercício 4.1.17 Considere a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por

$$f(x, y, z) = (4x^2 + y^2 + 2z^2, (x+y-z)^{2n+1}, (x-y)^{4n+3})$$

para $n \in \mathbb{N}$. Determine se ela é injectiva. Determine se a restrição de f a $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$ é ou não injectiva.

Exercício 4.1.18 Mostre que se uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ verificar para todos os pontos \mathbf{x}, \mathbf{y}

$$\|f(\mathbf{x}) - f(\mathbf{y})\| \geq c\|\mathbf{x} - \mathbf{y}\|^p,$$

para alguns $p, c > 0$ então f é injectiva.

Exercício 4.1.19 Prove que a função $f : \{(x, y) \in \mathbb{R}^2 : x > 0, 0 \leq y < 2\pi\} \rightarrow \mathbb{R}$ definida por $f(x, y) = (x \cos y, x \sin y)$ é injectiva e determine a sua inversa.

4.1.2 Sugestões para os exercícios

4.1.7

1. Repare que para $a, b > 0$, as curvas de nível definidas por $\frac{x}{y} = a$ são as rectas definidas por $x = ay$ e as curvas definidas por $x^2 + y^2 = b$ são circunferências de raio \sqrt{b} .
2. Repare que para $a, b \neq 0$, as curvas de nível definidas por $xy = a$ são hipérbolas bem como as definidas por $x^2 - y^2 = (x+y)(x-y) = b$ são também hipérbolas.
3. Ambas as curvas de nível são elipses.
4. $f(1, 0) = f(0, 1)$. Tente descobrir geometricamente porque é que f não é injectiva.

4.1.8 $f(g(x)) = f(g(y)) \Rightarrow g(x) = g(y) \Rightarrow x = y$. ◀

4.1.9 A equação $Ax = y$ tem solução única em \mathbb{R}^n sse $\det A \neq 0$. ◀

4.1.10

1. f é estritamente monótona sse $x < y$ então $f(x) < f(y)$ ou $f(x) > f(y)$.
2. Por exemplo $f(x) = 1$ para $x \in \mathbb{R}$.
3. Por exemplo $f(x) = 1/x$ para $x \in \mathbb{R} \setminus \{0\}$, $f(0) = 0$.

4.1.11 Escreva a definição de função estritamente monótona e compare com a definição de função monótona. ◀

4.1.12 Recorde o que fez no exercício anterior e utilize as propriedades das funções contínuas. ◀

4.1.13 A função é estritamente crescente e portanto injectiva. O seu contradomínio é \mathbb{R} . ◀

4.1.14 Se f é par então $f(x) = f(-x)$. ◀

4.1.15 Se a derivada for sempre positiva ou sempre negativa a função é monótona. ◀

4.1.16 Repare que a função é a composição da transformação linear $(x, y, z) \rightarrow (x+z, x+y, x-y)$ com a função $(x, y, z) \rightarrow (e^x, y^3, z^5)$. Se ambas as funções forem injectivas f também será.

Alternativamente poderá aplicar o método gráfico. ◀

4.1.17 As soluções da equação $f(x, y, z) = (a, b, c)$ estão sobre a intersecção de um elipsóide com dois planos. ◀

4.1.18 Se $f(x) = f(y)$ temos $0 \geq \|f(x) - f(y)\| \geq c\|x - y\|^p$ o que implica $x = y$. ◀

4.1.19 Repare que a inversa pode ter de ser escrita “por ramos” (veja o exemplo 4.1.4). ◀

4.2 Teorema do valor médio para funções vectoriais

Vai ser necessário, em particular ao iniciar o estudo do teorema da função inversa, estimar distâncias no contradomínio de uma função em termos de distâncias no domínio, isto é, estimar $\|F(\mathbf{x}) - F(\mathbf{y})\|$ em termos de $\|\mathbf{x} - \mathbf{y}\|$. Para tal necessitaremos do

Lema 4.2.1 (Teorema do valor médio)

Seja $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função de classe $C^1(S)$. Sejam $\mathbf{x}, \mathbf{y} \in S$ e tais que o segmento de recta que une \mathbf{x} a \mathbf{y} está contido em S . Então

$$\|F(\mathbf{x}) - F(\mathbf{y})\| \leq \sup_{t \in [0,1]} \|DF(t\mathbf{x} + (1-t)\mathbf{y})(\mathbf{x} - \mathbf{y})\|.$$

Ideia da demonstração. Mais uma vez recorreremos ao teorema do valor médio para funções escalares através de uma função auxiliar. Seja $g(t) = (F(\mathbf{x}) - F(\mathbf{y})) \cdot F(t\mathbf{x} + (1-t)\mathbf{y})$. Aplique-se o teorema do valor médio a g no intervalo $[0, 1]$ e estime-se usando a desigualdade de Cauchy-Schwarz.

Este resultado ainda não tem a forma pretendida. Para isso introduzimos

Definição 4.2.1 (Norma de aplicações lineares e de matrizes) Seja $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Definimos a norma de L como sendo

$$\|L\| \equiv \sup_{\|\mathbf{x}\|=1} \|L(\mathbf{x})\|.$$

Seja $A \in \mathbb{M}^{m \times n}$. Definimos a norma de A através de

$$\|A\| = \|L_A\|.$$

em que L_A é a aplicação linear definida canonicamente pela matriz A via $L_A(x) = Ax$. Por vezes consideraremos outras normas para matrizes reais como $\|A\|_2 = \sqrt{\text{tr } A^T A}$ ou $\|A\|_\infty = \max_{i,j} |a_{ij}|$ em que $A = (a_{ij})_{i,j=1,\dots,n}$. Continua a valer nesta situação a observação feita para normas em \mathbb{R}^n de que todas estas normas são equivalentes. O problema seguinte formaliza isso de alguma forma.

Problema 4.2.1 Seja E um espaço vectorial real ou complexo. Designe-se $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ conforme o caso. Uma função $\eta : E \rightarrow \mathbb{R}$ diz-se uma norma em E se verifica as propriedades enumeradas na definição 2.1.1 substituindo $\mathbb{R}^n \leftrightarrow E$ e $x \in \mathbb{R} \leftrightarrow x \in \mathbb{K}$.

1. Verifique que as normas de aplicações lineares e matrizes da definição 4.2.1 são normas nesta acepção geral.
2. Verifique que quaisquer duas normas num espaço vectorial de dimensão finita são equivalentes (adapte o enunciado e solução do exercício 2.1.13).
3. Quais são as melhores constantes na equivalência entre as normas de matrizes mencionadas na definição 4.2.1?

Corolário 4.2.2

Sob as mesmas hipóteses do lema 4.2.1 vale

$$\|F(\mathbf{x}) - F(\mathbf{y})\| \leq \max_{t \in [0,1]} \|DF(t\mathbf{x} + (1-t)\mathbf{y})\| \|\mathbf{x} - \mathbf{y}\|. \quad (4.2)$$

Mais geralmente se $F \in C^1(K)$, com K um conjunto limitado, fechado e convexo², então para todos os $\mathbf{x}, \mathbf{y} \in K$ temos

$$\|F(\mathbf{x}) - F(\mathbf{y})\| \leq \max_{t \in [0,1]} \|DF(t\mathbf{x} + (1-t)\mathbf{y})\| \|\mathbf{x} - \mathbf{y}\|.$$

Ambos os máximos atrás referidos são finitos (porquê?).

Problema 4.2.2 *Convém notar que não existe uma versão do teorema do valor médio para funções vectoriais análoga à conhecida para funções escalares e que envolva uma igualdade da forma $f(\mathbf{b}) - f(\mathbf{a}) = Df(\mathbf{a} + \theta(\mathbf{b} - \mathbf{a}))(\mathbf{b} - \mathbf{a})$. Com efeito, pode verificar que para a função $g: \mathbb{R} \rightarrow \mathbb{R}^2$ definida por $g(t) = (\cos t, \sin t)$ não existe $\theta \in]0, 2\pi[$ tal que $g(2\pi) - g(0) = Dg(\theta)(2\pi)$ embora a desigualdade 4.2.*

Problema 4.2.3 *Seja $A \in \mathbb{M}^{n \times n}$ e L_A a aplicação linear definida canonicamente por A como definido anteriormente. Obtenha uma expressão para o valor de $\|L_A\|$ em termos dos valores próprios de $A^T A$.*

4.3 Teorema da Função Inversa

Em primeira aproximação o teorema da função inversa diz respeito à resolução de sistemas de equações não lineares da forma

$$F(\mathbf{x}) = \mathbf{y} \quad (4.3)$$

em que $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Pretende-se obter, sob condições apropriadas, a garantia de existência de uma função que nos dê \mathbf{x} em função de \mathbf{y} satisfazendo a equação, avaliar a regularidade de tal função e relacionar a derivada da inversa com a derivada de F . Tais objectivos só são exequíveis sob condições particulares e desde que entendamos a existência de inversa num sentido *local*, isto é, dado um ponto \mathbf{x}_0 no domínio de F estabelece-se a existência de vizinhanças V de \mathbf{x}_0 e W de $F(\mathbf{x}_0)$ e de uma função $G: W \rightarrow V$ tal que para todo o $\mathbf{x} \in V$ temos $G(F(\mathbf{x})) = \mathbf{x}$. Nota-se que são casos particulares já conhecidos os seguintes:

Exemplo 4.3.1 (Caso linear) *Suponha-se que $A \in \mathbb{M}$, em que \mathbb{M} designa as matrizes reais $n \times n$, e $\mathbf{b} \in \mathbb{R}^n$. Considere-se*

$$F(\mathbf{x}) \equiv A\mathbf{x} + \mathbf{b}.$$

Então o sistema (4.3) é solúvel se e só se $\det A \neq 0$ e nesse caso podemos obter explicitamente

$$\mathbf{x} = A^{-1}(\mathbf{y} - \mathbf{b}) \equiv F^{-1}(\mathbf{y}).$$

Note-se que neste caso $DF = A$, F^{-1} é diferenciável e $(DF)^{-1} = A^{-1}$.

Exemplo 4.3.2 (Dimensão 1) *Seja $f:]a, b[\rightarrow \mathbb{R}$, $f \in C^1(]a, b[)$, $a < x_0 < b$, $y_0 \equiv f(x_0)$, $f'(x_0) \neq 0$. Então f' mantém o seu sinal numa vizinhança V de x_0 e consequentemente f é estritamente monótona em V . Assim a restrição de f a V , $f|_V$, é invertível, diferenciável e se $g \equiv (f|_V)^{-1}$ temos $g'(y_0) = [f'(x_0)]^{-1}$.*

Nesta situação podemos abdicar de alguma regularidade de f , supondo f unicamente diferenciável em vez de C^1 desde que suponhamos que f' mantém o seu sinal num intervalo J contendo x_0 . Podemos então concluir que f é invertível em J .

²Um subconjunto de um espaço vectorial diz-se convexo se contém qualquer segmento de recta definido por um par dos seus pontos.

A segunda parte do exemplo anterior deve ser contrastado com

Problema 4.3.1 Considere a aplicação³ $\mathbb{R}^2 \ni (x, y) \mapsto (e^x \cos y, e^x \sin y)$. Verifique que o determinante da matriz jacobiana desta função mantém o sinal em \mathbb{R}^2 e no entanto a função não é invertível. No entanto, dado um ponto existe uma vizinhança tal que a restrição da função a essa vizinhança é invertível.

Basta ter em conta o caso linear descrito no exemplo 4.3.1 para constatar que a generalização do teorema da função inversa que procuramos não terá entre as suas hipóteses $DF(\mathbf{x}_0) \neq 0$ por esta hipótese não ser suficiente para garantir invertibilidade. O caso linear sugere fortemente que uma hipótese a considerar seja $DF(\mathbf{x}_0)$ invertível e, de facto, assim é. Uma forte sugestão de que assim será decorre também do seguinte problema

Problema 4.3.2 Considere uma função F definida num aberto, diferenciável e que possui inversa diferenciável.

- a) Verifique a relação $(DF)^{-1} = D(F^{-1})$.
- b) Verifique que se $F \in C^1$ então $F^{-1} \in C^1$.

A importância do teorema da função inversa vai decorrer não só do resultado em si mas também dos métodos a aplicar na demonstração serem susceptíveis de generalização a outras áreas de Matemática⁴. Por isso vamos dedicar algum tempo a motivar e descrever as principais ideias da sua demonstração. No entanto, antes de iniciar a discussão do teorema propriamente dito convém notar alguns factos elementares.

A ideia base consiste na construção da *inversa local* através dum limite de *aproximações sucessivas* sendo cada aproximação construída através da resolução dum problema linear em que um dos dados é o termo anterior da sucessão. Designaremos tal sucessão por $(\mathbf{x}_i)_{i \in \mathbb{N}}$ e a aplicação que associa a cada termo da sucessão o termo seguinte por $T_{\mathbf{y}}$, isto é $T_{\mathbf{y}}(\mathbf{x}_{i-1}) = \mathbf{x}_i$ para $i \in \mathbb{N}$. O domínio de $T_{\mathbf{y}}$ será uma vizinhança V de \mathbf{x}_0 e $\mathbf{y} \in W$ uma vizinhança de $\mathbf{y}_0 \equiv F(\mathbf{x}_0)$. Tais vizinhanças serão escolhidas suficientemente pequenas de maneira a aí se verificarem condições que garantam a convergência da sucessão $(\mathbf{x}_i)_{i \in \mathbb{N}}$. Existem pelo menos duas hipóteses naturais para a escolha de $T_{\mathbf{y}}$ via a substituição de F por uma sua aproximação linear e resolução do sistema linear correspondente:

$$T_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} + DF(\mathbf{x})^{-1}(\mathbf{y} - F(\mathbf{x})) \quad (4.4)$$

ou

$$T_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} + DF(\mathbf{x}_0)^{-1}(\mathbf{y} - F(\mathbf{x})). \quad (4.5)$$

O significado de cada uma destas duas escolhas é ilustrado para o caso unidimensional na figura 4.3. Embora a primeira possa parecer mais natural a segunda tem a vantagem de não ser necessário ter de controlar a variação da derivada $DF(\mathbf{x})$ o que permitirá alguma simplificação do argumento⁵. Consideramos então que $T_{\mathbf{y}}$ está definida por (4.5) com \mathbf{x} e \mathbf{y} em vizinhanças a especificar de \mathbf{x}_0 e \mathbf{y}_0 .

Convencionamos que $T_{\mathbf{y}}^0 \equiv T_{\mathbf{y}}$ e $T_{\mathbf{y}}^k \equiv T_{\mathbf{y}} \circ T_{\mathbf{y}}^{k-1}$. Pretendemos provar que a inversa local, G , é dada por

$$G(\mathbf{y}) = \lim_{k \rightarrow \infty} T_{\mathbf{y}}^k(\mathbf{x}_0). \quad (4.6)$$

Exercício 4.3.1 Experimente aplicar o algoritmo descrito ao problema de determinar zeros de $f(x) = x - x^2$ começando com $x_0 = 1/2$. E com $x_0 = 4$?

³Verificaremos mais tarde que se identificarmos \mathbb{R}^2 a \mathbb{C} da maneira habitual esta aplicação é simplesmente $\mathbb{C} \ni z \mapsto e^z$.

⁴Para a maioria dos alunos tais exemplos serão encontrados ao estudar Análise Numérica e Equações Diferenciais Ordinárias.

⁵A opção pela segunda hipótese é também natural do ponto de vista do método numérico, conhecido por *método de Newton*, a que corresponde, pois evita recalcular e inverter uma matriz em cada iteração

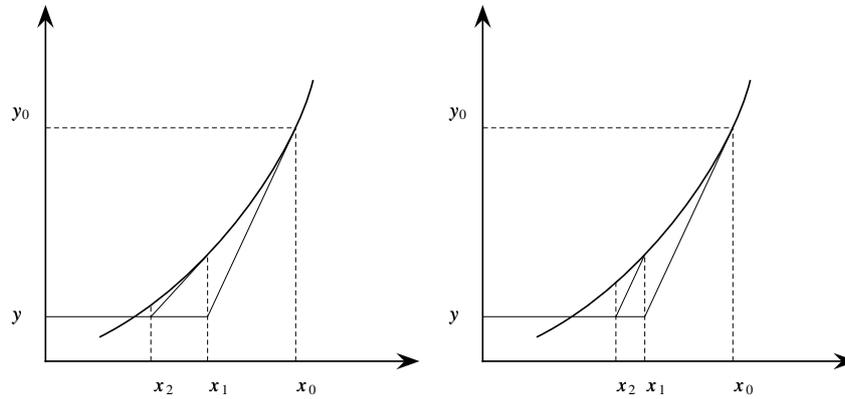


Figura 4.3: Duas hipóteses para a determinação de raízes de uma equação não linear por iterações sucessivas.

Claro que uma hipótese necessária para que estas ideias funcionem será exigir que $DF(\mathbf{x}_0)$ seja invertível ou de forma equivalente que o determinante da matriz jacobiana em \mathbf{x}_0 seja não nulo.

Primeiro verificaremos que o limite em 4.6 existe e só depois que o limite é a solução pretendida. Finalmente verificar-se-á a regularidade da inversa local construída.

Resumindo, os passos essenciais da demonstração do teorema da função inversa são:

1. Mostrar que a sucessão de aproximações sucessivas $(\mathbf{x}_k)_{k \in \mathbb{N}}$ é convergente.
2. Mostrar que o limite da sucessão define uma inversa local.
3. Mostrar que a inversa local é de classe C^1 .

Começemos então por tentar provar que a sucessão de aproximações sucessivas é convergente. Tal será feito à custa de um desenvolvimento “telescópico” dos termos da sucessão da seguinte forma

$$\mathbf{x}_i = \mathbf{x}_0 + \sum_{j=1}^i (\mathbf{x}_j - \mathbf{x}_{j-1}).$$

Tal permite reduzir o estudo da convergência da sucessão ao estudo da convergência da série $\sum \mathbf{x}_j - \mathbf{x}_{j-1}$. Para isso iremos utilizar

Problema 4.3.3 *Seja $\sum_{k=1}^{+\infty} \mathbf{x}_k$ uma série de termo geral em \mathbb{R}^n . Prove que se a série $\sum_{k=1}^{+\infty} \|\mathbf{x}_k\|$ é convergente em \mathbb{R} então a série é convergente⁶.*

Para provar a convergência da série $\sum_j \|\mathbf{x}_j - \mathbf{x}_{j-1}\|$ tentaremos estabelecer condições que garantem que o seu termo geral é majorado pelo de uma série geométrica convergente. Isto equivale a exigir que

$$\|\mathbf{x}_{j+1} - \mathbf{x}_j\| \leq \rho \|\mathbf{x}_j - \mathbf{x}_{j-1}\|$$

para alguma constante ρ , com $0 \leq \rho < 1$. Ora

$$\mathbf{x}_{j+1} - \mathbf{x}_j = T_{\mathbf{y}}(\mathbf{x}_j) - T_{\mathbf{y}}(\mathbf{x}_{j-1})$$

pelo que tal objectivo estará garantido se a aplicação $T_{\mathbf{y}}$ verificar para todo o $\mathbf{x}, \mathbf{z} \in V$ e todo o $\mathbf{y} \in W$

$$\|T_{\mathbf{y}}(\mathbf{x}) - T_{\mathbf{y}}(\mathbf{z})\| \leq \rho \|\mathbf{x} - \mathbf{z}\|. \quad (4.7)$$

⁶Nestas condições diz-se que a série é *absolutamente convergente*.

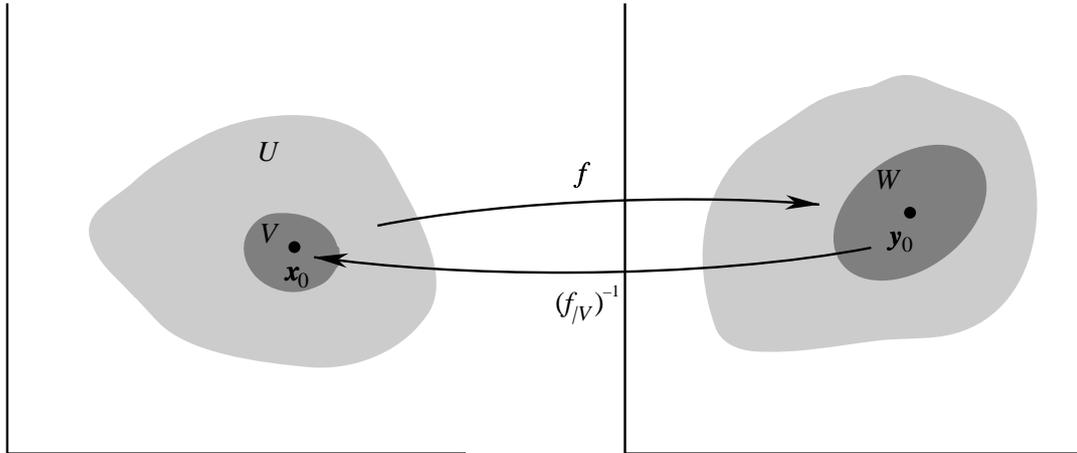


Figura 4.4: Algumas das convenções na demonstração do teorema da função inversa.

Com efeito

$$\begin{aligned}
 \|T_{\mathbf{y}}(\mathbf{x}) - T_{\mathbf{y}}(\mathbf{z})\| &= \|\mathbf{x} - \mathbf{z} - DF(\mathbf{x}_0)^{-1}(F(\mathbf{x}) - F(\mathbf{z}))\| \\
 &= \|DF(\mathbf{x}_0)^{-1}(F(\mathbf{x}) - F(\mathbf{z}) - DF(\mathbf{x}_0)(\mathbf{x} - \mathbf{z}))\| \\
 &\leq M\|(F(\mathbf{x}) - F(\mathbf{z}) - DF(\mathbf{x}_0)(\mathbf{x} - \mathbf{z}))\| \\
 &\leq M\alpha\|\mathbf{x} - \mathbf{z}\|,
 \end{aligned}$$

em que $M = \|DF(\mathbf{x}_0)^{-1}\|$ e na última passagem usou-se o teorema do valor médio aplicado à função $H(\mathbf{x}) = F(\mathbf{x}) - DF(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ sendo $\alpha = \sup_{\mathbf{x} \in V} \|DH(\mathbf{x})\|$ e exigindo que V seja convexa (uma bola). Notando que $DH(\mathbf{x}_0) = 0$ e usando a continuidade das derivadas parciais de F , concluímos que podemos fazer $M\alpha < 1$ desde que V seja suficientemente pequena (uma bola com raio suficientemente pequeno $B_\epsilon(\mathbf{x}_0)$).

Provamos agora que os termos das sucessões e os seus limites pertencem à vizinhança V de \mathbf{x}_0 desde que W e V sejam escolhidas suficientemente pequenas. Isto completará a justificação da convergência da sucessão de aproximações sucessivas. Para isso estimamos

$$\begin{aligned}
 \|T_{\mathbf{y}}(\mathbf{x}) - \mathbf{x}_0\| &= \|\mathbf{x} - \mathbf{x}_0 + DF(\mathbf{x}_0)^{-1}(y - F(\mathbf{x}_0) + F(\mathbf{x}_0) - F(\mathbf{x}))\| \\
 &\leq \|DF(\mathbf{x}_0)^{-1}(y - F(\mathbf{x}_0))\| + \|\mathbf{x} - \mathbf{x}_0 + DF(\mathbf{x}_0)^{-1}(F(\mathbf{x}_0) - F(\mathbf{x}))\| \\
 &\leq M\|y - F(\mathbf{x}_0)\| + M\|(F(\mathbf{x}_0) - F(\mathbf{x})) - DF(\mathbf{x}_0)(\mathbf{x}_0 - \mathbf{x})\| \\
 &\leq M\|y - F(\mathbf{x}_0)\| + M\alpha\|\mathbf{x}_0 - \mathbf{x}\|
 \end{aligned}$$

de onde podemos concluir que, se escolhermos $\epsilon > 0$ de maneira a que $\|\mathbf{x}_0 - \mathbf{x}\| < \epsilon$ garanta $M\alpha < 1/3$, podemos escolher também $\|y - F(\mathbf{x}_0)\| < r$ de maneira a $Mr < \epsilon/3$, donde $\|T_{\mathbf{y}}(\mathbf{x}) - \mathbf{x}_0\| < \frac{2}{3}\epsilon$. Logo todos os termos das sucessões e os seus limites estarão em $\overline{B}_{\frac{2}{3}\epsilon}(\mathbf{x}_0) \subset B_\epsilon(\mathbf{x}_0) \equiv V$.

Notamos também que $G(\mathbf{y})$ satisfaz $F(G(\mathbf{y})) = \mathbf{y}$ se $T_{\mathbf{y}}(G(\mathbf{y})) = G(\mathbf{y})$. Esta última igualdade decorre da continuidade de $T_{\mathbf{y}}$ que por sua vez decorre da desigualdade (4.7). Assim G satisfaz $F(G(\mathbf{y})) = \mathbf{y}$.

Ainda não provámos que numa vizinhança suficientemente pequena de \mathbf{x}_0 a função F é injectiva. A não injectividade corresponderia à existência de $\mathbf{x}, \mathbf{z}, \mathbf{x} \neq \mathbf{z}$ tais que $F(\mathbf{x}) = F(\mathbf{z}) = \mathbf{y}$. Teríamos então também $T_{\mathbf{y}}(\mathbf{x}) = T_{\mathbf{y}}(\mathbf{z})$. Assim $\|\mathbf{x} - \mathbf{z}\| = \|T_{\mathbf{y}}(\mathbf{x}) - T_{\mathbf{y}}(\mathbf{z})\| \leq \rho\|\mathbf{x} - \mathbf{z}\|$ com $0 < \rho < 1$, para $\mathbf{x}, \mathbf{z} \in B_\epsilon(\mathbf{x}_0)$, o que é impossível a não ser que $\mathbf{x} = \mathbf{z}$.

Resta agora estabelecer propriedades da inversa local. Note-se que, se estabelecermos que a inversa local é diferenciável com derivada dada por $DG(F(\mathbf{x})) = (DF(\mathbf{x}))^{-1}$, a continuidade de DF mais a fórmula explícita para a matriz jacobiana de G estabelecem que $DG \in C^{17}$. A

⁷De forma análoga se $F \in C^k$ então $G \in C^k$ com $k \geq 2$ ou $k = \infty$

unicidade local da inversa permite limitarmo-nos a analisar a diferenciabilidade em $\mathbf{y}_0 = F(\mathbf{x}_0)$. Para tal, convencionando $F(\mathbf{x}) = \mathbf{y}$ e $\mathbf{x} \neq \mathbf{x}_0$, considera-se

$$\begin{aligned} E &\equiv \|G(\mathbf{y}) - G(\mathbf{y}_0) - DF(\mathbf{x}_0)^{-1}(\mathbf{y} - \mathbf{y}_0)\| \\ &= \|\mathbf{x} - \mathbf{x}_0 - DF(\mathbf{x}_0)^{-1}(F(\mathbf{x}) - F(\mathbf{x}_0))\| \\ &= \|\mathbf{x} - \mathbf{x}_0 - DF(\mathbf{x}_0)^{-1}(DF(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|))\| \\ &= \|-DF(\mathbf{x}_0)^{-1}(o(\|\mathbf{x} - \mathbf{x}_0\|))\| \\ &\leq \|DF(\mathbf{x}_0)^{-1}\| \|o(\|\mathbf{x} - \mathbf{x}_0\|)\|. \end{aligned}$$

Demonstrar a diferenciabilidade de G em \mathbf{y}_0 com derivada $DF(\mathbf{x}_0)^{-1}$ corresponde a mostrar que $E/\|\mathbf{y} - \mathbf{y}_0\| \rightarrow 0$ quando $\mathbf{y} \rightarrow \mathbf{y}_0$ o que a desigualdade anterior permite reduzir a mostrar que $\frac{\|\mathbf{x} - \mathbf{x}_0\|}{\|\mathbf{y} - \mathbf{y}_0\|}$ é limitado para \mathbf{x} numa vizinhança de \mathbf{x}_0 e que G é contínua em \mathbf{x}_0 (o que garante que $E/\|\mathbf{x} - \mathbf{x}_0\| \rightarrow 0$ quando $\mathbf{y} \rightarrow \mathbf{y}_0$). Quanto à primeira destas questões observamos que

$$\frac{\|\mathbf{x} - \mathbf{x}_0\|}{\|\mathbf{y} - \mathbf{y}_0\|} = \frac{\|\mathbf{x} - \mathbf{x}_0\|}{\|DF(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)\|} \leq \frac{1}{2\|DF(\mathbf{x}_0)\|^{-1}}$$

em que no último passo escolheu-se \mathbf{x} numa vizinhança de \mathbf{x}_0 de forma a termos $\frac{\|o(\|\mathbf{x} - \mathbf{x}_0\|)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \leq \frac{1}{2\|DF(\mathbf{x}_0)\|^{-1}}$ e usou-se a estimativa $\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \geq \frac{1}{\|A^{-1}\|}$ válida para um operador linear não singular $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (demonstre-a!). Quanto à continuidade de G em \mathbf{y}_0 deixamos ao cuidado do leitor estabelecer que estimativas já obtidas permitem afirmar que dado $\epsilon > 0$ existem $\bar{\epsilon}$ com $0 < \bar{\epsilon} < \epsilon$ e $r > 0$ tais que $\|\mathbf{y} - \mathbf{y}_0\| < r$ e $\|\mathbf{x} - \mathbf{x}_0\| < \bar{\epsilon}$ implicam $\|T_{\mathbf{y}}(\mathbf{x}) - \mathbf{x}_0\| < \bar{\epsilon}$. Consequentemente, por indução obtém-se que $\|T_{\mathbf{y}}^k(\mathbf{x}_0) - \mathbf{x}_0\| < \bar{\epsilon}$ para todo o $k \in \mathbb{N}$ e por passagem ao limite $\|G(\mathbf{y}) - G(\mathbf{y}_0)\| < \epsilon$.

Provou-se assim:

Teorema 4.3.1 (Função Inversa)

Seja $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função de classe $C^1(U)$ em que U é um aberto e seja $\mathbf{x}_0 \in U$ tal que $DF(\mathbf{x}_0)$ é não singular, isto é,

$$\det DF(\mathbf{x}_0) \neq 0.$$

Então

- i) existem vizinhanças V de \mathbf{x}_0 e W de $F(\mathbf{x}_0)$ tais que F é uma bijecção de V sobre W e portanto $F|_V^{-1} : W \rightarrow V$ está bem definida;
- ii) $G = F|_V^{-1} \in C^1(W)$;
- iii) a derivada da função $G = F|_V^{-1}$ no ponto $\mathbf{y} = f(\mathbf{x})$ verifica

$$D(G)(\mathbf{y}) = (DF(\mathbf{x}))^{-1},$$

para todo o $\mathbf{x} \in V$ ou todo o $\mathbf{y} \in W$. Adicionalmente se $F \in C^k(U)$ com $k \in \mathbb{N}$ ou $k = \infty$ então $G \in C^k(W)$.

Convém acentuar que o teorema da função inversa não garante invertibilidade global e não é susceptível de ser melhorado nesse sentido devido a exemplos como o do problema 4.3.1.

Exemplo 4.3.3 Consideremos a função $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$ definida por

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

O seu jacobiano⁸ é dado por

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \det \begin{bmatrix} y & x \\ -2x & 2y \end{bmatrix} = 2(x^2 + y^2).$$

⁸ Jacobiano é uma abreviatura de determinante da matriz jacobiana.

Como este valor é sempre positivo (notemos que a origem foi excluída do domínio), o teorema da função inversa garante a invertibilidade local desta função. No entanto $f(1, 1) = f(-1, -1)$ e logo a função não pode ser globalmente invertível por não ser injectiva.

Exercício 4.3.2 Seja $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$ definida por

$$f(x, y) = (xy, y^6 - x^6)$$

Mostre que f é localmente injectiva mas não globalmente injectiva.

Exercício 4.3.3 Mostre que $f : \mathbb{R} \rightarrow \mathbb{R}^2$ definida por $f(\theta) = (\cos \theta, \sin \theta)$ é localmente mas não globalmente injectiva (não utilize o teorema da função inversa para provar injectividade local mas sim um raciocínio ad hoc).

O teorema da função inversa não só garante, em determinadas condições, a existência da inversa de uma função f , como também permite calcular, graças à garantia de diferenciabilidade numa vizinhança, todas as derivadas de f^{-1} até à ordem m , o grau de diferenciabilidade de f . O exemplo seguinte ilustra este facto:

Exemplo 4.3.4 Seja $f(x) = x + x^2$. O teorema da função inversa garante que f é invertível numa vizinhança de $x = 0$. Seja g a inversa de f . Então temos $g(0) = 0$ e

$$f(g(y)) = g(y) + g(y)^2 = y.$$

Diferenciando a identidade anterior obtemos

$$\frac{\partial g}{\partial y}(y) + 2g(y) \frac{\partial g}{\partial y}(y) = 1. \quad (4.8)$$

Portanto em $y = 0$ temos

$$\frac{\partial g}{\partial y}(0) = 1.$$

Diferenciando a identidade 4.8 obtemos

$$\frac{\partial^2 g}{\partial y^2}(y) + 2 \frac{\partial g}{\partial y}(y)^2 + 2g(y) \frac{\partial^2 g}{\partial y^2}(y) = 0$$

de onde se tira

$$\frac{\partial^2 g}{\partial y^2}(0) = -2.$$

Exercício 4.3.4 Calcule a inversa da função $f(x) = x + x^2$ e confirme o resultado do exemplo anterior.

Exemplo 4.3.5 Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função C^∞ . Se $f'(x) \neq 0$, então é possível calcular todas as derivadas de f^{-1} no ponto $y = f(x)$ usando o método seguinte:

Consideremos a identidade

$$f(f^{-1}(y)) = y.$$

Derivando ambos os membros da expressão anterior obtemos, pela regra da derivação da função composta,

$$\frac{\partial f}{\partial x}(f^{-1}(y)) \frac{\partial f^{-1}}{\partial y}(y) = 1. \quad (4.9)$$

Da equação anterior obtemos

$$\frac{\partial f^{-1}}{\partial y}(y) = \frac{1}{\frac{\partial f}{\partial x}(f^{-1}(y))}.$$

Para o cálculo de $\frac{\partial^2 f^{-1}}{\partial y^2}(y)$ derivamos novamente (4.9), obtendo

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

de onde concluímos

$$\frac{\partial^2 f^{-1}}{\partial y^2}(y) = - \frac{\frac{\partial^2 f}{\partial x^2}(f^{-1}(y)) \left(\frac{\partial f^{-1}}{\partial y}(y) \right)^2}{\frac{\partial f}{\partial x}(f^{-1}(y))^2}.$$

Assim, calculando sucessivamente as diversas derivadas de f^{-1} podemos desenvolver esta função em fórmula de Taylor em torno do ponto y e portanto, numa vizinhança suficientemente pequena, aproximá-la com precisão arbitrária.

Exercício 4.3.5 Utilizando as ideias do exemplo anterior calcule

$$\frac{\partial^3 f^{-1}}{\partial y^3}(y).$$

Aplicando o exemplo anterior a uma função podemos obter a fórmula de Taylor de $f^{-1}(y)$ em torno de um ponto

Exercício 4.3.6 Seja $f(x) = x + e^x$.

1. Prove que f é injectiva e portanto a inversa f^{-1} existe.
2. Calcule o desenvolvimento de Taylor de $f^{-1}(y)$, em torno de $y = f(0) = 1$ até à terceira ordem.

Para o cálculo de primeiras derivadas da inversa de uma função o teorema da função inversa dá-nos uma expressão explícita, que pode ser aplicada directamente.

Exemplo 4.3.6 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a função dada por

$$f(x, y) = (x + y + x^3 y - xy + 1, x - y + x^4 \operatorname{ch} y).$$

Sabemos que $f(0, 0) = (1, 0)$. Podemos facilmente provar que a função f admite inversa local definida numa vizinhança do ponto $(1, 0)$, sendo $f^{-1}(1, 0) = (0, 0)$. De facto temos

$$J_f = \begin{bmatrix} 1 + 3x^2 y - y & 1 + x^3 - x \\ 1 + 4x^3 \operatorname{ch} y & -1 + x^4 \operatorname{sh} y \end{bmatrix}.$$

No ponto $(x, y) = (0, 0)$ obtemos

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

Assim, como $\det Df = -2 \neq 0$ e a função é de classe C^∞ existe inversa $f^{-1} = (f_1^{-1}, f_2^{-1})$ também C^∞ numa vizinhança de $(u, v) = (1, 0)$ e verificando

$$Df^{-1}(1, 0) = \left[\begin{array}{cc} \frac{\partial f_1^{-1}}{\partial u} & \frac{\partial f_1^{-1}}{\partial v} \\ \frac{\partial f_2^{-1}}{\partial u} & \frac{\partial f_2^{-1}}{\partial v} \end{array} \right] \Big|_{(u,v)=(1,0)} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}.$$

Podemos também aplicar ideias semelhantes às do exemplo (4.3.5) para funções de \mathbb{R}^n em \mathbb{R}^n . Utilizando um procedimento análogo, resolva então o seguinte exercício:

Exercício 4.3.7 Determine $\frac{\partial^2 f^{-1}}{\partial u^2}(u, v)$ com $(u, v) = f(x, y) = (xy, x^2 - y^2)$.

Com o teorema da função inversa também podemos dar uma condição de injectividade local sobre funções de $\mathbb{R}^n \rightarrow \mathbb{R}^m$, com $m \geq n$, como se pode verificar no exemplo seguinte:

Exemplo 4.3.7 Se $\text{car}[Df] = n$ (isto é a característica de Df ou seja o número de linhas ou colunas linearmente independentes de Df for n) então a função f é localmente injectiva.

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, com $f(x) = (f_1(x), \dots, f_m(x))$. Se $\text{car}[Df] = n$ no ponto x_0 então existem índices i_1, \dots, i_n tais que a matriz

$$\begin{bmatrix} \frac{\partial f_{i_1}}{\partial x_1} & \cdots & \frac{\partial f_{i_1}}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_{i_n}}{\partial x_1} & \cdots & \frac{\partial f_{i_n}}{\partial x_n} \end{bmatrix}.$$

tem determinante não nulo. Então, pelo teorema da função inversa a função $g(x) = (f_{i_1}, \dots, f_{i_n})$ é localmente injectiva, pelo que f também será localmente injectiva.

Exercício 4.3.8 Mostre, usando o exemplo anterior, que a aplicação $\mathbb{R} \ni x \mapsto (\sin x, \cos x)$ é localmente injectiva.

4.3.1 Exercícios Suplementares

Exercício 4.3.9 Considere o sistema de equações

$$\begin{cases} u = xy + \sin(x+y), \\ v = e^{-x+y-2} + \frac{x}{y}. \end{cases}$$

Mostre que existem vizinhanças de $(u, v) = (-1, 0)$ e de $(x, y) = (-1, 1)$ tais que o sistema define (x, y) como uma função C^1 de (u, v) desde que as variáveis estejam nessas vizinhanças. Calcule $\frac{\partial x}{\partial u}(-1, 0)$.

Exercício 4.3.10 Considere o sistema de equações não lineares

$$\begin{cases} u = x^2y^3 + \sin(x+y) - 1, \\ v = \sin(xy) + x - y + 1. \end{cases}$$

- Mostre que existem vizinhanças de $(x, y) = (0, 0)$ e de $(u, v) = (-1, 1)$ tais que aquele sistema define (x, y) como uma função C^∞ de (u, v) em tais vizinhanças.
- Calcule a matriz jacobiana da função cuja existência garantiu na alínea anterior no ponto $(-1, 1)$.

Exercício 4.3.11 Considere a função real de variável real definida por $f(x) = \cos x$.

- Qual a maior vizinhança V do ponto $-\pi/4$ tal que $f|_V$ é injectiva? Calcule a inversa de f em V .
- Existe alguma vizinhança de π na qual a função f seja injectiva?
- Calcule uma inversa local de f , $f_{\pi/2}^{-1}$, tal que $f_{\pi/2}^{-1}(0) = \frac{\pi}{2}$.
- Calcule uma inversa local de f , $f_{3\pi/2}^{-1}$, tal que $f_{3\pi/2}^{-1}(0) = \frac{3\pi}{2}$.

Exercício 4.3.12 Mostre que a função $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$ é localmente mas não globalmente injectiva

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

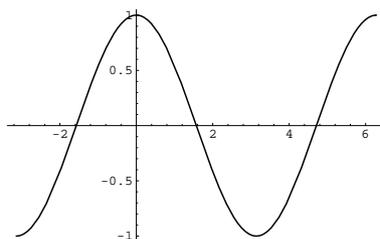


Figura 4.5: Gráfico do cosseno

Exercício 4.3.13 Mostre que a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, definida por

$$f(x, y, z) = (\sin(x + y), \sin(x - y), z^3 - z^5),$$

é localmente injetiva em torno $(0, 0, 0)$ mas que não é globalmente injetiva.

Exercício 4.3.14 Considere o sistema de equações

$$\begin{cases} u = xy + \sin(x + y), \\ v = e^{-x+y-2} + \frac{x}{y}. \end{cases}$$

Mostre que existem vizinhanças de $(u, v) = (-1, 0)$ e de $(x, y) = (-1, 1)$ tais que o sistema define (x, y) como uma função C^1 de (u, v) desde que as variáveis estejam nessas vizinhanças. Calcule $\frac{\partial x}{\partial u}(-1, 0)$.

Exercício 4.3.15 Seja $f : A \rightarrow \mathbb{R}^n$ uma função de classe C^1 , onde A é um conjunto aberto. Prove que para cada conjunto compacto C , $C \subset A$ onde $Jf \neq 0$ existe um número finito de conjuntos abertos U_i tais que $C \subset \cup U_i$ e f é invertível em cada U_i . Sugestão: Utilize o teorema de Heine-Borel.

Exercício 4.3.16 Considere a função

$$f(x, y) = (\sin x \operatorname{arccotg} y, \cos x \operatorname{arccotg} y).$$

Prove que

1. $J_f(x, y) \neq 0$ para todo x e y .
2. f não é injetiva.

4.3.2 Sugestões para os exercícios

4.3.11

1. Observe a figura 4.5 e repare que a restrição da função $\cos x$ ao intervalo $[-\pi, 0]$ é injetiva.
2. Utilize a figura 4.5.
3. Uma inversa possível será $\arccos x$, para $x \in [-1, 1]$.
4. Uma inversa nas condições requeridas é $2\pi - \arccos x$, para $x \in [-1, 1]$.

4.3.12 Temos

$$Df = \begin{bmatrix} y & x \\ 4x & -10y \end{bmatrix}$$

e portanto $\det Df = -10y^2 - 4x^2 \neq 0$ para $x, y \neq 0$. Deste modo o teorema da função inversa garante a injectividade local de f . Temos também que $f(1, 1) = f(-1, -1)$ pelo que f não é injectiva. ◀

4.3.13 Repare que a função $(\sin(x + y), \sin(x - y))$ é localmente injectiva numa vizinhança da origem. é também fácil verificar que numa vizinhança de $z = 0$ a função $z^3 - z^5$ é injectiva. Com estes resultados é fácil provar que f é localmente injectiva em torno da origem. $f(x, y, 0) = f(x, y, 1)$, pelo que f não pode ser injectiva. ◀

4.3.14 Definindo $f(x, y) = (xy + \sin(x + y), e^{-x+y-2} + \frac{x}{y})$, temos

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

Como o determinante desta matriz é não nulo a função é localmente invertível e

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

4.3.15 Como o jacobiano não se anula, para cada ponto existe uma vizinhança U_x onde a função é invertível. O conjunto de todas estas vizinhanças é uma cobertura de C . Como este conjunto é compacto podemos extrair uma subcobertura finita (pelo teorema de Heine-Borel). ◀

4.3.16 O primeiro resultado obtém-se pelo método usual, sendo o jacobiano dado por

$$Jf = -\frac{\operatorname{arccotg} y}{1 + y^2},$$

que nunca se anula. Note que a função $\operatorname{arccotg}$ não está definida na origem. ◀

4.4 Teorema da Função Implícita

Frequentemente necessitamos de resolver equações na forma

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

em ordem a \mathbf{x} ou \mathbf{y} . Porém, mesmo para expressões simples, esta tarefa pode ser extremamente difícil. Para se convencer disto...

Exercício 4.4.1 Tente resolver a equação

$$y + \sin y - x = 0 \tag{4.10}$$

de modo a obter y em função de x . Aparentemente não existe solução explícita elementar mas nunca se sabe...

No entanto, conhecemos uma solução da equação 4.10, nomeadamente $y = x = 0$. Para além disso numa vizinhança da origem, a função $x = y + \sin y$ tem inversa pois é injectiva. Portanto, aplicando o teorema da função inversa, poderíamos mostrar a existência de uma função $y(x)$ tal que $y(x) + \sin(y(x)) - x = 0$.

Exercício 4.4.2 Confirme as afirmações anteriores.

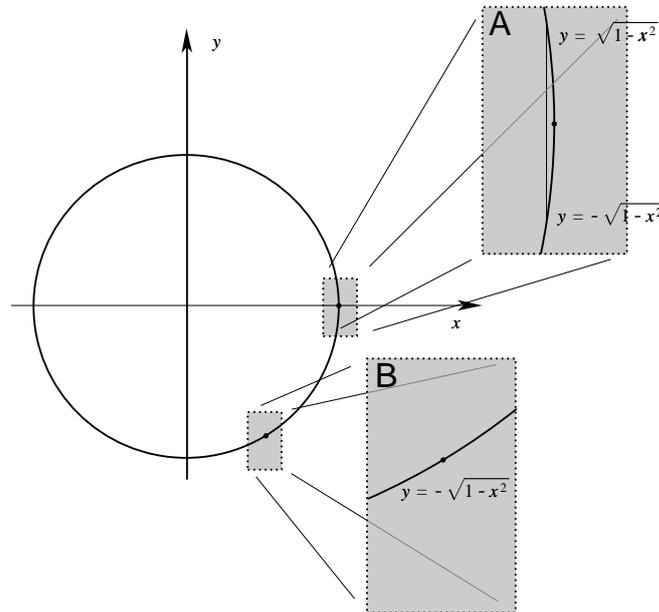


Figura 4.6: A porção da circunferência $x^2 + y^2 - 1 = 0$ ampliada em **A** não é um gráfico de uma função da forma $y = g(x)$. O mesmo acontece com a intersecção da circunferência com uma qualquer vizinhança de $(1, 0)$.

Nesta secção estudaremos um teorema que nos dá condições suficientes para se poder resolver equações $f(\mathbf{x}, \mathbf{y}) = 0$, uma vez conhecidos pontos \mathbf{x}_0 e \mathbf{y}_0 para os quais $f(\mathbf{x}_0, \mathbf{y}_0) = 0$ e que é conhecido por teorema da *função implícita*. Este teorema é um resultado intimamente ligado ao teorema da função inversa. De facto são equivalentes e o estabelecer essa equivalência é relativamente fácil embora mostrar que o teorema da função implícita é uma consequência do teorema da função inversa possa parecer, numa primeira análise, pouco natural.

Começemos por algumas observações ainda não inteiramente precisas para estabelecer um primeiro paralelismo entre os dois resultados. De um ponto de vista de solução de equações não lineares, o teorema da função inversa lida, como vimos, com a solução local de equações da forma $\mathbf{y} = F(\mathbf{x})$ em ordem a \mathbf{x} , em que $F \in C^1$ e uma certa aplicação linear (a derivada de F) é não singular num ponto \mathbf{x}_0 em torno do qual a invertibilidade local é assegurada. De forma análoga, o teorema da função implícita lida com a solução local de equações da forma $F(\mathbf{x}, \mathbf{y}) = 0$ em ordem a \mathbf{y} , em que $F \in C^1$ e uma certa aplicação linear relacionada com a derivada de F é não singular num ponto $(\mathbf{x}_0, \mathbf{y}_0)$ em torno do qual a resolubilidade fica assegurada. Começamos por ilustrar um tal problema numa situação em que se pode explicitamente chegar às mesmas conclusões e as dimensões dos espaços envolvidos são as mais baixas possíveis.

Exemplo 4.4.1 (Caso particular em dimensão 2) *Considere-se a equação da circunferência $x^2 + y^2 - 1 = 0$. Na vizinhança de que pontos em que é verificada é que esta equação define y como função de x ? Resolvendo a equação em ordem a y , ou melhor ainda esboçando o seu gráfico (ver a fig. 4.6), facilmente se reconhece que qualquer que seja o ponto sobre o gráfico desta circunferência, excepto os pontos $(-1, 0)$ e $(1, 0)$, é possível escolher uma vizinhança suficiente pequena desse ponto cuja intersecção com o conjunto definido pela equação verifica $y = \sqrt{1 - x^2}$ ou $y = -\sqrt{1 - x^2}$. É o que se ilustra na fig. 4.6 em **B**. Por outro lado numa vizinhança de um dos dois pontos excepcionais tal é sempre impossível, é o que se ilustra na fig. 4.6 em **A**.*

O carácter excepcional dos pontos $(1, 0)$ e $(-1, 0)$ obviamente tem a ver com o facto da tangente à circunferência nestes pontos ser vertical ou, se recordarmos que o gradiente de uma campo

escalar é ortogonal às suas linhas de nível, com o facto de designando $f(x, y) = x^2 + y^2 - 1$ temos $\frac{\partial f}{\partial y}(1, 0) = \frac{\partial f}{\partial y}(-1, 0) = 0$.

Uma outra fonte de inspiração para compreender este tipo de problemas é a álgebra linear.

Exemplo 4.4.2 (Caso linear) *Seja T uma transformação linear de \mathbb{R}^{m+n} em \mathbb{R}^m , com $n, m \geq 1$. Suponhamos que $T(\mathbf{x}) = A\mathbf{x}$ onde A é uma matriz $m \times (m+n)$. Se $\text{car } A = m$ (recorde que a característica de A , $\text{car } A$, é o número de linhas ou colunas linearmente independentes) então a equação $T(\mathbf{x}) = 0$ permite definir m coordenadas de \mathbf{x} em função das restantes n .*

Observemos agora alguns factos. Primeiro, temos $T(0) = 0$. Segundo, se $\text{car } A = m$ então existem m colunas linearmente independentes. Podemos, sem perda de generalidade, supor que são as m primeiras (se isto não fosse verdade seria sempre possível permutar as colunas da matriz, fazendo uma mudança de variáveis). A matriz A pode ser escrita como

$$A = \begin{bmatrix} a_{11} & \dots & a_{1m} & \dots & a_{1\ m+n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mm} & \dots & a_{m\ m+n} \end{bmatrix}.$$

A derivada de T em ordem às primeiras m variáveis, (x_1, \dots, x_m) , é representada pela matriz

$$\tilde{A} = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mm} \end{bmatrix}$$

e $\det \tilde{A} \neq 0$. Note que é o facto de \tilde{A} ter determinante não nulo que permite determinar as primeiras m coordenadas em função das $n - m$ restantes.

Exercício 4.4.3 *Resolva o sistema*

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

de modo a obter x e z como função de y .

O teorema da função implícita, que enunciamos de seguida, generaliza (de um forma bastante poderosa) o exemplo anterior. Vai ser conveniente no seu enunciado e para cálculos posteriores introduzir a notação $\frac{\partial(f_{j_1}, f_{j_2}, \dots, f_{j_k})}{\partial(x_{i_1}, x_{i_2}, \dots, x_{i_l})}$, com $1 \leq j_1 < j_2 < \dots < j_k \leq m$, $1 \leq i_1 < i_2 < \dots < i_l \leq n$, para representar a derivada de uma função que se obtém de uma função dada $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ só considerando k das suas componentes $f_{j_1}, f_{j_2}, \dots, f_{j_k}$ e fixando $n - l$ das suas variáveis, ou seja considerando-a só como função de l variáveis $x_{i_1}, x_{i_2}, \dots, x_{i_l}$. Tal derivada é representada pela submatriz da matriz jacobiana correspondente a considerar as colunas de índices i_1, i_2, \dots, i_l e as linhas de índices j_1, j_2, \dots, j_k a que também nos referiremos usando a mesma notação. No exemplo 4.4.2 podíamos ter escrito $\tilde{A} = \frac{\partial T}{\partial(x_1, \dots, x_m)}$.

Teorema 4.4.1 (Função Implícita)

Seja $f : U \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ uma função de classe $C^p(\text{int } U)$. Suponha-se que, no ponto $(\mathbf{x}_0, \mathbf{y}_0) \in \text{int } U$ ($\mathbf{x}_0 \in \mathbb{R}^n$ e $\mathbf{y}_0 \in \mathbb{R}^m$) verifica-se $f(\mathbf{x}_0, \mathbf{y}_0) = 0$ e

$$\det \frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \neq 0.$$

Então, existe uma vizinhança V de \mathbf{x}_0 na qual a equação

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

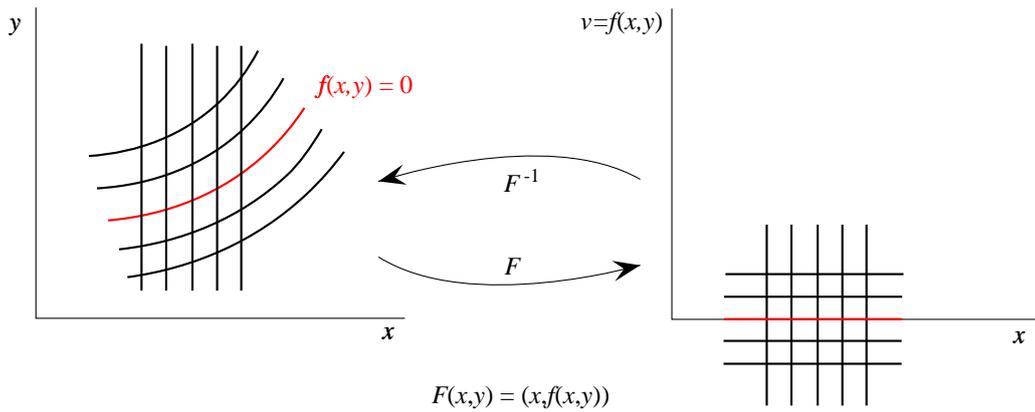


Figura 4.7: Na demonstração do teorema da função implícita pelo processo sugerido note que a função é construída pela composição das funções $x \mapsto (x, 0)$, F^{-1} e $(x, y) \mapsto y$ por esta ordem. Claro que F^{-1} designa uma inversa local.

define uma única função $g \in C^p(V)$, $g : V \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, para a qual

$$f(\mathbf{x}, g(\mathbf{x})) = 0$$

para todo o $\mathbf{x} \in V$. Adicionalmente, a derivada de g em V satisfaz

$$Dg(\mathbf{x}) = - \left[\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, g(\mathbf{x})) \right]^{-1} \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, g(\mathbf{x})). \tag{4.11}$$

Notemos que o resultado do teorema é local, ao contrário do exemplo 4.4.2 que é global, isto é se $\text{car } A = m$, podemos sempre resolver a equação $A\mathbf{x} = 0$ em ordem a m coordenadas. No entanto, isto só é válido porque o sistema naquele exemplo é linear. Em geral não temos nenhuma garantia de que possamos resolver uma equação da forma $F(\mathbf{x}, \mathbf{y}) = 0$ em ordem, por exemplo, a \mathbf{y} , para qualquer valor de \mathbf{x} , mesmo que num dado ponto sejam verificadas as condições do teorema da função implícita (o teorema apenas garante a existência de soluções na vizinhança do ponto).

O teorema da função implícita pode ser demonstrado de uma forma análoga à do teorema da função inversa por linearização e aproximações sucessivas. É, no entanto, muito mais simples demonstrá-lo à custa do teorema da função inversa.

Problema 4.4.1 *Demonstre o teorema da função implícita. Para tal, usando as convenções do enunciado do teorema, considere a aplicação $F : U \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ definida por $F(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, f(\mathbf{x}, \mathbf{y}))$ e aplique-lhe o teorema da função inversa relativamente ao ponto $(\mathbf{x}_0, \mathbf{y}_0)$.*

Exemplo 4.4.3 *Consideremos a equação $f(x, y) = x^2 + y + \text{sen}(x^2 + y^2) = 0$. Então como $f(0, 0) = 0$ e $\frac{\partial f}{\partial y}(0, 0) = 1 \neq 0$, existe uma função $g(x)$, definida para $|x|$ suficientemente pequeno tal que $f(x, g(x)) = 0$. Para além disto temos*

$$\frac{\partial g}{\partial x}(0) = - \frac{\frac{\partial f}{\partial x}(0, 0)}{\frac{\partial f}{\partial y}(0, 0)} = 0.$$

Exemplo 4.4.4 *Consideremos o sistema*

$$\begin{cases} x + y + xyz = 0, \\ x - y + xz + yz = 0. \end{cases}$$

Vamos agora estudar a existência de funções $X(z)$ e $Y(z)$ tais que $X(0) = Y(0) = 0$ e para z numa vizinhança da origem, $(x, y, z) = (X(z), Y(z), z)$ seja solução do sistema. Definido

$$f(x, y, z) = (f_1(x, y, z), f_2(x, y, z)) = (x + y + xyz, x - y + xz + yz),$$

temos

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x}(0, 0, 0) & \frac{\partial f_1}{\partial y}(0, 0, 0) \\ \frac{\partial f_2}{\partial x}(0, 0, 0) & \frac{\partial f_2}{\partial y}(0, 0, 0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Como $\det A \neq 0$ o teorema da função implícita garante a existência das funções $X(z)$ e $Y(z)$. Se estivermos interessados em calcular as derivadas de X e Y a expressão para estas será

$$\begin{bmatrix} X'(0) \\ Y'(0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial f_1}{\partial z}(0, 0, 0) \\ \frac{\partial f_2}{\partial z}(0, 0, 0) \end{bmatrix}.$$

Exercício 4.4.4 Mostre que a equação

$$\begin{bmatrix} \sin x + \sin y + \sin z \\ \sin x - \sin y + \sin^3 z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

admite uma solução da forma $(x, y, z) = (X(z), Y(z), z)$, para $|z|$ suficientemente pequeno, onde $X(z)$ e $Y(z)$ são funções convenientes que verificam $X(0) = Y(0) = 0$. Calcule

$$\frac{\partial X}{\partial z}(0).$$

Exercício 4.4.5 Mostre que a equação $\sin x + y = 0$ não tem solução $x(y)$ definida para todos valores de y . Porque que é que isto não contradiz o teorema da função implícita apesar de $\frac{d \sin x}{dx}|_{x=0} \neq 0$ e $\sin 0 + 0 = 0$?

O facto de não serem cumpridas as condições do teorema da função implícita não implica que na vizinhança de um ponto não exista solução (ou que ela não seja única) de uma determinada equação, tal como se pode verificar pelo exemplo seguinte:

Exemplo 4.4.5 Consideremos a equação

$$f(x, y) = x^3 - y^3 = 0.$$

É fácil verificar que $f(0, 0) = 0$. É também imediato que $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Assim não estamos nas condições do teorema da função implícita. No entanto a equação tem solução global, única e diferenciável $x = y$.

O seguinte exercício mostra que é possível generalizar ligeiramente o teorema da função implícita de modo a tratar casos semelhantes ao anterior.

Exercício 4.4.6 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função de classe C^1 , $g, h : \mathbb{R} \rightarrow \mathbb{R}$ funções contínuas bijectivas. Mostre que se $f(0, 0) = 0$, $h(0) = g(0) = 0$ e $\frac{\partial f}{\partial y}(0, 0) \neq 0$ então para x numa vizinhança da origem a equação

$$f(g(x), h(y)) = 0$$

pode ser unicamente resolvida em ordem a y , sendo a solução da forma $y = j(x)$ com $j : \mathbb{R} \rightarrow \mathbb{R}$ uma função real de variável real.

4.4.1 Exercícios suplementares

Exercício 4.4.7 Considere o sistema de equações

$$\begin{cases} \operatorname{sen}(x + y + z) = z^4 \\ x - y + z = \operatorname{sen}(x^4 + y^4 + z^4). \end{cases}$$

1. Prove que existem funções reais e diferenciáveis $g_x(z)$ e $g_y(z)$, definidas para $|z|$ suficientemente pequeno, tais que $g_x(0) = g_y(0) = 0$ e $(x, y, z) = (g_x(z), g_y(z), z)$ é solução do sistema.
2. Calcule $g'_x(0)$ e $g'_y(0)$.
3. Desenvolva g_x em série de Taylor até à terceira ordem.

Exercício 4.4.8 Seja $\alpha \in \mathbb{R}$ e considere as funções $f_\alpha : \mathbb{R}^3 \rightarrow \mathbb{R}$ definidas por

$$f_\alpha(x, y, z) = \alpha z \operatorname{ch}(x + y + z) - x^2 e^y$$

1. Determine para que valores de α a equação $f_\alpha(x, y, z) = 0$ define implicitamente, numa vizinhança da origem, uma função $z = \Psi_\alpha(x, y)$
2. Verifique que as funções Ψ_α têm um ponto de estacionaridade na origem, isto é, $\nabla \Psi_\alpha(0, 0) = 0$.

Exercício 4.4.9 Considere a equação

$$\|x\|^2 Ax = f(x, \epsilon)$$

onde $x \in \mathbb{R}^n$, $\epsilon \in \mathbb{R}^k$, A é uma matriz $n \times n$ não singular e $f : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^n$ uma função C^∞ . Mostre que se se verificar

$$\lim_{\epsilon \rightarrow 0} \left(\lim_{\|x\| \rightarrow 0} \frac{f(x, \epsilon)}{\|x\|^3} \right) = 0$$

a primeira equação define x como função diferenciável de ϵ para (x, ϵ) numa vizinhança de $(0, 0)$.

Exercício 4.4.10 (Função Implícita Topológica) Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, contínua. Suponha que para cada x fixo se tem

$$\lim_{y \rightarrow -\infty} f(x, y) = -\infty \quad \lim_{y \rightarrow +\infty} f(x, y) = +\infty.$$

1. Prove que existe pelo menos uma função $y(x)$ tal que $f(x, y(x)) = 0$ para todo o $x \in \mathbb{R}$.
2. Dê um exemplo em que a função $y(x)$ não seja única
3. Dê um exemplo em que a função $y(x)$ não seja contínua.

4.4.2 Sugestões para os exercícios

4.4.7

1. Defina $u(x, y, z) = \operatorname{sen}(x + y + z) - z^4$ e $v(x, y, z) = x - y + z - \operatorname{sen}(x^4 + y^4 + z^4)$. Observe que $u(0, 0, 0) = v(0, 0, 0) = 0$ e que

$$\begin{bmatrix} \frac{\partial u}{\partial x}(0, 0, 0) & \frac{\partial u}{\partial y}(0, 0, 0) \\ \frac{\partial v}{\partial x}(0, 0, 0) & \frac{\partial v}{\partial y}(0, 0, 0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

tem determinante não nulo. Portanto podemos aplicar o teorema da função implícita.

2. $g'_x(0) = -1$ e $g'_y(0) = 0$

3. $g_x(z) = -z + O(z^4)$.

4.4.9 Aplique o teorema da função implícita à equação ◀

$$Ax = \frac{f(x, \epsilon)}{\|x\|^2}$$

e mostre que este problema é equivalente ao original. ◀

4.4.10

1. Utilize o teorema de Bolzano.
2. Escolha f de modo a que para cada x tenha pelo menos dois zeros, $y_1(x)$ e $y_2(x)$, distintos.
3. Utilize a função da alínea anterior e defina

$$y(x) = \begin{cases} y_1(x) & \text{se } x > 0, \\ y_2(x) & \text{caso contrário.} \end{cases}$$
◀

Bibliografia

- [1] J. Campos Ferreira. *Introdução à Análise em \mathbb{R}^n* . AEIST, 1978.
- [2] F. R. Dias Agudo. *Lições de Análise Infinitesimal: I. Cálculo Diferencial em \mathbb{R}^n* . 1977.
- [3] Luís Torres Magalhães. *Integrais Múltiplos*. Texto Editora, Lisboa, 2^a edição, 1996.
- [4] Luís Torres Magalhães. *Álgebra Linear*. Texto Editora, Lisboa, 1985.
- [5] Luís Torres Magalhães. *Integrais em Variedades*. Texto Editora, Lisboa, 1994.
- [6] Luís Torres Magalhães. *Complementos de Cálculo Diferencial em \mathbb{R}^n* . AEIST, Lisboa, 1983.

Índice

- aberto, 8
- aproximações sucessivas, 54
- classe
 - C^1 , 14
 - C^∞ , 20
 - C^k , 19
- compacto, 10
- convexo, 53
- derivada, 12
 - direcciona, 13
 - dirigida, 13
 - de ordem superior à primeira, 21
 - parcial, 13
 - de ordem superior à primeira, 19
- diferenciabilidade, 12
- direcções singulares, 42
- fórmula
 - de Taylor, *ver* teorema de Taylor
- fechado, 9
- fecho, 10
- forma
 - de grau k , 41
 - quadrática
 - definida negativa, 39
 - definida positiva, 39
 - indefinida, 39
 - semidefinida negativa, 39
 - semidefinida positiva, 39
- fronteira, 10
- função
 - diferenciável, 7, 12
 - implícita, 63
 - inversa, 50
- gradiente, 16
- interior, 10
- inversa local, 54
- jacobiano, 57
- método de Newton, 54
- matriz
 - definida negativa, 39
 - definida positiva, 39
 - hessiana, 34, 39
 - indefinida, 39
 - jacobiana, 14
 - semidefinida negativa, 39
 - semidefinida positiva, 39
- norma, 52
 - de um vector de \mathbb{R}^n , 7
 - de uma aplicação linear, 52
 - de uma matriz, 52
- normas equivalentes, 10
- polinómio
 - de Taylor, 23
 - homogéneo, 22
- ponto
 - crítico, 30
 - de estacionaridade, 30
 - de sela, 31
- pré-imagem, 48
- resto da fórmula de Taylor, 23
- série
 - absolutamente convergente, 55
- Schwarz, 20
- sistema de estacionaridade, 27, 40
- Taylor, 23
- teorema
 - da função inversa, 53
 - de Lagrange, *ver* teorema do valor médio
 - de Schwarz, 20
 - de Taylor, 23
 - do valor médio
 - para funções vectoriais, 52
 - para funções escalares, 17
- Weierstrass, 29