



Departamento de Produção e Sistemas

Escola de Engenharia

Texto de Apoio

MÉTODOS NUMÉRICOS E OTIMIZAÇÃO NÃO LINEAR

Isabel Espírito Santo

Lino Costa

Conteúdo

1	Erros e números	1
1.1	Formato de vírgula flutuante	1
1.2	Erro de arredondamento	2
1.3	Erros inerentes aos dados	2
1.4	Erro absoluto e erro relativo	3
1.5	Algarismos significativos	4
1.6	Fórmula fundamental dos erros	5
1.7	Erros de truncatura	6
1.7.1	Métodos iterativos	6
1.7.2	Métodos de discretização	7
1.8	Exercícios	8
1.9	Soluções	10
2	Introdução ao MATLAB	11
2.1	Variáveis	11
2.2	Números	12
2.3	Funções, constantes e variáveis especiais	12
2.4	Operações básicas com matrizes	13
2.5	Vetores e matrizes	14
2.6	Aceder a partes de uma matriz	15
2.7	Matrizes especiais	16
2.8	Ficheiros m	16
2.9	Exercícios para resolver em MATLAB	18

3	Sistemas de equações lineares	21
3.1	Forma geral do problema	21
3.2	Existência e unicidade da solução	21
3.3	Métodos para a resolução de $Ax = b$	22
3.4	EGPP para resolução de sistemas lineares	22
3.5	EGPP para o cálculo do determinante de uma matriz	26
3.6	EGPP para o cálculo da inversa de uma matriz	26
3.7	Exercícios	29
3.8	Soluções	32
3.9	MATLAB	32
3.9.1	Método EGPP	32
3.9.2	Exercícios para resolver em MATLAB	34
4	Equações não lineares	37
4.1	Forma geral do problema	37
4.2	Solução de uma equação não linear	38
4.2.1	Métodos gráficos	38
4.2.2	Métodos iterativos	41
4.2.3	Método da secante	42
4.2.4	Método de Newton	46
4.2.5	Método da secante <i>versus</i> método de Newton	49
4.3	Solução de um sistema de equações não lineares	49
4.3.1	Forma geral do problema	49
4.3.2	Método de Newton	50
4.4	Exercícios	54
4.5	Soluções	59
4.6	MATLAB	60
4.6.1	fsolve	60
4.6.2	Exercícios para resolver em MATLAB	63
5	Polinômio interpolador de Newton	65
5.1	Erro da aproximação	66

5.2	Diferenças divididas	68
5.2.1	Definição	68
5.2.2	Propriedades das diferenças divididas	69
5.3	Polinómio interpolador de Newton	69
5.3.1	Interpolação direta	70
5.4	Erro de truncatura	71
5.5	Exercícios	75
5.6	Soluções	77
5.7	MATLAB	78
5.7.1	polyfit	78
5.7.2	polyval	78
5.7.3	Exercícios para resolver em MATLAB	80
6	Interpolação segmentada - 'spline'	83
6.1	Definição	84
6.2	'Spline' linear	85
6.2.1	Definição	85
6.2.2	Limite superior do erro de truncatura	85
6.3	'Spline' cúbica	86
6.3.1	Definição	86
6.3.2	'Spline' cúbica natural	87
6.3.3	'Spline' cúbica completa	87
6.3.4	Limite superior do erro de truncatura	88
6.4	Exercícios	92
6.5	Soluções	94
6.6	MATLAB	95
6.6.1	spline	95
6.6.2	Exercícios para resolver em MATLAB	98
7	Integração numérica	101
7.1	Forma geral do problema	101
7.2	Fórmulas simples de Newton-Cotes	102

7.2.1	Regra do retângulo	102
7.2.2	Regra do ponto médio	103
7.2.3	Regra do trapézio	103
7.2.4	Regra de Simpson	104
7.2.5	Regra dos três oitavos	105
7.2.6	Erros de truncatura	105
7.3	Fórmulas compostas	107
7.3.1	Fórmula composta do trapézio	107
7.3.2	Fórmula composta de Simpson	109
7.3.3	Fórmula composta dos três oitavos	110
7.4	Intervalos de amplitude não constante	112
7.5	Escolha da melhor fórmula	112
7.6	Exercícios	115
7.7	Soluções	119
7.8	MATLAB	120
7.8.1	trapz	120
7.8.2	integral	120
7.8.3	Exercícios para resolver em MATLAB	121
8	Aproximação dos mínimos quadrados	123
8.1	Forma geral do problema	123
8.2	Modelo polinomial	124
8.2.1	Polinômios ortogonais	125
8.3	Modelo linear não polinomial	127
8.3.1	Sistema das equações normais	127
8.4	Aproximação dos mínimos quadrados	132
8.5	Soluções	135
8.6	MATLAB	136
8.6.1	polyfit e polyval	136
8.6.2	lsqcurvefit	136
8.6.3	Exercícios para resolver em MATLAB	139

9	Otimização não linear sem restrições	141
9.1	Forma geral do problema	142
9.2	Classificação de mínimos e máximos	143
9.3	Mínimos <i>versus</i> máximos	144
10	Otimização unidimensional	147
10.1	Condições de otimalidade	147
10.2	Exercícios	149
10.3	Soluções	150
11	Método de DSC	151
11.1	Exercícios	160
11.2	Soluções	162
12	Otimização multidimensional sem restrições	163
12.1	Notação	164
12.2	Condições de otimalidade	166
12.3	Exercícios	168
12.4	Soluções	169
13	Métodos do gradiente	171
13.1	Técnicas de globalização	171
13.2	Procura unidimensional aproximada - critério de Armijo	172
13.3	Método de Newton	172
13.3.1	Propriedades do método de Newton	174
13.3.2	Limitações do método de Newton	174
13.3.3	Desvantagens do método de Newton	176
13.4	Método de segurança de Newton	177
13.5	Método quasi-Newton	180
13.5.1	Características da matriz H	181
13.5.2	Propriedades do método quasi-Newton	182
13.6	Exercícios	191
13.7	Soluções	194

13.8	MATLAB	195
13.8.1	fminunc	195
13.8.2	Exercícios para resolver em MATLAB	198
14	Método de Nelder-Mead	201
14.1	Exercícios	211
14.2	Soluções	212
14.3	MATLAB	213
14.3.1	fminsearch	213
14.3.2	Exercícios para resolver em MATLAB	214

Lista de Algoritmos

4.1	Método da secante	43
4.2	Método de Newton	47
4.3	Método de Newton para sistemas de equações não lineares	52
11.1	Método de Davies, Swann e Campey	155
13.2	Método do gradiente	173
13.3	Critério de Armijo	173
13.4	Método de segurança de Newton	177
13.5	Método quasi-Newton	183
14.6	Método de Nelder-Mead	208

Capítulo 1

Erros e números

A base da computação numérica são os números. É inevitável que se cometam erros nos cálculos que se vão efetuando ao longo de qualquer processo numérico, isto sem contar com os erros crassos. Basta, para isso, o facto de serem usados números. Assim, o primeiro tipo de erro que se comete, surge da representação do próprio número.

Quando se usam números que podem ser representados por uma sequência finita de dígitos, estes podem ser usados exatamente nos cálculos e, por isso, não se comete qualquer erro na sua representação (por exemplo, $1/8 = 0.125$). No entanto, outros há que originam uma sequência infinita de dígitos, pelo que a sua representação não é exata, sendo apenas possível utilizar-se um número limitado de dígitos (por exemplo $\sqrt{2} = 1.414213562\dots$). Assim sendo, comete-se um erro na sua representação.

1.1 Formato de vírgula flutuante

Uma das formas mais usuais de se representar um número é o formato de vírgula flutuante,

$$fl(x) = \pm 0.d_1d_2\dots d_k \times b^e,$$

em que d_k , $k = 1, \dots, k$ representam os k dígitos da mantissa, b é a base de representação (o mais usual é considerar-se $b = 10$) e e é o expoente.

A mantissa contém um número limitado de dígitos e esse número k define o comprimento da palavra no computador. Os cálculos podem ser feitos com precisão simples, quando se usam os k dígitos da mantissa ou com precisão dupla, caso se usem o dobro de dígitos da mantissa. No caso da precisão dupla, garante-se uma maior precisão nos cálculos.

1.2 Erro de arredondamento

Para se obter um número no formato de vírgula flutuante com uma mantissa de k dígitos pode recorrer-se a duas técnicas.

1. A **truncatura**, em que o formato da mantissa é o que está mais próximo de x e que se encontra entre x e 0.

Exemplo 1.1

$$x = \sqrt{2} = 1.414213562... = 0.1414213562... \times 10^1$$

$$se\ k = 8, fl(x) = 0.14142135 \times 10^1$$

$$se\ k = 6, fl(x) = 0.141421 \times 10^1$$

O erro de aproximação será, assim, a diferença $x - fl(x)$.

2. O **arredondamento**, em que o formato da mantissa é o que está mais próximo de x .

Exemplo 1.2

$$x = \sqrt{2} = 1.414213562... = 0.1414213562... \times 10^1$$

$$se\ k = 8, fl(x) = 0.14142136 \times 10^1$$

$$se\ k = 6, fl(x) = 0.141421 \times 10^1$$

O erro de aproximação será, assim, a diferença $x - fl(x)$.

1.3 Erros inerentes aos dados

Por vezes, surgem erros por não ser possível atribuir, com exactidão, valores numéricos corretos aos dados obtidos por leitura experimental, proveniente de um equipamento. Por exemplo, medir um segmento de reta com uma régua, efetuar uma pesagem com uma balança... estes são os erros que são inerentes aos próprios dados.

1.4 Erro absoluto e erro relativo

A primeira questão que surge quando se fala de erros, quer de arredondamento, quer inerentes aos dados, é: como se avalia a proximidade do valor aproximado em relação ao valor exato?

Se x for o valor aproximado usado nos cálculos e x^* o valor exato, então

$$\text{erro absoluto} = x^* - x.$$

Esta diferença pode ser positiva, negativa ou nula.

O que acontece, em geral, é que o valor exato x^* é desconhecido e por isso também o erro. No entanto, é possível estabelecer um limite máximo para o erro que se comete, o que permite ter uma ideia da sua grandeza. Este conhecimento é, na maior parte das vezes, suficiente para se conhecer a exactidão do valor em causa.

Define-se, assim, como limite superior do erro absoluto a quantidade não negativa, δ_x , tal que

$$|x - x^*| \leq \delta_x$$

ou

$$x - \delta_x \leq x^* \leq x + \delta_x.$$

Apesar de útil, o erro absoluto, por si só, não é suficiente para que se perceba a exactidão do valor aproximado. Um certo erro absoluto pode ser considerado pequeno numas situações e grande noutras. Tudo depende do valor exato, x^* , que se está a medir. Para ultrapassar este problema, define-se o erro relativo, r_x , que depende, não apenas do erro absoluto, mas também do valor de x^* :

$$r_x = \frac{|x^* - x|}{|x^*|}.$$

Se δ_x for pequeno quando comparado com x^* , então define-se como limite superior do erro relativo a quantidade

$$r_x \simeq \frac{|x^* - x|}{|x|} \leq \frac{\delta_x}{|x|}.$$

Esta aproximação é muito importante e é a mais usada na prática porque, mais uma vez, o valor de x^* é quase sempre desconhecido.

1.5 Algarismos significativos

Quando se usa um valor aproximado x nem todos os algarismos são de confiança. Apenas os são os denominados algarismos significativos.

Os algarismos significativos são os que se encontram na parte de confiança do número e são

- diferentes de zero;
- iguais a zero mas não estão no número a indicar a posição do ponto decimal.

Para se determinar a parte de confiança de x , usa-se o limite superior do erro absoluto δ_x . Considera-se o majorante do erro imediatamente superior ao calculado ou conhecido e que comece pelo algarismo 5, isto é, que esteja na forma 0.5×10^n . Coloca-se x a concordar, isto é, na forma $\dots \times 10^n$. A parte de confiança de x inclui os algarismos que se encontram à esquerda da posição que é equivalente à posição onde se encontra o algarismo 5 em δ_x .

Exemplo 1.3 Seja $x = 50.1234 \times 10^0$ e $\delta_x = 0.5 \times 10^{-3}$. Então

$$\begin{aligned} x &= 50123.\textcolor{red}{4} \times 10^{-3} \\ \delta_x &= 0.\textcolor{red}{5} \times 10^{-3} \end{aligned}$$

ou

$$\begin{aligned} x &= 50.123\textcolor{red}{4} \times 10^0 \\ \delta_x &= 0.000\textcolor{red}{5} \times 10^0. \end{aligned}$$

Logo, x tem 5 algarismos significativos.

Exemplo 1.4 Seja $x = 13.5550 \times 10^0$. Então

$$\begin{aligned} x &= 13.5550 \times 10^0 \\ \delta_x &= 0.000\textcolor{red}{5} \times 10^0. \end{aligned}$$

Logo, x tem 5 algarismos significativos se $\delta_x = 0.5 \times 10^{-3}$, ou

$$\begin{aligned} x &= 13.\text{ }5550 \times 10^0 \\ \delta_x &= 0.\text{ }0000\textcolor{red}{5} \times 10^0. \end{aligned}$$

Logo, x tem 6 algarismos significativos se $\delta_x = 0.5 \times 10^{-4}$.

1.6 Fórmula fundamental dos erros

Quando se efetua um certo número de operações, há necessidade também de se avaliar o erro cometido. Para isso usa-se a fórmula fundamental dos erros, que permite calcular o limite superior do erro absoluto de uma expressão que envolve várias operações e várias variáveis, $f(x, y, z, \dots)$. Conhecidos os valores aproximados das variáveis envolvidas na operação x, y, z, \dots e os correspondentes limites superiores dos erros absolutos $\delta_x, \delta_y, \delta_z, \dots$, calculam-se os intervalos de valores possíveis para as variáveis,

$$I = \begin{cases} x - \delta_x \leq x^* \leq x + \delta_x \\ y - \delta_y \leq y^* \leq y + \delta_y \\ z - \delta_z \leq z^* \leq z + \delta_z \\ \dots \end{cases}$$

De seguida, calculam-se os majorantes das derivadas parciais de f em ordem a cada uma das variáveis x, y, z, \dots :

$$\left| \left[\frac{\partial f}{\partial x} \right]_I \right| \leq M_x, \left| \left[\frac{\partial f}{\partial y} \right]_I \right| \leq M_y, \left| \left[\frac{\partial f}{\partial z} \right]_I \right| \leq M_z, \dots$$

O limite superior do erro absoluto em f é dado por

$$\delta_f \leq M_x \delta_x + M_y \delta_y + M_z \delta_z + \dots \quad (1.1)$$

Exemplo 1.5 *Seja $x = 1.1$ com 2 algarismos significativos, $y = 2.04$ com 3 algarismos significativos e $z = 0.5$ com 1 algarismo significativo. Quantos algarismos significativos tem o resultado de $f(x, y, z) = -x + y + \sin z$?*

$$\begin{cases} x = 1.1 & \Rightarrow \delta_x = 0.05 \\ y = 2.04 & \Rightarrow \delta_y = 0.005 \\ z = 0.5 & \Rightarrow \delta_z = 0.05 \end{cases} \quad e \quad I = \begin{cases} 1.05 & \leq x^* < 1.15 \\ 2.035 & \leq y^* < 2.045 \\ 0.45 & \leq z^* < 0.55. \end{cases}$$

Sendo que

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

vem

$$\frac{\partial f}{\partial x} = -1, \quad \frac{\partial f}{\partial y} = 2y, \quad \frac{\partial f}{\partial z} = \cos(z),$$

ou seja,

$$M_x = 1, \quad M_y = 2 \times 2.045 = 4.09, \quad M_z = \cos(0.45) = 0.9004471024.$$

Assim, pela fórmula fundamental dos erros (1.1) obtém-se

$$\begin{aligned} \delta_f &\leq 1 \times 0.05 + 4.09 \times 0.005 + 0.9004471024 \times 0.05 \\ &= 0.115423551 \\ &\leq 0.5 \times 10^0. \end{aligned}$$

Como

$$f(1.1, 2.04, 0.5) = -1.1 + 2.04^2 + \sin(0.5) = 3.541025539,$$

$$f = 3.541025539 \times 10^0$$

$$\delta_f = 0.5 \times 10^0$$

$\Rightarrow 1$ algarismo significativo.

1.7 Erros de truncatura

Quando se resolvem problemas matemáticos recorrendo a um certo tipo de Métodos Numéricos, podem cometer-se erros de truncatura. Estes acontecem quando se utilizam métodos iterativos ou métodos de discretização.

1.7.1 Métodos iterativos

Um método iterativo é definido por uma equação iterativa, ou seja,

$$x_{k+1} = F(x_k),$$

a partir da qual é gerada uma sucessão de aproximações à solução exata do problema, x^* . k indica o índice da iteração e x_k representa a aproximação à solução numa determinada iteração k .

Um método iterativo gera uma sucessão de aproximações, que é iniciada numa aproximação inicial x_1 :

$$x_1 \curvearrowright x_2 \curvearrowright x_3 \curvearrowright \dots \curvearrowright x_n \curvearrowright x_{n+1} \dots$$

Se o processo iterativo estiver a convergir, esta sucessão converge para x^* .

Ao usar-se um método iterativo, a solução x^* é atingida ao fim de um número infinito de operações. Por isso, face aos recursos limitados em termos de tempo e de memória, o processo

iterativo tem de ser terminado ao fim de um número finito de iterações. Para isso usa-se um critério de paragem que garante que a aproximação calculada, x_{n+1} , está suficientemente próxima de x^* .

O erro de truncatura é dado pela diferença

$$x^* - x_{n+1}.$$

1.7.2 Métodos de discretização

Este tipo de métodos transforma um problema matemático que envolve conceitos de natureza contínua, tais como a diferenciação e a integração, num problema discreto, que envolve apenas operações algébricas. A este processo chama-se discretização.

Exemplo 1.6 *Dado o integral*

$$I = \int_1^{1.2} \sqrt{x} \, dx,$$

este pode ser transformado em

$$I \approx S(0.05) = \frac{0.05}{3} [f(1) + 4f(1.05) + 2f(1.1) + 4f(1.15) + f(1.2)]$$

através da fórmula composta de Simpson.

O erro de truncatura, neste caso, é dado pela diferença $I - S(0.05)$.

1.8 Exercícios

1. Com base no limite superior do erro absoluto no cálculo da expressão

$$f(\pi, \sqrt{3}, \sqrt{2}) = \frac{2\pi\sqrt{3}}{\pi^2 + \sqrt{2}},$$

e sabendo que são usados os seguintes valores aproximados

$$\pi = 3.1416, \quad \sqrt{3} = 1.732 \quad e \quad \sqrt{2} = 1.4142,$$

quantos algarismos significativos tem o valor calculado de f ?

2. O perímetro P de um triângulo rectângulo de hipotenusa h e com um dos ângulos agudos α , pode ser dado pela expressão

$$P = (\sin(\alpha) + \cos(\alpha)) h.$$

Supondo que $\alpha = 0.34$ rad, qual o erro absoluto com que se deve medir h , de valor aproximado 16.7 m, para que o erro absoluto em P não exceda 0.5?

3. Uma corrente eléctrica atravessa uma resistência (R) de 20Ω . A resistência foi medida com um erro relativo que não excede 0.01. A intensidade da corrente (I) é 3.00 ± 0.01 A. Sabendo que a tensão da corrente é dada por $V = RI$, determine um limite superior do erro absoluto no calculo da tensão da corrente. Quantos algarismos significativos garante para o valor calculado da tensão?

4. Tomando $\overline{\sqrt{2}} = 1.41$, $\overline{\sqrt{3}} = 1.73$ e $\overline{\pi} = 3.14$:

a) Calcule um limite superior do módulo do erro absoluto que se comete no cálculo de $N = \frac{\sqrt{3} - \sqrt{2}}{\pi}$.

b) Com quantas casas decimais exactas devem ser tomados os valores aproximados de $\sqrt{2}$, $\sqrt{3}$ e π para obter uma aproximação de N com 3 casas decimais exactas?

5. Seja

$$A = \frac{3\sqrt{3}a^2}{2}$$

a área de um hexágono regular de lado a . Seja 1m o valor aproximado para o lado do hexágono. Considerando um valor aproximado de $\sqrt{3}$ com quatro algarismos significativos, com que aproximação se deve medir o lado de modo a que o limite superior do erro absoluto no cálculo da área não exceda 100cm^2 ?

6. Pretende-se calcular a área de um círculo, de raio aproximadamente igual a 25 cm, com erro absoluto que em módulo não excede 0.5 cm^2 . Com que aproximação se deve medir o raio do círculo e quantos algarismos significativos se devem usar no valor aproximado de π ?
7. O rendimento η de um transformador depende da potência de entrada z , da potência de saída a e da perda de potência b , pelas relações:

$$\eta = \frac{a}{z} = \frac{a}{a + b}$$

Podem medir-se z e a a menos de 1%, enquanto que o erro na medida de b pode ser de 20%, sendo η cerca de 0.95. Qual das relações usaria para a determinação de η ?

1.9 Soluções

1. Três algarismos significativos.
2. $\delta_h \leq 0.349632$.
3. Um algarismo significativo.
4. a) $\delta_f \leq 0.5 \times 10^{-2}$.
b) São necessárias pelo menos três casas decimais.
5. $\delta_a \leq 0.00177603$.
6. $\delta_r \leq 0.000617928008$. Para aproximar π devem usar-se quatro.
7. A segunda fórmula é a mais adequada para determinar η .

Capítulo 2

Introdução ao MATLAB

O MATLAB é um pacote de *software* comercial cujo nome deriva de MATrix LABoratory. Começou por ser um programa iterativo para cálculos com matrizes, tal como o nome indica, e transformou-se numa linguagem matemática de alto nível. O seu desenvolvimento permite agora, por exemplo, a resolução de equações diferenciais, o desenho de gráficos a duas e três dimensões e a resolução problemas de otimização. Possuindo o MATLAB uma linguagem de programação, qualquer algoritmo pode ser implementado em MATLAB.

Pode ser obtida ajuda no MATLAB sobre qualquer função digitando apenas na janela de comandos a palavra 'help' seguida do nome da função sobre a qual pretendemos obter informações.

Os comentários em MATLAB podem ser introduzidos precedidos do sinal %.

2.1 Variáveis

O MATLAB não requer nenhum tipo de declaração de variáveis ou dimensões. Quando um novo nome de variável é encontrado, o MATLAB automaticamente cria a variável e aloca a memória necessária. Se a variável já existir, o seu conteúdo é substituído pelo novo e, caso seja necessário, faz uma nova alocação de memória. O nome das variáveis é 'case sensitive', ou seja $a \neq A$, por exemplo. Os nomes das variáveis devem começar por uma letra, seguida de outras letras, dígitos ou **underscore**. Outros caracteres nunca devem ser usados. Não devem ser usados também nomes de variáveis que já estejam a ser usados pelo MATLAB, como por exemplo `exp`, `log`, As variáveis, por defeito, são matrizes, sendo que podem ser vetores

ou escalares, já que estes são casos particulares de matrizes.

Exemplo 2.1 *Introdução de uma variável*

```
>>exemplo_1=2
```

Este comando cria uma matriz 1×1 , com o nome 'exemplo_1' e guarda o valor '2' no seu único elemento.

2.2 Números

O MATLAB usa a notação decimal convencional nos números, com um ponto decimal opcional, podendo o número ser precedido de mais ou menos. A notação científica usa a letra 'e' para especificar potência de 10. Os números imaginários usam 'i' ou 'j' como sufixo. Todos os números são armazenados internamente no formato especificado pela norma IEEE floating-point, que tem uma precisão finita de cerca de 16 algarismos significativos e uma gama finita de 10^{-308} a 10^{308} , aproximadamente.

O comando `format` define a forma como os números são apresentados, não tendo qualquer influência na forma como são armazenados. O formato que está definido por defeito é o `format short`, que apresenta os números com cinco dígitos e um ponto decimal fixo. O `format long` apresenta um formato de 15 dígitos e um ponto decimal fixo. Existem outros formatos disponíveis em MATLAB.

2.3 Funções, constantes e variáveis especiais

Apresenta-se de seguida uma lista de algumas funções, constantes e variáveis das mais usadas. Esta lista não pretende ser exaustiva, mas apenas um suporte para um utilizador inexperiente.

`ans` - variável *built in* criada quando não é atribuído um nome à variável

`pi` - valor de π (3.14159265...)

`i, j` - unidade imaginária (como sufixo)

`inf` - infinito

`NaN` - 'Not-a-Number'

`exp(x)` - exponencial de x (base neperiana)

`log(x)` - logaritmo natural de x

`log10(x)` - logaritmo na base 10 de x

`sqrt(x)` - raiz quadrada de x

`abs(x)` - valor absoluto de x

`rem(x,y)` - resto da divisão de x por y

`round(x)` - converte x para o valor inteiro mais próximo

`sin(x)` - seno de x

`cos(x)` - cosseno de x

`tan(x)` - tangente de x

`cot(x)` - co-tangente de x

2.4 Operações básicas com matrizes

Apresentam-se de seguida os operadores matemáticos mais usados, bem como algumas operações básicas com matrizes. As operações são sempre efetuadas segundo as regras de operações algébricas entre matrizes. Se se pretender efetuar uma operação para cada ponto da matriz, deve anteceder-se a operação em questão de '.

`+` adição

`-` subtração

`*` multiplicação

`/` divisão

`^` potenciação

`'` transposta de uma matriz

2.5 Vetores e matrizes

Uma das grandes vantagens do MATLAB é que considera a matriz como elemento básico, sendo que a sua dimensão não tem de ser definida *a priori*. Por esse motivo, permite a manipulação e criação de matrizes de uma forma muito mais rápida e intuitiva que em outras linguagens.

Um vetor ou uma matriz deve ser introduzido entre parêntesis retos, []. Se a variável for um escalar, não é necessário colocar parêntesis.

Para se introduzir um vetor linha, os elementos devem ser separados por um espaço ou por uma vírgula.

Exemplo 2.2 *Introduzir o vetor*

(1 2 3).

```
>> u=[1 2 3]
```

ou

```
>> u=[1,2,3]
```

Um vetor linha cujos os elementos tenham uma distância constante entre si, podem ser criados de uma forma mais fácil usando ':'. Indica-se o primeiro elemento do vetor, de seguida o espaçamento entre os elementos e o elemento final do vetor (ou valor máximo até onde se pretende que os elementos do vetor cheguem). Se o espaçamento for a unidade, pode suprimir-se o espaçamento.

Exemplo 2.3 *Criar o vetor*

(1 3 5 7).

```
>>i=1:2:7
```

Para se introduzir um vetor coluna, os elementos devem ser separados por ponto e vírgula, ou cria-se através de um vetor linha fazendo a sua transposta (usando ').

Exemplo 2.4 *Introduzir o vetor*

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$


```
>> v=[1;2;3]
```

ou

```
>> v=[1 2 3]'
```

Para introduzir uma matriz combinam-se as regras anteriores.

Exemplo 2.5 *Introduzir a matriz*

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

```
>> A=[1 2 3;4 5 6;7 8 9]
```

ou

```
A=[1 2 3
    4 5 6
    7 8 9]
```

Para se obter a dimensão de uma matriz, usa-se o comando **size**. Esta dimensão é devolvida como um vetor de dois elementos, em que o primeiro representa o número de linhas e o segundo o número de colunas. Para se obter o número de elementos num vetor usa-se o comando **length**. Este comando aplicado a uma matriz dá o valor máximo entre o número de linhas e o número de colunas.

2.6 Aceder a partes de uma matriz

Pode aceder-se a partes específicas de uma determinada matriz, bastando para isso indicar-se as coordenadas dos elementos aos quais se pretende aceder. As posições indicadas à esquerda da vírgula indicam as coordenadas das linhas, à direita da vírgula as coordenadas das colunas. Se quisermos incluir todas as linhas (ou todas as colunas) de uma matriz, de uma forma simplificada basta indicar ':' na respetiva posição.

Exemplo 2.6 *Considerando a matriz A definida no Exemplo 2.5, se quisermos aceder apenas aos elementos 1,2,4 e 5, criando a submatriz $Aa = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$, basta usar o comando*

```
>>Aa=A(1:2,1:2)
```

2.7 Matrizes especiais

O MATLAB gera algumas matrizes de forma automática. A dimensão da matriz é indicada entre parêntesis, sendo que (n, m) cria uma matriz com n linhas e m colunas e (n) cria uma matriz quadrada $n \times n$.

`>>rand(m,n)` ou `>>rand(n)` - gera uma matriz com números aleatórios entre zero e um.

`>>eye(n,m)` ou `>>eye(n)` - gera uma matriz com os elementos da diagonal principal iguais a um e os restantes iguais a zero. No caso da matriz ser quadrada, o comando `eye` gera a matriz identidade.

`>>zeros(m,n)` ou `>>zeros(n)` - gera uma matriz com todos os elementos iguais a zero.

`>>ones(m,n)` ou `>>ones(n)` - gera uma matriz com todos os elementos iguais a um.

`>>magic(n)` - gera uma matriz em que a soma dos elementos de cada linha e a soma dos elementos de cada coluna é sempre igual.

2.8 Ficheiros m

Os ficheiros usados em MATLAB devem ter a extensão `m`. Os mais usados são os que definem funções, com parâmetros de entrada e de saída, e os do tipo `script`, que executam um conjunto de comandos sequencialmente, pela ordem que aparecem escritos.

Para se criar um novo ficheiro, a forma mais fácil é usar o editor do MATLAB, que no caso das funções já aparece por defeito com a estrutura definida. Um ficheiro do tipo função começa sempre pela palavra `function` na primeira linha. De seguida define-se a estrutura da função. Primeiro aparecem os parâmetros de saída (se for mais que um terão de aparecer entre parêntesis retos, `[]`), seguidos do sinal `'='`, o nome da função, que deve coincidir com o nome do ficheiro e os parâmetros de entrada, que aparecem entre parêntesis curvos, `()`. Nas linhas seguintes aparecem as relações funcionais entre os parâmetros de entrada e os parâmetros de saída.

Exemplo 2.7 *Para se criar um ficheiro `.m` do tipo função que contenha o sistema*

$$\begin{cases} f_1 = x_1 + x_2 \\ f_2 = \frac{\sin(x_2)}{x_1} \end{cases},$$

cria-se o ficheiro

```
function f=fun(x)
f(1)=x(1)+x(2);
f(2)=sin(x(2))/x(1);
```

Exemplo 2.8 *Criar um programa que faça a soma de $x=2$ com $y=3$.*

O ficheiro a criar é, neste caso, do tipo `script`, que pode gravar-se como `soma.m`.

```
x=2;
y=3;
z=x+y
```

Basta executar este ficheiro ou então digitar na janela de comandos `>>soma`. O ponto e vírgula serve para suprimir a saída na janela de comandos (o comando é executado mas não é exibido).

2.9 Exercícios para resolver em MATLAB

1. Escreva os seguintes vetores ou matrizes:

a) $u = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix};$

b) $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix};$

c) um vetor linha com os números naturais menores ou iguais a 10;

d) um vetor linha com os números pares naturais menores ou iguais a 12;

e) $A = \begin{pmatrix} 2 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$

2. Com base na matriz A da pergunta anterior, verifique o que resulta dos seguintes comandos:

a) $B=A(2:3,1:2);$

b) $C=A(:,1:2);$

c) $D=[A;4 \ 4 \ 4];$

d) $E=D([2 \ 4],:);$

e) $F=[0:3:9;2:2:8;5:5:20];$

f) $A([1 \ 2],1)=2*A([1 \ 2],1);$

g) $A([1 \ 3],1)=A([3,1],1).$

3. Gere as seguintes matrizes:

a) a matriz identidade 5×5 ;

b) uma matriz 3×3 com elementos aleatórios entre 0 e 1;

c) uma matriz 4×3 com elementos aleatórios entre -1 e 1;

d) uma matriz nula 2×3 ;

e) uma matriz 2×2 com todos os elementos iguais a 1;

f) uma matriz 10×10 com todos os elementos iguais a 10;

- g) uma matriz com os elementos da diagonal da matriz A da pergunta 1 e os restantes iguais a zero.

4. Considere a matriz $A = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 4 & 1 \\ 2 & 2 & 1 \end{pmatrix}$, a matriz B que consiste numa matriz de dimensão

3×3 , com todos os elementos iguais a um, o vetor $a = \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}$ e o vetor $b = \begin{pmatrix} 0 & 3 & 5 \end{pmatrix}$. Efetue as seguintes operações.

- a) $A + B$;
 - b) $A \times B$;
 - c) o produto de cada um dos elementos de a por b ;
 - d) o produto de cada um dos elementos de A por B .
5. Escreva um programa que lê dois números e escreve a sua soma e o seu produto.
6. Escreva um programa que lê uma sequência de n números e escreve a sua soma e o seu produto.
7. Escreva um programa que lê uma sequência de n números e escreve o maior deles.

Capítulo 3

Sistemas de equações lineares

3.1 Forma geral do problema

O problema que se aborda neste capítulo é um sistema de n equações e n variáveis que tem a forma geral

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n &= b_3 \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{cases}$$

ou

$$Ax = b,$$

em que $A_{n \times n}$ é a matriz dos coeficientes do sistema com n linhas e n colunas, $x = (x_1 \ x_2 \ x_3 \ \dots, x_n)^T$ é o vetor solução, $b = (b_1 \ b_2 \ b_3 \ \dots, b_n)^T$ é o vetor dos termos independentes e $(A|b)$ é a matriz ampliada do sistema.

3.2 Existência e unicidade da solução

Um sistema de equações lineares, em termos do número de soluções, pode ser classificado como

- **possível e determinado**, quando tem uma solução única,
- **possível e indeterminado** quando tem um número infinito de soluções e

- **impossível** quando não tem solução.

A existência e unicidade da solução do sistema linear depende da característica das matrizes A e $(A|b)$.

A característica de uma matriz A , $c(A)$, é dada pelo número de linhas ou colunas linearmente independentes.

Existe uma relação direta entre $c(A)$, o determinante de A , $\det(A)$, e a existência da inversa de A , A^{-1} .

- Se $c(A) = n$, então $\det(A) \neq 0$, A^{-1} existe e o sistema é **possível e determinado**, ou seja, tem uma solução única.
- Se $c(A) < n$, então $\det(A) = 0$, A^{-1} não existe e
 - se $c(A) = c(A|b)$ o sistema é **possível e indeterminado**, ou seja, tem uma infinidade de soluções,
 - se $c(A) < c(A|b)$ o sistema é **impossível**, ou seja, não tem solução.

3.3 Métodos para a resolução de $Ax = b$

Há vários métodos para a resolução de sistemas de equações lineares, no entanto, de um modo geral, podem agrupar-se em dois tipos:

- **Métodos diretos**

São aplicados a sistemas de pequena ou média dimensão e a solução exata é obtida num número finito de operações.

- **Métodos iterativos**

São aplicados a sistemas de grande dimensão e a solução exata só se obtém ao fim de um número infinito de operações.

3.4 Método directo de eliminação de Gauss com pivotagem parcial para resolução de sistemas lineares

Um dos métodos mais usados para resolver sistemas de equações lineares é o método de eliminação de Gauss com pivotagem parcial (EGPP). Este distingue-se do método de eliminação de

Gauss (EG) por permitir trocas de linhas, que garantem que o pivot tenha um valor numérico não superior a um. O sucesso deste método deve-se ao facto de ser direto, simples e numericamente estável, ao contrário, por exemplo, do EG. Ao resolver-se um sistema por EGPP, uma vez que se garante que o multiplicador nunca é superior a um, os erros de arredondamento nunca serão propagados, podendo ser, pelo contrário, minimizados. Atente-se ao exemplo 3.1

Exemplo 3.1 Considere o sistema linear, cuja solução é $(1, 1)^T$

$$\begin{cases} 10^{-20}x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases}$$

- Se for resolvido por EG,

$$\begin{pmatrix} 10^{-20} & 1 & | & 1 \\ 1 & 1 & | & 2 \end{pmatrix} \xrightarrow{m_{21}=-10^{20}} \begin{pmatrix} 10^{-20} & 1 & | & 1 \\ 0 & -10^{20} & | & -10^{20} \end{pmatrix}$$

$$\rightarrow \begin{cases} 10^{-20}x_1 + x_2 = 1 \\ -10^{20}x_2 = -10^{20} \end{cases} \rightarrow \begin{cases} x_1 = 0 \\ x_2 = 1 \end{cases}$$

- Se for resolvido por EGPP,

$$\begin{pmatrix} 10^{-20} & 1 & | & 1 \\ 1 & 1 & | & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & | & 2 \\ 10^{-20} & 1 & | & 1 \end{pmatrix} \xrightarrow{m_{21}=-10^{-20}} \begin{pmatrix} 1 & 1 & | & 2 \\ 0 & 1 & | & 1 \end{pmatrix}$$

$$\rightarrow \begin{cases} x_1 + x_2 = 2 \\ x_2 = 1 \end{cases} \rightarrow \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$$

Com base neste exemplo pode verificar-se que o método EGPP é numericamente estável, ao contrário do EG, que pode tornar-se numericamente instável.

O método de EGPP para resolver um sistema de equações lineares, pode dividir-se em 2 passos:

1. transformar $Ax = b$ em $Ux = c$ usando EGPP

em que U é uma matriz triangular superior. Os sistemas $Ax = b$ e $Ux = c$ são equivalentes, isto é, têm a a mesma solução.

Há operações elementares de matrizes que se podem efetuar de forma a transformar A em U :

- (a) trocar duas linhas paralelas,

- (b) multiplicar uma linha por um escalar diferente de zero,
- (c) substituir uma linha pela que dela é obtida adicionando o produto de outra linha paralela por um escalar.

2. resolver $Ux = c$ por substituição inversa

Exemplo 3.2 Considere a matriz ampliada do sistema

$$\left(\begin{array}{ccc|c} 0.3 & -0.2 & 10 & 71.4 \\ 0.1 & 7 & -0.3 & -19.3 \\ 3 & -0.1 & -0.2 & 7.85 \end{array} \right)$$

Este sistema tem dimensão $n = 3$, o que significa que o primeiro passo de resolução, isto é, a transformação de A em U , tem $n - 1 = 2$ etapas.

- Etapa 1*
- Colocar 'pivot' na posição $(1,1)$. O 'pivot' é o valor de maior módulo de entre todos os valores da primeira coluna.
 - Trocar a linhas 1 com a linha 3.

$$\left(\begin{array}{ccc|c} \mathbf{3} & -0.1 & -0.2 & 7.85 \\ 0.1 & 7 & -0.3 & -19.3 \\ 0.3 & -0.2 & 10 & 71.4 \end{array} \right) \quad \Leftrightarrow \quad \text{com } I_{1,3} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$I_{1,3}(A|b)$

Para reduzir a zero os elementos 0.1 e 0.3, calculam-se os escalares - denominados multiplicadores:

- $m_{21} = -\frac{0.1}{\mathbf{3}} = -0.033333$
- $m_{31} = -\frac{0.3}{\mathbf{3}} = -0.1$

Nota: não esquecer que $|\text{multiplicador}| \leq 1$ conserva a estabilidade numérica.

$m_{21} \times (\text{linha } 1) + \text{linha } 2 \quad \text{e} \quad m_{31} \times (\text{linha } 1) + \text{linha } 3 \Rightarrow$

$$\left(\begin{array}{ccc|c} \mathbf{3} & -0.1 & -0.2 & 7.85 \\ 0 & 7.003333 & -0.293333 & -19.561664 \\ 0 & -0.19 & 10.02 & 70.615 \end{array} \right) \quad \Leftrightarrow \quad \text{com } J_1 = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & 0 & 1 \end{pmatrix}.$$

$J_1 I_{1,3}(A|b)$

Etapa 2 colocar o 'pivot' na posição (2,2). O pivot é o elemento de maior módulo na coluna dois, abaixo da linha 1.

- não é necessário trocarem-se linhas,
- para reduzir a zero o elemento -0.19 , calcula-se o multiplicador

$$m_{32} = \frac{-0.19}{\mathbf{7.003333}} = 0.027130$$
- $m_{32} \times (\text{linha } 2) + (\text{linha } 3) \Rightarrow$

$$\left(\begin{array}{ccc|c} 3 & -0.1 & -0.2 & 7.85 \\ 0 & \mathbf{7.003333} & -0.293333 & -19.561664 \\ 0 & 0 & 10.012042 & 70.084292 \end{array} \right) \xLeftrightarrow{J_2 J_1 I_{1,3}(A|b)} \text{ com } J_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & m_{32} & 1 \end{pmatrix}.$$

- A matriz ampliada está, nesta fase, na forma $(U|c)$ que corresponde ao sistema $Ux = c$.
- Os cálculos efetuados sobre a matriz ampliada correspondem ao procedimento:

$$J_2 J_1 I_{1,3}(A|b) = (U|c).$$

Para resolver o sistema $Ux = c$ por substituição inversa:

$$\left\{ \begin{array}{rcl} 3x_1 & -0.1x_2 & -0.2x_3 = 7.85 \\ & 7.003333x_2 & -0.293333x_3 = -19.561664 \\ & & 10.012042x_3 = 70.084292 \end{array} \right.$$

$$x_3 = \frac{70.084292}{10.012042} = 7,$$

$$x_2 = \frac{-19.561664 + 0.293333(7)}{7.003333} = -2.5,$$

$$x_1 = \frac{7.85 + 0.2(7) + 0.1(-2.5)}{3} = 3.$$

A solução do sistema é

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2.5 \\ 7 \end{pmatrix}.$$

3.5 EGPP para o cálculo do determinante de uma matriz

Para calcular o determinante de uma matriz quadrada pode usar-se o método de eliminação de Gauss com pivotagem parcial. Para isso usa-se EGPP para transformar A na matriz triangular superior U . De seguida, usam-se as propriedades do determinante. Atente-se no Exemplo 3.3.

Exemplo 3.3 *Calcular o determinante de*

$$A = \begin{pmatrix} 0.3 & -0.2 & 10 \\ 0.1 & 7 & -0.3 \\ 3 & -0.1 & -0.2 \end{pmatrix}$$

O procedimento para transformar A em U foi (Exemplo 3.2)

$$J_2 J_1 I_{1,3} A = U.$$

Então,

$$\begin{aligned} \det(J_2 J_1 I_{1,3} A) &= \det(U) \Leftrightarrow \\ \Leftrightarrow \det(J_1) \times \det(J_1) \times \det(I_{1,3}) \times \det(A) &= \det(U) \Leftrightarrow \\ \Leftrightarrow 1 \times 1 \times (-1) \det(A) &= 3 \times 7.003333 \times 10.012042 \Leftrightarrow \\ \Leftrightarrow \det(A) &= -210.352994. \end{aligned}$$

Como regra, estabelece-se que

$$\det(A) = (-1)^r \times U_{ii}, \quad i = 1 \dots, n,$$

sendo r o número de trocas de linhas efetuadas durante o processo de EGPP.

3.6 EGPP para o cálculo da inversa de uma matriz

Dada uma matriz quadrada $A_{n \times n}$, se $\det(A) \neq 0$, então existe uma matriz inversa de A , A^{-1} , tal que $A \times A^{-1} = A^{-1} \times A = I$, sendo I a matriz identidade. Quando existe a inversa de uma matriz A , esta matriz A diz-se não singular.

Quando $\det(A) = 0$, a matriz A não tem inversa e diz-se singular.

Para calcular a inversa de uma matriz A pode usar-se o método de eliminação de Gauss com pivotagem parcial. As n colunas de A^{-1} são as soluções dos n sistemas

$$AX = I \quad \Rightarrow \quad X = A^{-1}.$$

Usa-se EGPP para transformar A na matriz triangular superior U , fazendo as mesmas operações em simultâneo sobre a matriz I , ou seja, transforma-se a matriz ampliada $(A|I)$ em $(U|J)$.

De seguida, resolvem-se os n sistemas, em que os termos independentes são as colunas de J , por substituição inversa.

Atente-se no Exemplo 3.4.

Exemplo 3.4 Calcular A^{-1} sendo

$$A = \begin{pmatrix} 2.71 & 1.63 & 0.32 \\ 4.11 & 2.44 & 0.19 \\ 2.69 & 1.64 & 0.36 \end{pmatrix}$$

A matriz ampliada dos $n = 3$ sistemas é

$$(A|I) = \left(\begin{array}{ccc|ccc} 2.71 & 1.63 & 0.32 & 1 & 0 & 0 \\ 4.11 & 2.44 & 0.19 & 0 & 1 & 0 \\ 2.69 & 1.64 & 0.36 & 0 & 0 & 1 \end{array} \right)$$

Etapa 1:

$$I_{1,2}(A|I) \Leftrightarrow \left(\begin{array}{ccc|ccc} 4.11 & 2.44 & 0.19 & 0 & 1 & 0 \\ 2.71 & 1.63 & 0.32 & 1 & 0 & 0 \\ 2.69 & 1.64 & 0.36 & 0 & 0 & 1 \end{array} \right) \begin{matrix} m_{21} = -0.65937 \\ m_{31} = -0.65450 \end{matrix}$$

$$J_1 I_{1,2}(A|I) \Leftrightarrow \left(\begin{array}{ccc|ccc} 4.11 & 2.44 & 0.19 & 0 & 1 & 0 \\ 0 & 0.02114 & 0.19472 & 1 & -0.65937 & 0 \\ 0 & 0.04302 & 0.23565 & 0 & -0.65450 & 1 \end{array} \right)$$

Etapa 2:

$$I_{2,3}J_1I_{1,2}(A|I) \Leftrightarrow \left(\begin{array}{ccc|ccc} 4.11 & 2.44 & 0.19 & 0 & 1 & 0 \\ 0 & 0.04302 & 0.23565 & 0 & -0.65450 & 1 \\ 0 & 0.02114 & 0.19472 & 1 & -0.65937 & 0 \end{array} \right) m_{32} = -0.49140$$

$$J_2I_{2,3}J_1I_{1,2}(A|I) \Leftrightarrow \left(\begin{array}{ccc|ccc} 4.11 & 2.44 & 0.19 & 0 & 1 & 0 \\ 0 & 0.04302 & 0.23565 & 0 & -0.65450 & 1 \\ 0 & 0 & 0.07892 & 1 & -0.33775 & -0.49140 \end{array} \right)$$

A inversa calcula-se coluna a coluna resolvendo os $n = 3$ sistemas por substituição inversa:

• **1ª coluna**

$$\left(\begin{array}{ccc|c} 4.11 & 2.44 & 0.19 & 0 \\ 0 & 0.04302 & 0.23565 & 0 \\ 0 & 0 & 0.07892 & 1 \end{array} \right) \longrightarrow \left(\begin{array}{c} 40.62 \\ -69.40807 \\ 12.67106 \end{array} \right)$$

• **2ª coluna**

$$\left(\begin{array}{ccc|c} 4.11 & 2.44 & 0.19 & 1 \\ 0 & 0.04302 & 0.23565 & -0.65450 \\ 0 & 0 & 0.07892 & -0.33775 \end{array} \right) \longrightarrow \left(\begin{array}{c} 4.44403 \\ 8.22872 \\ -4.27965 \end{array} \right)$$

• **3ª coluna**

$$\left(\begin{array}{ccc|c} 4.11 & 2.44 & 0.19 & 0 \\ 0 & 0.04302 & 0.23565 & 1 \\ 0 & 0 & 0.07892 & -0.49140 \end{array} \right) \longrightarrow \left(\begin{array}{c} -33.76063 \\ 57.35214 \\ -6.22656 \end{array} \right)$$

A inversa da matriz A é, assim,

$$A^{-1} = \left(\begin{array}{ccc} 40.62 & 4.44403 & -33.76063 \\ -69.40807 & 8.22872 & 57.35214 \\ 12.67106 & -4.27965 & -6.22656 \end{array} \right)$$

3.7 Exercícios

1. Um engenheiro supervisiona a produção de 3 modelos de automóveis. Para a sua produção, são necessários 3 tipos de materiais: metal, tecido e plástico. As quantidades para produzir um carro de cada modelo são:

	metal (kg./carro)	tecido(kg./carro)	borracha(Kg./carro)
‘Jeep’	2.71	4.11	2.69
‘coupé’	1.63	2.44	1.64
‘V6’	0.32	0.19	0.36

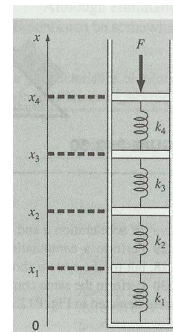
Existem em *stock*, respectivamente 38.48, 56.69, 38.54 kg. de metal, tecido e borracha. Quantos automóveis podem ser produzidos com a quantidade de *stock* existente?

Resolva o sistema por um método direto e estável usando 4 casas decimais nos cálculos.

2. Considere a figura representando um sistema de 4 molas ligadas em série sujeito a uma força F de 2000 Kg. Numa situação de equilíbrio, as equações força-balanço deduzidas definem inter-relações entre as molas:

$$\begin{cases} k_2(x_2 - x_1) = k_1 x_1 \\ k_3(x_3 - x_2) = k_2(x_2 - x_1) \\ k_4(x_4 - x_3) = k_3(x_3 - x_2) \\ F = k_4(x_4 - x_3) \end{cases}$$

em que $k_1 = 150$, $k_2 = 50$, $k_3 = 75$ e $k_4 = 225$ são as constantes das molas (kg/s^2).



Resolva o sistema pelo método EGPP, usando 5 casas decimais nos cálculos.

3. Considere os três sistemas e equações lineares

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 8 \\ 9 \end{pmatrix}, \quad A \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 9 \\ 1 \\ 8 \\ 4 \end{pmatrix} \quad \text{e} \quad A \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ -7 \\ 5 \end{pmatrix}$$

com

$$A = \begin{pmatrix} 4 & 2 & 1 & -3 \\ 1 & 2 & -1 & 0 \\ 3 & -1 & 2 & 4 \\ 0 & 2 & 4 & 3 \end{pmatrix}.$$

Calcule as três soluções de uma só vez, usando um método direto e estável.

4. Considere o sistema

$$\begin{cases} x_1 + 0.5x_2 + 0.5x_3 = 2 \\ 0.5x_1 + x_2 + 0.5x_3 = 2 \\ 0.5x_1 + 0.5x_2 + x_3 = 2 \end{cases}$$

a) Use o método da eliminação de Gauss com pivotagem parcial para calcular a sua solução.

b) Calcule o determinante da matriz dos coeficientes.

5. Considere o seguinte sistema de equações lineares

$$\begin{cases} 0.8x_1 + 1.4x_2 + 3.0x_3 = 12.6 \\ 0.6x_1 + 0.9x_2 + 2.8x_3 = 10.8 \\ 2.0x_1 + 1.0x_2 + 1.1x_3 = 4.0 \end{cases}$$

a) Calcule a inversa da matriz dos coeficientes por um método direto e estável.

b) Calcule o determinante da matriz dos coeficientes por um método direto e estável.

c) Resolva o sistema por um método direto e estável.

6. Dada a matriz

$$A = \begin{pmatrix} 2.4 & 6.0 & -2.7 & 5.0 \\ -2.1 & -2.7 & 5.9 & -4.0 \\ 3.0 & 5.0 & -4.0 & 6.0 \\ 0.9 & 1.9 & 4.7 & 1.8 \end{pmatrix}$$

e o vetor $b = (14.6, -11.4, 14.0, -0.9)^T$,

- a) resolva o sistema correspondente por um método direto e estável,
- b) calcule o determinante de A por um método direto e estável e
- c) calcule A^{-1} usando o método de eliminação de Gauss com pivotagem parcial.

3.8 Soluções

1. $x \approx (10, 6, 5)^T$.

2. $x = (13.33334, 53.33334, 80.00001, 88.88889)^T$.

3. $x = (1, 1, 1, 1)^T$, $y = (2, 0, 1, 0)^T$ e $z = (-1, 2, 1, -1)^T$.

4. a) $x = (1, 1, 1)^T$.

b) $\det(A) = 0.5$.

5. a) $A^{-1} = \begin{pmatrix} -0.968951 & 0.781584 & 0.653105 \\ 2.644540 & -2.740899 & -0.235546 \\ -0.642398 & 1.070664 & -0.064240 \end{pmatrix}$

b) $\det(A) = 1.868$.

c) $x = (-1.155247, 2.777303, 3.211991)^T$.

6. a) $x = (1, 2, -1, -0.5)$.

b) $\det(A) = -4.872$.

c) $A^{-1} = \begin{pmatrix} 3.009031 & -13.009031 & -13.222084 & 6.806240 \\ 0.726601 & -0.726601 & -1.185961 & 0.320197 \\ -0.049261 & 0.049261 & 0.029557 & 0.147783 \\ -2.142857 & 7.142857 & 7.785714 & -3.571429 \end{pmatrix}$.

3.9 MATLAB

3.9.1 Método EGPP

O MATLAB permite resolver sistemas lineares, calcular a inversa e o determinante de uma matriz quadrada, através de EGPP. Sendo \mathbf{A} a matriz dos coeficientes e \mathbf{b} o vetor dos termos independentes, os comandos a usar são $\mathbf{A} \backslash \mathbf{b}$, para resolver o sistema linear correspondente, $\text{inv}(\mathbf{A})$ para calcular a inversa de \mathbf{A} e $\det(\mathbf{A})$ para calcular o determinante de \mathbf{A} .

Exemplo 3.5 Dado o sistema linear do Exemplo 3.2, para se resolver o sistema, depois de se introduzir a matriz \mathbf{A} e o vetor \mathbf{b} em MATLAB, basta executar o comando

```
>>x=A\b
```

Exemplo 3.5 *Dada a matriz do Exemplo 3.3, depois de introduzida a matriz A basta executar o comando*

```
>>determinante=det(A)
```

Exemplo 3.5 *Dada a matriz do Exemplo 3.4, depois de introduzida a matriz A, basta executar o comando*

```
inversa=inv(A)
```

3.9.2 Exercícios para resolver em MATLAB

1. Resolva os seguintes sistemas através de um método direto e estável.

$$\text{a) } \begin{cases} 4x_1 + 13x_2 + 2x_3 = -15 \\ -8x_1 + 10x_2 + 8x_3 = 6 \\ 2x_1 + 6.5x_2 + 5.5x_3 = -3 \end{cases}$$

$$\text{b) } \begin{cases} 2x_1 + 3x_2 = 1 \\ 2x_1 + 3.0001x_2 = 0.9999 \end{cases}$$

$$\text{c) } \begin{cases} 2x_1 + 3x_2 = 1 \\ 2x_1 + 3.0001x_2 = 2 \end{cases}$$

$$\text{d) } \begin{cases} -30x_1 + 9x_2 + 9x_3 = 10 \\ 10x_1 - 2.9999x_2 - 2.9999x_3 = -3.3333 \\ 6x_1 - 6x_2 - 20x_3 = 10 \end{cases}$$

2. Calcule o determinante e a inversa das seguintes matrizes:

$$\text{a) } A = \begin{pmatrix} 2 & 3 \\ 2 & 3.0001 \end{pmatrix}$$

$$\text{b) } B = \begin{pmatrix} -602.9 & -0.4762 & 301.0 \\ -248.8 & -0.1048 & 124.2 \\ -200.6 & 0 & 101.7 \end{pmatrix}$$

$$\text{c) } C = \begin{pmatrix} 10 & 1 & 4 & 0 \\ 1 & 10 & 5 & -1 \\ 4 & 5 & 10 & 7 \\ 0 & -1 & 7 & 9 \end{pmatrix}$$

3. Considere a matriz A e o vetor b .

$$\begin{pmatrix} 2.4 & 6.0 & -2.7 & 5.0 \\ -2.1 & -2.7 & 5.9 & -4.0 \\ 3.0 & 5.0 & -4.0 & 6.0 \\ 0.9 & 1.9 & 4.7 & 1.8 \end{pmatrix} b = \begin{pmatrix} 14.6 \\ -11.4 \\ 14.0 \\ -0.9 \end{pmatrix}.$$

- a) Resolva o sistema correspondente por um método direto e estável.
- b) Calcule o determinante de A por um método direto e estável.
- c) Calcule A^{-1} usando o método de eliminação de Gauss com pivotagem parcial.

4. Considere o sistema linear:

$$\begin{cases} 6x_1 + x_2 + 2x_3 + x_5 & = & 10 \\ 2x_1 + 8x_2 + x_3 + 2x_4 + 2x_5 & = & 15 \\ x_1 - 2x_2 + 8x_3 + x_4 & = & 8 \\ -x_3 + 9x_4 + 2x_5 & = & 10 \\ x_1 + x_2 - x_4 + 7x_5 & = & 8 \end{cases}$$

- a) Resolva o sistema por EGPP.
- b) Calcule o determinante da matriz dos coeficientes.
- c) Calcule a inversa da matriz dos coeficientes.

Capítulo 4

Equações não lineares

4.1 Forma geral do problema

O problema que se pretende tratar neste capítulo tem a forma geral

$$f(x) = 0, \quad f : \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

No caso de uma equação não linear com uma variável, $n = 1$, o problema reduz-se a

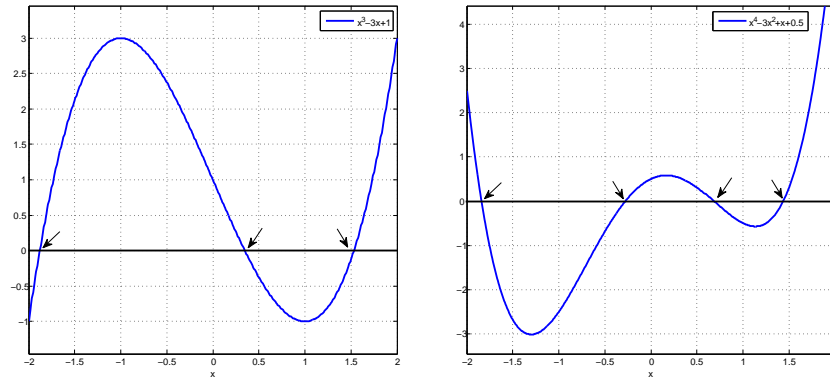
$$f(x) = 0, \quad \text{com } f : \mathbb{R} \rightarrow \mathbb{R},$$

em que x é a variável independente do problema e $y = f(x)$ é a variável dependente do problema.

Quando f é não linear em x , esta equação pode não ter soluções reais, ter uma só solução real, ter várias soluções reais, ter soluções complexas e ter soluções reais e complexas. Uma solução real da equação $f(x) = 0$ é equivalente a uma raiz real da equação $f(x) = 0$, a um zero real da função f ou ainda à interseção de f com o eixo do X no plano real XOY .

Podem dividir-se as equações não lineares em dois tipos genéricos.

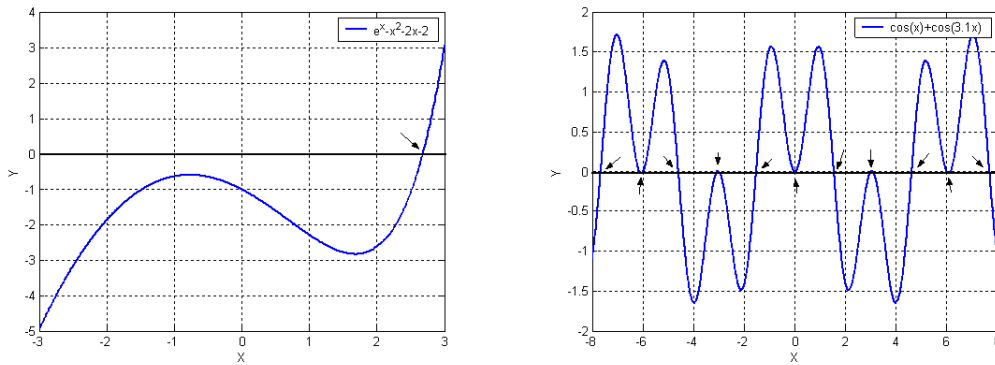
1. **as algébricas**, que envolvem apenas operações aritméticas básicas, de que são exemplo os polinómios (Figura 4.1),
2. **as transcendent**es, que envolvem funções trigonométricas, exponenciais, logarítmicas, entre outras, como são exemplos as equações $e^x - x^2 - 2x - 2 = 0$ e $\cos(x) + \cos(3.1x) = 0$ (Figura 4.2).



(a) $x^3 - 3x + 1 = 0$

(b) $x^4 - 3x^2 + x + 0.5 = 0$

Figura 4.1: Soluções de duas equações algébricas.



(a) $e^x - x^2 - 2x - 2 = 0$

(b) $\cos(x) + \cos(3.1x) = 0$

Figura 4.2: Soluções de duas equações transcendentais.

4.2 Solução de uma equação não linear

Nem sempre é fácil, ou mesmo possível, resolver uma equação não linear analiticamente. Os métodos mais usados na resolução deste tipo de problemas são os **métodos iterativos**. No entanto, estes métodos exigem que seja fornecida uma ou várias aproximações iniciais. Estas podem ser identificadas através da localização de raízes, por exemplo, recorrendo a **métodos gráficos**.

4.2.1 Métodos gráficos

Os métodos gráficos não permitem conhecer em rigor a solução de uma equação não linear, mas permitem localizá-la. Tornam-se, assim, de extrema importância quando se pretende

saber onde se encontra a solução ou soluções de uma equação não linear, como é o caso de se determinar uma aproximação ou aproximações iniciais que servirão de ponto de partida para a implementação de um método iterativo. Este processo pode fazer-se através da representação gráfica de:

- $f(x)$ no plano XOY se f for fácil de representar,
- $g(x)$ e $h(x)$ no plano XOY , em que

$$f(x) = 0 \Leftrightarrow g(x) = h(x)$$

se f for difícil de representar e g e h forem fáceis de representar.

Exemplo 4.1 *Localização gráfica de raízes*

Considere a equação $x^3 - 3x + 1 = 0$. Para as suas raízes pode fazer-se a representação gráfica de $f(x) = x^3 - 3x + 1$.

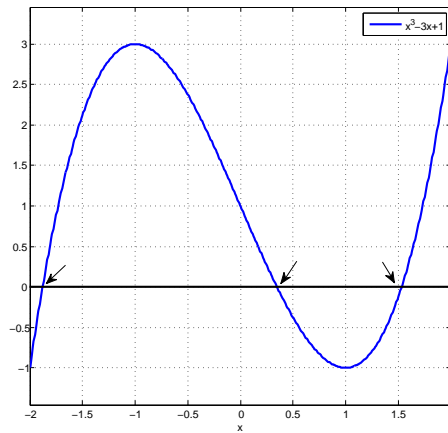


Figura 4.3: Representação gráfica de $f(x)$.

Pode verificar-se na Figura 4.3 a localização das três soluções desta equação. Uma próxima de -2, outra próxima de 0.5 e outra ainda próxima de 1.5.

No entanto, esta equação pode ser reformulada:

$$x^3 - 3x + 1 = 0 \Leftrightarrow x^3 - (3x - 1) = 0 \Leftrightarrow \begin{cases} x^3 &= 3x - 1 \\ g(x) &= h(x) \end{cases}$$

Desta forma, os zeros de $f(x) = x^3 - 3x + 1$ são os pontos de interseção de $g(x) = x^3$ com $h(x) = 3x - 1$. Como pode verificar-se na Figura 4.4, as soluções encontradas são exatamente as mesmas que pelo processo anterior.

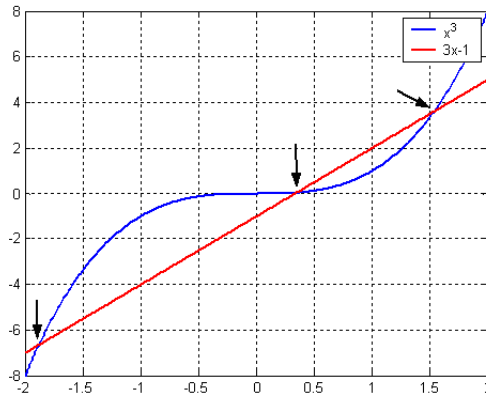


Figura 4.4: Representação gráfica de $g(x)$ e $h(x)$.

Exemplo 4.2 *Localização gráfica de raízes*

Considere a equação $\sin x + x - 2 = 0$. Esta equação pode resolver-se por métodos gráficos representando no plano XOY $f(x) = \sin x + x - 2$. Pode ver-se na Figura 4.5 que esta equação

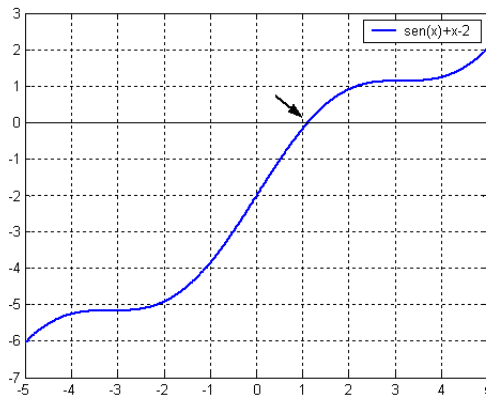


Figura 4.5: Representação gráfica de $f(x)$.

tem uma única solução, próxima de 1.

Da mesma forma que no Exemplo 4.1, pode desdobrar-se a função $f(x) = \sin x + x - 2$ em duas mais simples.

$$\sin x + x - 2 = 0 \Leftrightarrow \sin x - (-x + 2) = 0 \Leftrightarrow \begin{cases} \sin x &= -x + 2 \\ g(x) &= h(x) \end{cases}$$

Assim, os zeros de $f(x) = \sin x + x - 2$ são os pontos onde se verifica a interseção de $g(x) =$

$\sin x$ com $h(x) = -x + 2$. Mais uma vez, a solução obtida é exatamente a mesma que na representação gráfica anterior, como se pode verificar na Figura 4.6 ($x \approx 1$).

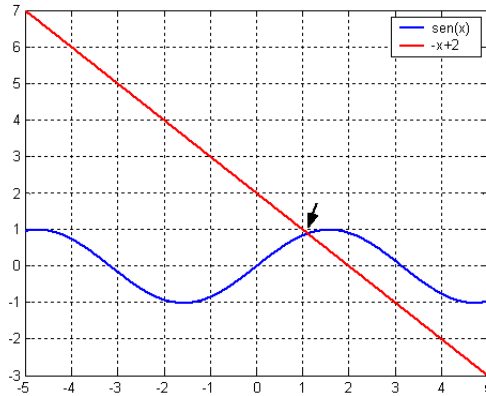


Figura 4.6: Representação gráfica de $g(x)$ e $h(x)$.

4.2.2 Métodos iterativos

Da representação gráfica, como pode verificar-se na Subsecção 4.2.1, é possível retirar um intervalo que contenha a solução pretendida ou apenas um valor que esteja próximo da solução pretendida. A aproximação obtida por métodos gráficos pode ser melhorada através de métodos iterativos. Entre estes métodos encontram-se

- o método da secante;
- o método de Newton.

Qualquer um destes métodos pode ser usado para calcular raízes reais de $f(x) = 0$. No entanto, podem também ser usados para calcular raízes complexas, desde que se introduza aritmética complexa nos cálculos e as aproximações iniciais sejam números complexos.

Como qualquer processo iterativo, o método da secante e o método de Newton exigem um critério de paragem, já que a solução exata só é obtida para um número infinito de iterações, o que em termos computacionais não é exequível. Os critérios de paragem usados nestes métodos são os que se apresentam em (4.1) e (4.2). Os valores ε_1 e ε_2 são quantidades positivas e próximas de zero. Quanto menores forem estas quantidades, mais próxima será a aproximação x_{k+1} da solução x^* .

Critérios de paragem

São usados como critérios de paragem dos métodos iterativos para a resolução de equações não lineares

- a estimativa do erro relativo da aproximação próxima de zero

$$\frac{|x_{k+1} - x_k|}{|x_{k+1}|} \leq \varepsilon_1, \quad (4.1)$$

- o valor absoluto da função na última aproximação próximo de zero

$$|f(x_{k+1})| \leq \varepsilon_2. \quad (4.2)$$

4.2.3 Método da secante

O método da secante precisa de duas aproximações iniciais para iniciar o processo iterativo. Em cada iteração, e com base em dois pontos (aproximações), o método aproxima $f(x)$ por uma reta definida por esses dois pontos (secante). O ponto de interseção desta reta com o eixo do X fornece a aproximação seguinte. Em termos gráficos, podem ver-se na Figura 4.7 as duas primeiras iterações do método da secante. Na primeira iteração (Figura 4.7(a)), parte-se

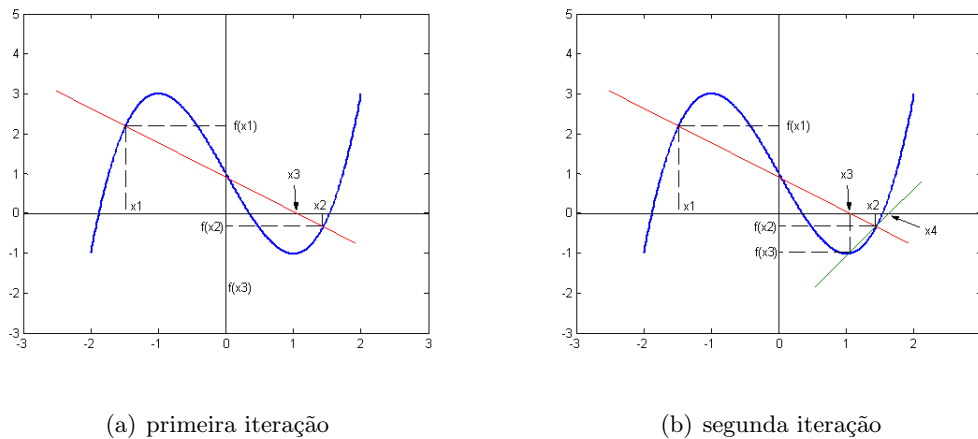


Figura 4.7: Duas iterações do método da secante

de duas aproximações iniciais $(x_1, f(x_1))$ e $(x_2, f(x_2))$. Traça-se a reta que passa nestes dois pontos e obtém-se a aproximação seguinte $(x_3, f(x_3))$. Na segunda iteração (Figura 4.7(b)), a reta é traçada a partir dos pontos $(x_2, f(x_2))$ e $(x_3, f(x_3))$. A nova aproximação é $(x_4, f(x_4))$. O processo iterativo é repetido até se obter uma solução que verifique os critérios de paragem.

Equação iterativa do método da secante

Sejam dois pontos $(x_{k-1}, f(x_{k-1}))$ e $(x_k, f(x_k))$. O ponto de interseção da reta secante que passa por estes dois pontos e o eixo do X é x_{k+1} dado por

$$x_{k+1} = x_k - \frac{(x_k - x_{k-1})f(x_k)}{f(x_k) - f(x_{k-1})}, \quad k = 2, 3, \dots$$

Descreve-se, de forma detalhada, o método da secante no Algoritmo 4.1.

Algoritmo 4.1 Método da secante

ler: x_1 e x_2 (aproximações iniciais)

$k \leftarrow 1$

repetir

$k \leftarrow k + 1$

$x_{k+1} \leftarrow x_k - \frac{(x_k - x_{k-1})f(x_k)}{f(x_k) - f(x_{k-1})}$

até $\frac{|x_{k+1} - x_k|}{|x_{k+1}|} \leq \varepsilon_1 \wedge |f(x_{k+1})| \leq \varepsilon_2$

$x^* \leftarrow x_{k+1}$

$f(x^*) \leftarrow f(x_{k+1})$

Exemplo 4.3 *Um certo equipamento de 20000 euros vai ser pago durante 6 anos. O pagamento anual é de 4000 euros. A relação entre o custo do equipamento P , o pagamento anual A , o número de anos n e a taxa de juro i é a seguinte:*

$$A = P \frac{i(1+i)^n}{(1+i)^n - 1}.$$

Utilize o método que não recorre à derivada para determinar a taxa de juro utilizada nos cálculos. O valor da taxa de juro pertence ao intervalo $[0.05, 0.15]$. Use $\varepsilon_1 = \varepsilon_2 = 0.005$. Use seis casas decimais nos cálculos.

Resolução:

Substituindo os valores de P , A e n , vem

$$4000 = 20000 \frac{i(1+i)^6}{(1+i)^6 - 1} \quad \Leftrightarrow \quad 20000 \frac{i(1+i)^6}{(1+i)^6 - 1} - 4000 = 0 \quad \Leftrightarrow \quad \frac{5i(1+i)^6}{(1+i)^6 - 1} - 1 = 0$$

Logo,

$$f(i) = \frac{5i(1+i)^6}{(1+i)^6 - 1} - 1.$$

• **1ª iteração** ($k = 2$)

$$i_1 = 0.05, \quad f(i_1) = -0.014913$$

$$i_2 = 0.15, \quad f(i_2) = 0.321185$$

$$i_3 = i_2 - \frac{(i_2 - i_1)f(i_2)}{f(i_2) - f(i_1)} = 0.054437$$

Critério de Paragem

$$f(i_3) = -0.000891$$

$$|f(i_3)| \leq \varepsilon_2 \Leftrightarrow 0.000891 \leq 0.005 \quad (\text{verdadeiro})$$

Nota: Uma vez que se verifica esta condição, para o processo iterativo poder terminar tem que se verificar também a outra.

$$\frac{|i_3 - i_2|}{|i_3|} \leq \varepsilon_1 \Leftrightarrow 1.755479 \leq 0.005 \quad (\text{falso})$$

• **2ª iteração** ($k = 3$)

$$i_2 = 0.15, \quad f(i_2) = 0.321185$$

$$i_3 = 0.054437, \quad f(i_3) = -0.000891$$

$$i_4 = i_3 - \frac{(i_3 - i_2)f(i_3)}{f(i_3) - f(i_2)} = 0.054701$$

Critério de Paragem

$$f(i_4) = -0.000054$$

$$|f(i_4)| \leq \varepsilon_2 \Leftrightarrow 0.000054 \leq 0.005 \quad (\text{verdadeiro})$$

Nota: Uma vez que se verifica esta condição, para o processo iterativo poder terminar tem que se verificar também a outra.

$$\frac{|i_4 - i_3|}{|i_4|} \leq \varepsilon_1 \Leftrightarrow 0.004826 \leq 0.005 \quad (\text{verdadeiro})$$

Uma vez que se verificam ambas as condições do critério de paragem, o processo iterativo termina com $i \approx 0.054701$ e $f(i) \approx 0.000054$.

Condições de convergência do método da secante

Nem sempre se pode garantir que o método da secante convirja. Para garantir a convergência para a solução é necessário que

- x^* é tal que $f(x^*) = 0$,
- $f(x)$ é continuamente diferenciável,
- $f'(x^*) \neq 0$,
- as aproximações iniciais x_1 e x_2 têm de se encontrar na vizinhança de x^* (convergência local).

Se se verificarem as condições anteriores, o método iterativo da secante converge e

$$\lim_{k \rightarrow \infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^p} = L, \quad L > 0, \quad p = 1.618.$$

Por essa razão diz-se que o método da secante exibe convergência superlinear.

Situação de divergência

Ao longo do processo iterativo, pode acontecer que duas aproximações tenham valores de f muito próximos. Esta situação leva a uma divergência do método (Figura 4.8).

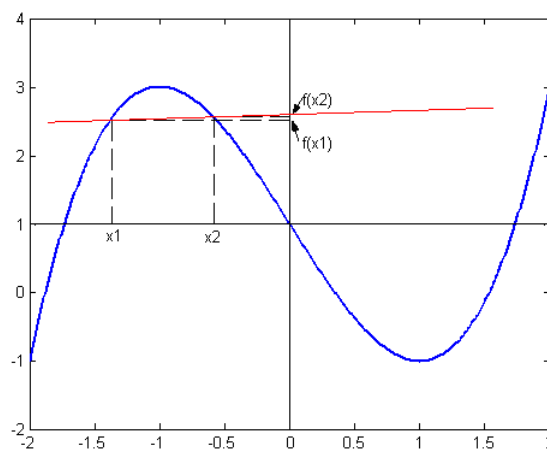
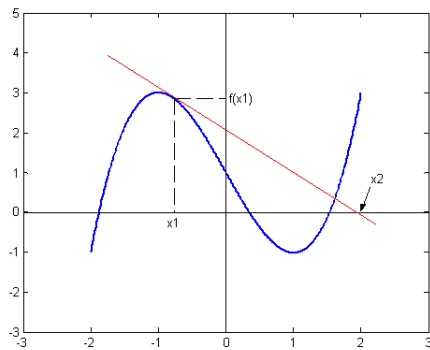


Figura 4.8: Situação de divergência no método da secante.

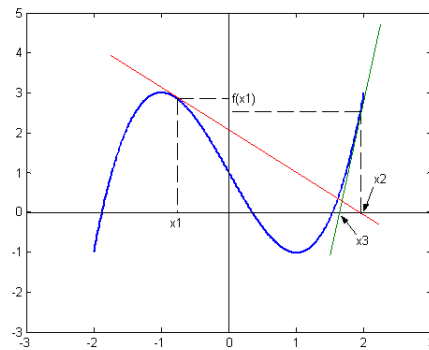
4.2.4 Método de Newton

O método de Newton precisa apenas de uma aproximação inicial. Em cada iteração usa informação de f e de f' relativa a um ponto, aproximando a função $f(x)$ por uma reta que é tangente a f nesse ponto. O ponto de interseção dessa recta com o eixo do X fornece a aproximação seguinte.

Na primeira iteração (Figura 4.9(a)), parte-se de uma aproximação inicial $(x_1, f(x_1))$. Traça-se a reta tangente a $f(x)$ nesse ponto e obtém-se a aproximação seguinte $(x_2, f(x_2))$. Na segunda iteração (Figura 4.9(b)), a reta é traçada a partir do ponto $(x_2, f(x_2))$. A nova aproximação é $(x_3, f(x_3))$. O processo iterativo é repetido até se obter uma solução com a precisão desejada.



(a) primeira iteração



(b) segunda iteração

Figura 4.9: Duas iterações do método de Newton

Equação iterativa do método de Newton

Considere-se o ponto $(x_k, f(x_k))$ e $f'(x_k)$. Então, o ponto de interseção da reta, que passa por este ponto, com declive definido por $f'(x_k)$ com o eixo do X é x_{k+1} , sendo

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 1, 2, \dots$$

Descreve-se, de forma detalhada, o método de Newton no Algoritmo 4.2.

Algoritmo 4.2 Método de Newton

ler: x_1 (aproximação inicial) $k \leftarrow 0$ **repetir** $k \leftarrow k + 1$

$$x_{k+1} \leftarrow x_k - \frac{f(x_k)}{f'(x_k)}$$

até $\frac{|x_{k+1} - x_k|}{|x_{k+1}|} \leq \varepsilon_1 \wedge |f(x_{k+1})| \leq \varepsilon_2$ $x^* \leftarrow x_{k+1}$ $f(x^*) \leftarrow f(x_{k+1})$

Exemplo 4.4 Considere a seguinte função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$f(x) = x^2 - e^x.$$

Calcule a raiz negativa utilizando o método iterativo de Newton. Considere $x_1 = 0.25$. Faça duas iterações e apresente o erro relativo.

Resolução:**• 1ª iteração ($k = 1$)**

$$x_1 = 0.25, \quad f(x_1) = -1.2215, \quad f'(x_1) = -0.7840$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = -1.3080$$

• 2ª iteração ($k = 2$)

$$x_2 = -1.3080, \quad f(x_2) = 1.4405, \quad f'(x_2) = -2.8864$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = -0.8089$$

Ao fim da segunda iteração, $x \approx -0.8089$ e o erro relativo é

$$\frac{|x_3 - x_2|}{|x_3|} = 0.6170.$$

Condições de convergência do método de Newton

O método de Newton nem sempre converge. Para se garantir uma convergência para a solução de uma equação não linear através deste método é necessário que

- x^* é tal que $f(x^*) = 0$,
- $f(x)$ é continuamente diferenciável,
- $f'(x^*) \neq 0$ e
- a aproximação inicial x_1 tem de se encontrar na vizinhança de x^* (convergência local).

Se se verificarem todas as condições anteriores, o método de Newton converge e

$$\lim_{k \rightarrow \infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^p} = L, \quad L > 0, \quad p = 2.$$

Por esta razão diz-se que o método de Newton exhibe convergência quadrática.

Situação de divergência

Ao longo do processo iterativo pode acontecer que o declive da reta que é tangente a f na aproximação, que vai ser usada para gerar a nova aproximação, seja um valor próximo de zero. Esta situação leva a uma divergência do método (Figura 4.10).

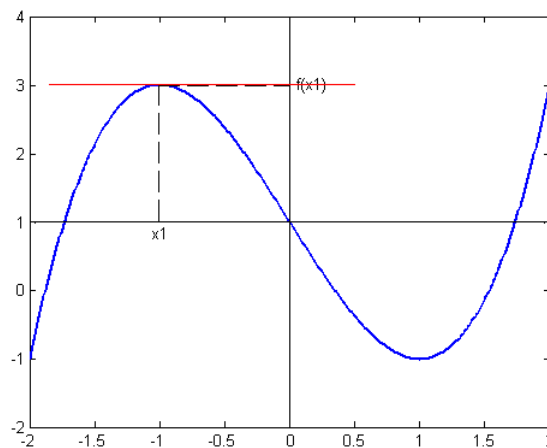


Figura 4.10: Situação de divergência no método de Newton.

4.2.5 Método da secante *versus* método de Newton

Os métodos iterativos da secante e de Newton têm vantagens e desvantagens. Como tal, devem ser escolhidos de forma a adequarem-se o melhor possível ao problema que se está a resolver. Apresentam-se de seguida algumas linhas gerais que podem ser seguidas de forma a ajudar na escolha de um ou outro em determinadas situações.

- Quando ambos os métodos convergem, o método de Newton é, em geral, mais rápido, já que exibe convergência quadrática, ao passo que o método da secante tem convergência superlinear.
- O método da secante necessita apenas de informação sobre $f(x)$ ao passo que o método de Newton exige também informação sobre $f'(x)$. Se $f(x)$ não for diferenciável ou se a sua expressão analítica for muito complexa, deve optar-se pelo método da secante, já que, ou não é possível o cálculo da derivada ou este é demasiado dispendioso, em termos de esforço computacional.

4.3 Solução de um sistema de equações não lineares

4.3.1 Forma geral do problema

O problema que se pretende resolver tem a forma geral

$$f(x) = 0, \quad f : \mathbb{R}^n \rightarrow \mathbb{R}^n \Leftrightarrow \begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ f_n(x_1, x_2, \dots, x_n) = 0, \end{cases}$$

em que f é o vetor $(f_1 \ f_2 \ \dots \ f_n)^T$, x é o vetor $(x_1 \ x_2 \ \dots \ x_n)^T$ e pelo menos uma das funções de f é não linear (Exemplo 4.5).

Exemplo 4.5 Considerando o sistema de equações não lineares com três equações e três variáveis $(x_1, x_2 \text{ e } x_3)$

$$\begin{cases} 3x_1 - \cos(x_2 x_3) - 0.5 = 0 \\ x_1^2 - 625x_2^2 = 0 \\ e^{-x_1 x_2} + 20x_3 + 9 = 0, \end{cases}$$

este pode ser reescrito na forma genérica

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ f_3(x_1, x_2, \dots, x_n) = 0 \end{cases} \Leftrightarrow f(x) = 0.$$

Assim, f é o vetor que contém as três funções f_1 , f_2 e f_3 , ou seja, $f = (f_1 \ f_2 \ f_3)^T$ e x é o vetor que contém as três variáveis x_1 , x_2 e x_3 , ou seja, $x = (x_1 \ x_2 \ x_3)^T$. Trata-se de um problema de dimensão $n = 3$.

4.3.2 Método de Newton

Para resolver um sistema de equações não lineares ($n > 1$) pode usar-se o método de Newton, que não é mais que uma generalização do caso em que este método se aplica à resolução de uma equação não linear ($n = 1$).

A equação iterativa do método de Newton para resolver uma equação não linear, $f(x) = 0$, $f : \mathbb{R} \rightarrow \mathbb{R}$, pode ser reformulada e escrita na forma

$$x_{k+1} = x_k + \Delta x_k, \quad k = 1, 2, \dots$$

com

$$\begin{aligned} \Delta x_k &= -\frac{1}{f'(x_k)} f(x_k) \\ &= -(f'(x_k))^{-1} f(x_k). \end{aligned}$$

Assim, para um sistema de n equações em n variáveis, a primeira derivada da função, que é um escalar, é substituída pela matriz do Jacobiano $n \times n$ do vetor de funções, e a equação iterativa fica

$$x_{k+1} = x_k + \Delta x_k, \quad k = 1, 2, \dots$$

com

$$\Delta x_k = -(J(x_k))^{-1} f(x_k), \tag{4.3}$$

em que $x \in \mathbb{R}^n$, $\Delta x \in \mathbb{R}^n$, $f \in \mathbb{R}^n$ e $J(x) \in \mathbb{R}^{n \times n}$.

A matriz do Jacobiano contém as primeiras derivadas parciais das funções f_1, f_2, \dots, f_n

em ordem às variáveis x_1, x_2, \dots, x_n .

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}.$$

Equação iterativa do método de Newton

A equação (4.3) pode ser reescrita da seguinte forma:

$$J(x_k)\Delta x_k = -f(x_k). \quad (4.4)$$

Significa que o vetor Δx_k em cada iteração se obtém através da resolução do sistema linear (4.4) pelo método de eliminação de Gauss com pivotagem parcial, por exemplo.

Critério de paragem

Tal como para as equações não lineares, podem usar-se como critério de paragem as medidas

- estimativa do erro relativo da aproximação

$$\frac{\|\Delta x_k\|_2}{\|x_{k+1}\|_2} \leq \varepsilon_1 \quad (4.5)$$

- aproximação ao zero de f

$$\|f(x_{k+1})\|_2 \leq \varepsilon_2 \quad (4.6)$$

sendo ε_1 e ε_2 quantidades positivas e próximas de zero. $\|\cdot\|_2$ representa a norma 2.

A implementação do método de Newton encontra-se descrita no Algoritmo 4.3.

Algoritmo 4.3 Método de Newton para sistemas de equações não lineares**ler:** x_1 (aproximação inicial) $k \leftarrow 0$ calcular $J(x)$ **repetir** $k \leftarrow k + 1$ calcular $J(x_k)$ calcular $f(x_k)$ resolver o sistema linear $J(x_k)\Delta x_k = -f(x_k)$ por EGPP para calcular o vetor Δx_k $x_{k+1} \leftarrow x_k + \Delta x_k$ **até** $\frac{\|\Delta x_k\|_2}{\|x_{k+1}\|_2} \leq \varepsilon_1 \wedge \|f(x_{k+1})\|_2 \leq \varepsilon_2$ $x^* \leftarrow x_{k+1}$ $f(x^*) \leftarrow f(x_{k+1})$

Exemplo 4.6 Considere o sistema do Exemplo 4.5. Tome $x^1 = (0, 1, 0)^T$ para valor inicial e faça apenas uma iteração. Apresente uma medida da proximidade ao zero das funções.

Resolução:**1ª iteração** ($k = 1$)

$$x_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad f(x_1) = \begin{pmatrix} -1.5 \\ -625 \\ 10 \end{pmatrix}, \quad J(x_1) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1250 & 0 \\ -1 & 0 & 20 \end{pmatrix}$$

Resolver o sistema linear $J(x_1)\Delta_1 = -f(x_1)$ por EGPP para calcular Δ_1 :

$$\left(\begin{array}{ccc|c} 3 & 0 & 0 & 1.5 \\ 0 & -1250 & 0 & 625 \\ -1 & 0 & 20 & -10 \end{array} \right)$$

$$\Delta_1 = \begin{pmatrix} -0.5 \\ -0.5 \\ -0.475 \end{pmatrix}$$

$$x_2 = x_1 + \Delta_1 = \begin{pmatrix} 0.5 \\ 0.5 \\ -0.475 \end{pmatrix}$$

Aproximação ao zero absoluto das funções

$$f(x_2) = \begin{pmatrix} 0.0281 \\ -156 \\ 0.7840 \end{pmatrix}$$

$$\|f(x_2)\|_2 = 156.0020$$

Condições de convergência do método de Newton

Para que o método de Newton convirja quando aplicado à resolução do sistema

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

em que $f = (f_1 \dots f_n)^T$ e $x = (x_1 \dots x_n)^T$, é necessário que

- x^* é tal que $f(x^*) = 0$,
- f é um vetor de funções continuamente diferenciáveis,
- $J(x^*)$ é uma matriz não singular ($\exists (J(x^*))^{-1}$),
- $(J(x^*))^{-1}$ é limitada ($\|J(x_k) - J(x^*)\| \leq \beta, \beta > 0$)
- $J(x)$ é matriz Lipschitz contínua na vizinhança de x^* , ou seja, $\exists \gamma > 0 : \|J(x_k) - J(x^*)\| \leq \gamma \|x_k - x^*\|$,
- A aproximação inicial, x_1 , tem de estar na vizinhança de x^* - convergência local.

Verificando-se todas estas condições,

$$\lim_{k \rightarrow \infty} \frac{\|x^* - x_{k+1}\|}{\|x^* - x_k\|^p} = L, \quad L > 0, \quad p = 2,$$

o que significa que o método de Newton para a resolução de sistemas não lineares exhibe convergência quadrática.

4.4 Exercícios

1. Localize através do método gráfico os zeros das funções não lineares em x ,

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

b) $f(x) = \sin x + x - 2$;

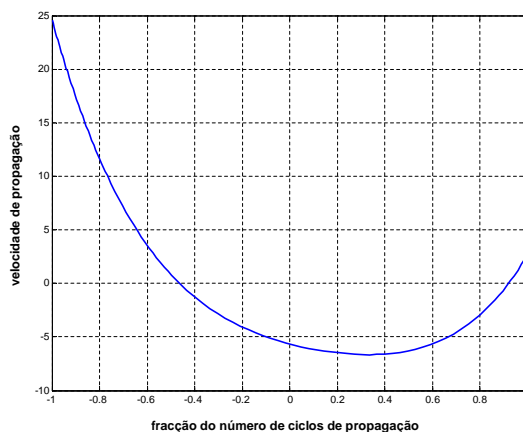
c) $f(x) = e^x + x - 1$;

d) $f(x) = x + \ln x$.

2. A função

$$a(x) = 2.02x^5 - 1.28x^4 + 3.06x^3 - 2.92x^2 - 5.66x + 6.08$$

é utilizada num estudo do comportamento mecânico de materiais, representando $a(x)$ o comprimento da fissura e x (> 0) uma fracção do número de ciclos de propagação. Pretende-se saber para que valores de x a velocidade de propagação da fissura é nula. Utilize um método que não recorre ao cálculo de derivadas, usando no critério de paragem $\varepsilon_1 = \varepsilon_2 = 10^{-2}$ ou no máximo três iterações. Use seis casas decimais nos cálculos.



3. Um certo equipamento de 20000 euros vai ser pago durante 6 anos. O pagamento anual é de 4000 euros. A relação entre o custo do equipamento P , o pagamento anual A , o número de anos n e a taxa de juro i é a seguinte:

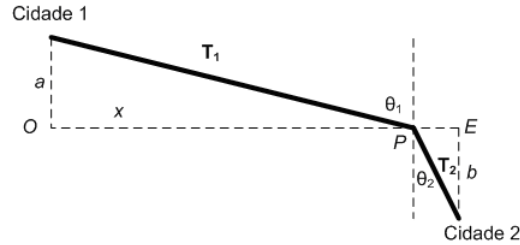
$$A = P \frac{i(1+i)^n}{(1+i)^n - 1}.$$

Utilize o método que não recorre à derivada para determinar a taxa de juro utilizada nos cálculos. O valor da taxa de juro pertence ao intervalo $[0.05, 0.15]$. Use $\varepsilon_1 = \varepsilon_2 = 0.005$. Use seis casas decimais nos cálculos.

4. O volume v de um líquido num tanque esférico de raio r está relacionado com a profundidade h do líquido da seguinte forma:

$$v = \frac{\pi h^2(3r - h)}{3}.$$

- a) Calcule, utilizando um método que não recorre ao cálculo de derivadas, a profundidade h , num tanque de raio $r = 1$ para um volume de 0.5. Utilize para aproximação inicial o intervalo $[0.25, 0.5]$. Faça 3 iterações e use seis casas decimais nos cálculos.
- b) Repita os cálculos, nas mesmas condições da alínea anterior, mas utilizando para aproximação inicial o intervalo $[2.5, 3]$. Comente os resultados e analise a viabilidade da solução encontrada.
5. Considere duas cidades localizadas como se mostra na figura. Uma petrolífera pretende construir uma conduta que ligue as duas cidades. Devido às diferenças no terreno, o custo para construir a conduta será C_1 milhões de euros por quilómetro para o troço \mathbf{T}_1 e C_2 milhões de euros por quilómetro para o troço \mathbf{T}_2 . Para tornar a construção mais económica, o ponto P de intersecção dos dois troços deve estar localizado de modo a que $C_1 \sin \theta_1 = C_2 \sin \theta_2$.



- (a) Usando a informação da figura e escrevendo esta equação em função de x (a distância de O a P), mostre que se obtém

$$C_2^2(L - x)^2(a^2 + x^2) = C_1^2x^2(b^2 + (L - x)^2),$$

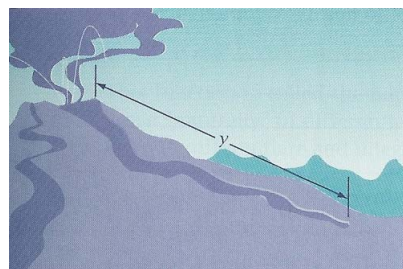
sendo L a distância de O a E .

- (b) Resolva a equação considerando $a = 3$, $b = 1$, $L = 4$, $C_1 = 1$ e $C_2 = 2$. Utilize o método de Newton e a aproximação inicial $x_1 = 3.75$ e $n_{\max} = 2$. Apresente uma estimativa do erro relativo. Use seis casas decimais nos cálculos.

6. A figura representa um vulcão em erupção.

A relação entre a distância y (milhas) percorrida pela lava e o tempo t (horas) é dada por:

$$y = 7(2 - 0.9^t).$$

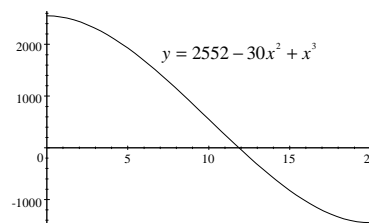
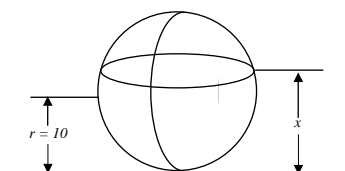


Existe uma aldeia no sopé da montanha a uma distância de $y = 10$. O gabinete de protecção civil advertiu os moradores da aldeia de que a lava chegaria às suas casas em menos de 6 horas. Calcule utilizando um método iterativo que recorre ao cálculo de derivadas o instante de tempo em que a lava do vulcão atinge a aldeia.

Considere $\varepsilon_1 = \varepsilon_2 = 10^{-3}$ ou no máximo três iterações. Use seis casas decimais nos cálculos.

Nota: $(a^x)' = a^x \ln(a)$, para a constante.

7. Uma bola esférica de raio $r = 10$ cm feita de uma substância cuja densidade é $\rho = 0.638$, foi colocada num recipiente com água.



Usando o método iterativo de Newton, calcule a distância x da parte submersa da bola sabendo que

$$f(x) \equiv \frac{\pi(x^3 - 3x^2r + 4r^3\rho)}{3} = 0.$$

Pare o processo iterativo quando o critério de paragem for verificado para $\varepsilon_1 = \varepsilon_2 = 0.001$, ou ao fim de três iterações. Use o método de Newton e seis casas decimais nos cálculos.

8. Num colector solar, um balanço de energia na placa absorvente e na placa de vidro produz o seguinte sistema de equações não lineares nas temperaturas absolutas da placa absorvente (x_1) e da placa de vidro (x_2)

$$\begin{cases} x_1^4 + 0.068x_1 - x_2^4 - 0.058x_2 &= 0.015 \\ x_1^4 + 0.058x_1 - 2x_2^4 - 0.117x_2 &= 0 \end{cases}.$$

Considerando a seguinte aproximação inicial $(x_1, x_2)_1 = (0.3, 0.3)$, implemente duas iterações do método de Newton. Apresente uma estimativa do erro relativo da aproximação calculada. Use seis casas decimais nos cálculos.

9. Considere o seguinte sistema

$$\begin{cases} -x_2 + 2x_1^n = 4 \\ -x_2 - x_2^m - x_1 = 8 \end{cases}$$

em que n e m são parâmetros.

Considere $m = 3$ e $n = 2$. Resolva o sistema utilizando para aproximação inicial o ponto $x_1 = (1, -2)^T$. Para o critério de paragem use $\varepsilon_1 = \varepsilon_2 = 10^{-2}$ (ou no máximo duas iterações). Use seis casas decimais nos cálculos.

10. Existe um par de valores que anula as primeiras derivadas parciais da função de duas variáveis

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

Usando um método iterativo, e a partir da aproximação inicial $(x, y)_1 = (-1, 1)$, determine esse par de modo que a estimativa do erro relativo da aproximação calculada não exceda 0.05 (duas iterações). Use seis casas decimais nos cálculos.

11. Usando o método de Newton, determine um dos pontos de interseção da circunferência

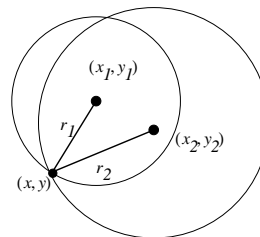
$$x_1^2 + x_2^2 = 2$$

com a hipérbole

$$x_1^2 - x_2^2 = 1.$$

Considere os valores iniciais $(x_1, x_2)_1 = (1.5, 0.5)$ e para a paragem do processo iterativo use $\varepsilon_1 = \varepsilon_2 = 0.05$ ou no máximo duas iterações. Use seis casas decimais nos cálculos.

12. Em problemas de navegação, é necessário encontrar a posição de um ponto (x, y) , através dos valores das distâncias r_1 e r_2 a dois pontos de posição conhecida (x_1, y_1) e (x_2, y_2) , como mostra a figura.



- a) Formule o problema como um sistema de equações não lineares em função das coordenadas do ponto (x, y) .
- b) Considerando $(x_1, y_1) = (10, 10)$, $(x_2, y_2) = (10, -10)$, $r_1 = 14$ e $r_2 = 16$, calcule as coordenadas do ponto (x, y) através do método iterativo de Newton considerando a aproximação inicial $(x, y)_1 = (0, 0)$. Apresente o valor ao fim de duas iterações com a correspondente estimativa do erro relativo. Use seis casas decimais nos cálculos.
13. A concentração de um poluente num lago depende do tempo t e é dada por

$$C(t) = 70e^{\beta t} + 20e^{\omega t}.$$

Efectuaram-se duas medições da concentração que foram registadas na seguinte tabela

t	1	2
$C(t)$	27.5702	17.6567

Utilize o método de Newton para determinar β e ω . Considere a aproximação inicial $(\beta, \omega)_1 = (-1.9, -0.15)$, efectue duas iterações e apresente uma estimativa do erro relativo.

4.5 Soluções

1. a) $x \approx -2$, $x \approx 0.5$ e $x \approx 1.5$.
b) $x \approx 1$.
c) $x \approx 0$.
d) $x \approx 0.5$.
2. três iterações; $x \approx 0.920524$ (com $x_1 = 0.8$ e $x_2 = 1$).
3. duas iterações; $i \approx 0.054701$.
4. a) três iterações; $x \approx 0.431128$.
b) três iterações; $x \approx 2.944958$ (não tem significado físico).
5. (a) (demonstração).
(b) duas iterações; $x \approx 3.583691$; erro relativo= 0.010255.
6. três iterações; $t \approx 5.311438$.
7. três iterações; $x \approx 11.861502$.
8. duas iterações; $x \approx (0.290677, 0.187842)^T$; erro relativo= 0.086900.
9. duas iterações; $x \approx (1.019367, -1.921783)^T$.
10. duas iterações; $(x, y) \approx (-0.694042, -1)$.
11. duas iterações; $x \approx (1.225, 0.708333)^T$.
12. a) (formulação).
b) duas iterações; $(x, y) \approx (-1.125664, 1.5)$.
13. duas iterações; $(\beta, \omega) \approx (-2.000015, -0.099987)$; erro relativo= 0.001828.

4.6 MATLAB

Para resolver equações não lineares o MATLAB tem três funções: a função **roots**, que calcula os zeros de um polinómio, a função **fzero**, que calcula a solução de qualquer equação não linear através de um método do tipo bissecção, e a função **fsolve**, que permite não só calcular a solução de equações não lineares, como também calcular a solução de sistemas de equações não lineares. Por esta última função ser mais abrangente, vai descrever-se de seguida.

4.6.1 fsolve

A função **fsolve** pode ser usada para resolver equações não lineares ou sistemas de equações não lineares, fornecendo-se ou não as primeiras derivadas da função ou funções. A sua sintaxe é

$$[x,f,exitflag,output] = fsolve('func',x0,options)$$

em que os parâmetros de saída definem a solução em **x**, o valor da função na solução **f** (que deverá ser próximo de zero), a **exitflag** define a forma como parou o algoritmo (a desejável é 1 - significa que o processo convergiu para uma raiz) e na estrutura **output** encontra-se informação sobre o processo iterativo (número de iterações, número de cálculos da função, algoritmo usado...). Os parâmetros de entrada a fornecer são a equação ou sistema de equações a resolver, que devem ser escritos numa **m-file** do tipo função. Opcionalmente, esta **m-file** poderá também conter as primeiras derivadas da função ou funções. **x0** é o valor inicial e na estrutura **options** podem alterar-se alguns dos valores que o MATLAB tem por defeito, nomeadamente definir se é pretendido usar as derivadas fornecidas na **m-file**. Para se saber que opções podem ser alteradas e quais os valores que se encontram por defeito no MATLAB, basta escrever na janela de comandos **fsolve('defaults')**. Para se alterar a estrutura **options** deve usar-se a sintaxe

$$options = optimset('param1',value1,'param2',value2,...)$$

Os valores que são *strings* devem indicar-se entre plicas (*'*).

Exemplo 4.7 Resolver a equação não linear $\sin(3x) + 4x^2 = 0$, sem fornecer derivadas, usando como ponto inicial $x_0 = 5$.

Resolução:

Basta executar os comandos

```
>>f=@(x) sin(3*x)+4*x^2
>>[x,f,extf,outp]=fsolve(f,5)
```

Em alternativa, a função **f** pode ser criada numa **m-file**.

Exemplo 4.8 Resolver a equação não linear do Exemplo 4.7, fornecendo a primeira derivada.

Primeiro cria-se uma **m-file** que contenha a função e a derivada.

Resolução:

```
function [f,d]=exEQNL(x)
f=sin(3*x)+4*x^2;
if nargout>1
    d=3*cos(3*x)+8*x;
end
```

O comando **if** aparece para tornar a função mais eficiente. Sempre que o **fsolve** precisar de recorrer ao cálculo da função, só passa ao cálculo da derivada caso isso seja solicitado. Se não se colocar o **if**, sempre que se aceder à função será calculada a função e a derivada.

Para que seja usada a derivada, não basta colocá-la na **m-file**, é necessário indicar nas opções que se pretende usá-la, e isso faz-se com a opção **Jacobian**. Assim, a primeira coisa a fazer é alterar a estrutura das opções:

```
>>op=optimset('Jacobian','on');
A sequência de comandos deverá ser
>>x0=5;
>>[x,f,extf,outp]=fsolve('exEQNL',x0,op)
```

Exemplo 4.9 Resolver o sistema não linear,

$$\begin{cases} 2x_1 - x_2 - e^{-x_1} = 0 \\ -x_1 + 2x_2 - e^{-x_2} = 0 \end{cases}$$

fornecendo a matriz do Jacobiano e tomando para valores iniciais o vetor $(-5, -5)$.

Resolução:

A forma de resolver este problema é a mesma que para o Exemplo 4.8. Cria-se a **m-file** com as funções e a matriz do Jacobiano.

```
function [f,J]=exSEQNL(x)
f=[2*x(1)-x(2)-exp(-x(1));-x(1)+2*x(2)-exp(-x(2))];
if nargin>1
    J=[2+exp(-x(1)) -1;-1 2+exp(-x(2))];
end
```

A sequência de comandos é

```
>>op=optimset('Jacobian','on');
>>x0=[5 5];
>>[x,f,extf,outp]=fsolve('exSEQNL,x0,op)
```


4.6.2 Exercícios para resolver em MATLAB

1. Calcule um zero da função

$$f(x) = e^x - x^2 - 2x - 2,$$

usando $x^{(1)} = 2$.

2. Resolva a seguinte equação não linear, fornecendo as primeiras derivadas.

$$\cos(x) - \cos(3.1x) = 0.$$

Considere

a) $x^{(1)} = -1$;

b) $x^{(1)} = 1$;

c) $x^{(1)} = -10$;

d) $x^{(1)} = 10$.

3. Resolva o seguinte sistema de equações não lineares nas variáveis x_1 e x_2 .

$$\begin{cases} \sin\left(\frac{x_1+x_2}{2}\right) = 2x_1 \\ \cos\left(\frac{x_1-x_2}{2}\right) = 2x_2 \end{cases}$$

a) Considere uma aproximação inicial $x^{(1)} = (0, 0)^T$;

b) Repita com $x^{(1)} = (1, 2)^T$;

c) Repita a alínea anterior fornecendo a matriz do Jacobiano.

4. Determine a solução do sistema de equações não lineares, fornecendo a matriz do Jacobiano,

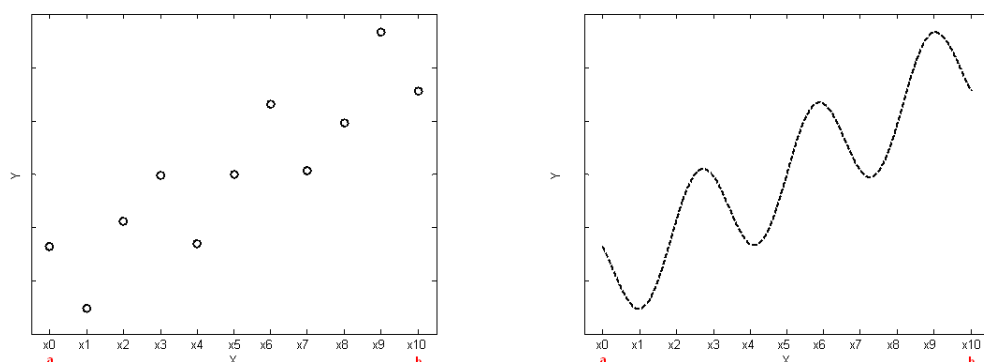
$$\begin{cases} x_1 = 0 \\ x_2^2 + x_2 = 0 \\ e^{x_3} - 1 = 0 \end{cases}$$

usando como aproximação inicial o vetor $x^{(1)} = (1, 1, -1)^T$.

Capítulo 5

Polinómio interpolador de Newton

O polinómio interpolador é uma das técnicas disponíveis para a aproximação de funções. O objetivo é encontrar uma aproximação, neste caso concreto através de um polinómio, $p_n(x)$, à função dada, $f(x)$, com o menor erro possível. Faz-se este tipo de aproximação em duas situações:



(a) Função dada por um conjunto de 11 pontos (b) Função dada por uma expressão matemática

Figura 5.1: Aproximação de funções.

1. Dado um conjunto discreto de valores (Figura 5.1(a))

$$(x_i, f_i), \quad i = 0, 1, \dots, n \quad (n+1 \text{ pontos})$$

pretende-se encontrar uma relação funcional (expressão matemática) entre as variáveis x e f para prever o comportamento entre as variáveis e estimar valores, em que x é a variável independente e f é a variável dependente.

2. Dada uma função $f(x)$ por uma expressão matemática complicada (Figura 5.1(b)), pretende-se conhecer uma expressão mais simples que descreva o melhor possível o comportamento de f como função de x .

5.1 Erro da aproximação

Teorema 5.1.1 Teorema de Weirstrass

Dadas a função $f(x)$, contínua num intervalo $[a, b]$, e uma quantidade $\varepsilon > 0$, existe sempre um polinómio $p_n(x)$, de grau menor ou igual a n , tal que o erro da aproximação $\|e_n(x)\| = \|f(x) - p_n(x)\| \leq \varepsilon$.

Pelo Teorema 5.1.1 pode assegurar-se que o erro seja igual a zero para um conjunto de $n + 1$ pontos seleccionados do intervalo $[a, b]$, isto é, o polinómio passa por esses $n + 1$ pontos da função,

$$f_i \equiv f(x_i) = p_n(x_i), \quad \text{para } i = 0, 1, \dots, n$$

e que este polinómio é único e de grau menor ou igual a n .

Os polinómios interpoladores mais conhecidos são

- o polinómio interpolador de Newton baseado em diferenças divididas,
- o polinómio interpolador de Lagrange,

que são ambos polinómios de colocação.

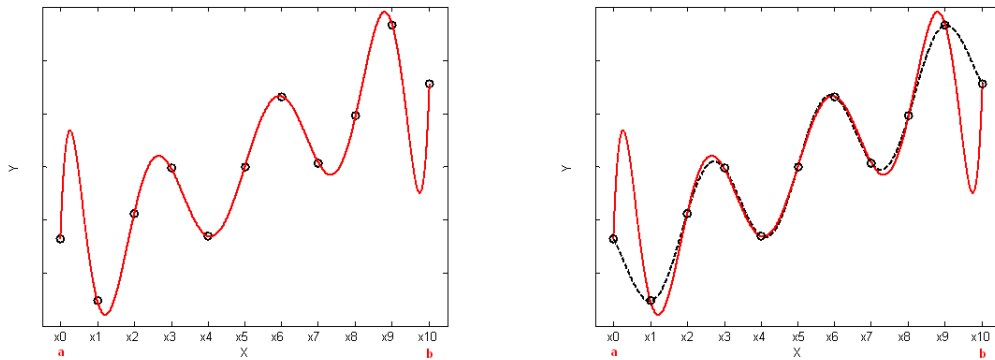
Pode ainda definir-se o polinómio dos mínimos quadrados, em que se assegura que a soma dos quadrados dos erros é mínima no intervalo $[a, b]$, isto é,

$$\min \sum_{i=0}^n (e_n(x))^2 \equiv \sum_{i=0}^n (f_i - p_n(x_i))^2,$$

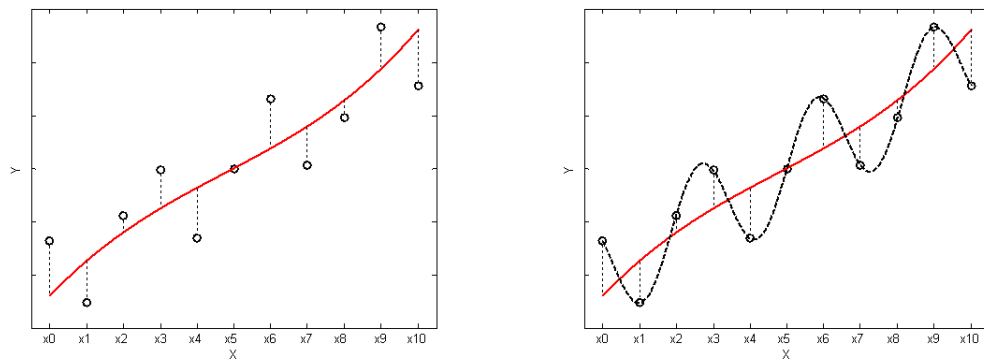
mas este assunto será discutido em detalhe noutra capítulo. Pode ver-se na Figura 5.2 a diferença entre usar-se um polinómio interpolador (Figuras 5.2(a) e 5.2(b)) ou um polinómio dos mínimos quadrados (Figuras 5.2(c) e 5.2(d)).

Deve escolher-se um ou outro método de acordo com a situação que se tem em mãos.

Se os dados forem precisos, isto é, não contêm erros de observação, é mais vantajoso usar-se uma função que passe pelos pontos dados, e por isso o método mais adequado, neste caso, é o polinómio interpolador.



(a) Polinómio interpolador de grau 10 (conjunto de 11 pontos) (b) Polinómio interpolador de grau 10 (função dada por uma expressão matemática)



(c) Polinómio de grau 3 dos mínimos quadrados (conjunto de 11 pontos) (d) Polinómio de grau 3 dos mínimos quadrados (função dada por uma expressão matemática)

Figura 5.2: Aproximação polinomial.

Se, pelo contrário, os dados possuem erros de observação, torna-se mais vantajoso encontrar uma função que descreva o comportamento dos dados, sem a preocupação da curva passar pelos pontos, e por isso deve usar-se o polinómio dos mínimos quadrados. Esta situação ocorre, por exemplo, quando se estão a recolher dados experimentais associados a um aparelho com um erro associado.

Deve ainda ter-se em conta que quando o número de pontos disponível é muito grande, o polinómio interpolador será de um grau muito elevado, tornando-se, por isso, muito ruidoso e irregular, e o seu uso é também, neste caso, desaconselhável. O que se faz muitas vezes é seleccionar um número limitado de pontos na região que interessa interpolar. Podendo assim usar-se um polinómio interpolador de grau mais baixo.

5.2 Diferenças divididas

5.2.1 Definição

Considere-se um conjunto de $n + 1$ pontos diferentes entre si $(x_0, f_0), (x_1, f_1), \dots, (x_n, f_n)$.

Podem definir-se as diferenças divididas como

- Diferenças divididas de 1ª ordem

$$[x_0, x_1] = \frac{f_0 - f_1}{x_0 - x_1} = \frac{f_1 - f_0}{x_1 - x_0}$$

...

$$[x_{n-1}, x_n] = \frac{f_{n-1} - f_n}{x_{n-1} - x_n} = \frac{f_n - f_{n-1}}{x_n - x_{n-1}}$$

- Diferenças divididas de 2ª ordem

$$[x_0, x_1, x_2] = \frac{[x_0, x_1] - [x_1, x_2]}{x_0 - x_2} = \frac{[x_1, x_2] - [x_0, x_1]}{x_2 - x_0}$$

...

$$[x_{n-2}, x_{n-1}, x_n] = \frac{[x_{n-2}, x_{n-1}] - [x_{n-1}, x_n]}{x_{n-2} - x_n} = \frac{[x_{n-1}, x_n] - [x_{n-2}, x_{n-1}]}{x_n - x_{n-2}}$$

- ...

- Diferença dividida de nª ordem

$$[x_0, x_1, \dots, x_{n-1}, x_n] = \frac{[x_0, x_1, \dots, x_{n-1}] - [x_1, x_2, \dots, x_n]}{x_0 - x_n} = \frac{[x_1, x_2, \dots, x_n] - [x_0, x_1, \dots, x_{n-1}]}{x_n - x_0}$$

As diferenças divididas podem ser colocadas numa tabela com a forma

x_0	f_0				
		$[x_0, x_1]$			
x_1	f_1		$[x_0, x_1, x_2]$		
		$[x_1, x_2]$		$[x_0, x_1, x_2, x_3]$	
x_2	f_2		$[x_1, x_2, x_3]$		\dots
		$[x_2, x_3]$		$[x_1, x_2, x_3, x_4]$	$[x_0, x_1, \dots, x_n]$
x_3	f_3		\dots		\dots
\vdots	\vdots	\dots		\dots	
x_{n-1}	f_{n-1}		$[x_{n-2}, x_{n-1}, x_n]$		
		$[x_{n-1}, x_n]$			
x_n	f_n				

5.2.2 Propriedades das diferenças divididas

As diferenças divididas gozam de algumas propriedades importantes que a seguir se listam.

1. Podem ser calculadas para qualquer espaçamento, mesmo sendo não constante, entre os pontos disponíveis $x_0, x_1, \dots, x_{n-1}, x_n$.

2. As diferenças divididas são funções simétricas dos seus argumentos, isto é

$$[x_0, x_1] = [x_1, x_0]$$

$$[x_0, x_1, x_2] = [x_2, x_1, x_0]$$

\dots

3. As diferenças divididas de n^{a} ordem de um polinómio de grau n são iguais entre si e diferentes de zero. Por consequência, as diferenças divididas de $(n+1)^{\text{a}}$ ordem são iguais a zero neste caso.

5.3 Polinómio interpolador de Newton baseado em diferenças divididas

O polinómio de grau menor ou igual a n , $p_n(x)$, é construído com base em $n+1$ pontos e é único.

Para simplificar a notação, considere-se $f_i \equiv f(x_i)$. Sejam os $n + 1$ pontos:

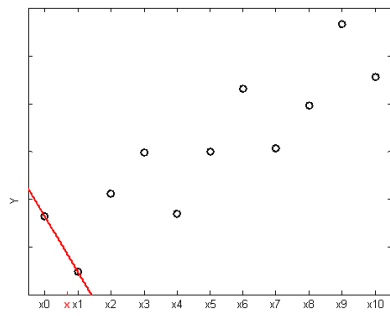
x_0	x_1	x_2	x_3	x_4	\cdots	x_{n-2}	x_{n-1}	x_n
f_0	f_1	f_2	f_3	f_4	\cdots	f_{n-2}	f_{n-1}	f_n

O polinómio interpolador de Newton de grau $\leq n$ é

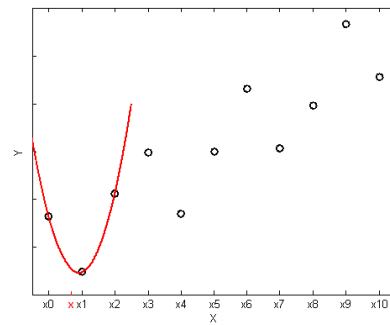
$$p_n(x) = f_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3] \\ + \cdots + (x - x_0)(x - x_1)(x - x_2) \cdots (x - x_{n-1})[x_0, x_1, x_2, \dots, x_{n-1}, x_n]$$

5.3.1 Interpolação direta

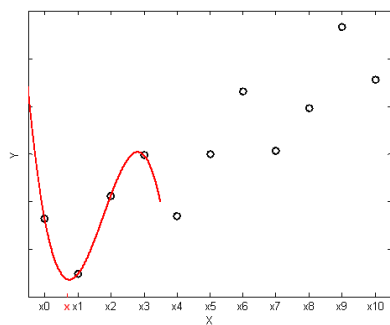
O objetivo da interpolação direta é estimar o valor de $f(\bar{x})$, sendo \bar{x} um ponto que não está na tabela.



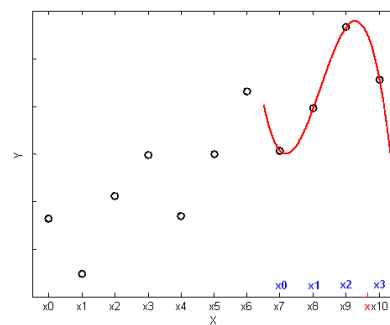
(a) $f(x) \approx p_1(x)$



(b) $f(x) \approx p_2(x)$



(c) $f(x) \approx p_3(x)$



(d) $f(x) \approx p_3(x)$

Figura 5.3: Alguns exemplos de polinómios interpoladores

Para construir um polinómio de grau n , devem escolher-se $n + 1$ pontos da tabela de pontos disponíveis, garantindo que pelo menos um dos pontos está à direita de \bar{x} e outro à

sua esquerda, se isso for possível. Os restantes pontos da tabela escolhem-se de forma a que estejam o mais próximo possível de \bar{x} .

Na Figura 5.3(a) apresentam-se os pontos selecionados (dois) e o polinómio de grau um resultante para estimar o valor da função no ponto representado a vermelho. Na Figura 5.3(b) apresenta-se a mesma situação mas para um polinómio de grau dois (três pontos). Nas Figuras 5.3(c) e 5.3(d) apresentam-se dois polinómios de grau três diferentes (quatro pontos) para dois pontos interpoladores diferentes, representados a vermelho.

No Exemplo 5.1 é apresentada a forma como os pontos que entram na construção de um polinómio interpolador devem ser escolhidos.

Exemplo 5.1 *Escolha de quatro pontos para construir um polinómio de grau três.*

Dada a tabela

x_i	-1	0	1	2	8	10	12	15	20
f_i	-5	-2	-1	3	0	-2	-1	4	6

que pontos devem ser escolhidos para construir um polinómio de grau três se o ponto interpolador for

a) $\bar{x} = 3$?

$$\begin{array}{cccc|c} x_0 & x_1 & x_2 & x_3 & \\ \hline f_0 & f_1 & f_2 & f_3 & \end{array} \Leftrightarrow \begin{array}{cccc|c} 0 & 1 & 2 & 8 & \\ \hline -2 & -1 & 3 & 0 & \end{array}$$

b) $\bar{x} = 13$?

$$\begin{array}{cccc|c} x_0 & x_1 & x_2 & x_3 & \\ \hline f_0 & f_1 & f_2 & f_3 & \end{array} \Leftrightarrow \begin{array}{cccc|c} 8 & 10 & 12 & 15 & \\ \hline 0 & -2 & -1 & 4 & \end{array}$$

5.4 Erro de truncatura

O erro de truncatura do polinómio interpolador é dado por

$$e_n(x) = f(x) - p_n(x) = (x - x_0)(x - x_1) \cdots (x - x_{n-1})(x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!},$$

com $\xi \in [a, b]$.

O majorante do erro de truncatura cometido com a aproximação, para um certo x do intervalo $[a, b]$, que contém os pontos usados para construir o polinómio de grau n , $x_0, x_1, x_2, \dots, x_{n-1}, x_n$, pode ser estimado de duas formas:

1. Se $f(x)$ for dada por uma expressão, então

$$|e_n(x)| \leq |(x - x_0)(x - x_1) \cdots (x - x_{n-1})(x - x_n)| \frac{1}{(n+1)!} M_{n+1}$$

em que

$$\left| \left[f^{(n+1)}(x) \right]_{[a,b]} \right| \leq M_{n+1}.$$

2. Se $f(x)$ for dada por um conjunto discreto de pontos,

$$|e_n(x)| \leq |(x - x_0) \cdots (x - x_{n-1})(x - x_n)| |(\text{dd de ordem } n+1)|$$

em que

$$(\text{dd de ordem } n+1) = [x_0, x_1, \dots, x_{n-1}, x_n, x_z]$$

se só existir uma, ou a maior delas em valor absoluto se existirem mais que uma.

Exemplo 5.2 Dada a tabela de valores de uma função $f(x)$

x_i	0.0	0.1	0.2	0.3	0.4	0.5	0.8	1.0
$f(x_i)$	0	1	1	2	2	3	3	4

- a) Pretende-se aproximar $f(0.6)$ usando um polinómio de grau 3. Use a fórmula interpoladora de Newton baseada em diferenças divididas.
- b) Estime o erro de truncatura cometido na alínea anterior.

Resolução:

- a) Para estimar $f(0.6)$ usando um polinómio de grau 3 são necessários 4 pontos. A tabela das diferenças divididas para os 4 pontos mais próximos de 0.6 é

x_i	f_i	$dd1$	$dd2$	$dd3$
0.3	2			
		0		
0.4	2		50	
		10		-150
0.5	3		-25	
		0		
0.8	3			

$$p_3(x) = 2 + 50(x - 0.3)(x - 0.4) - 150(x - 0.3)(x - 0.4)(x - 0.5).$$

A aproximação para $f(0.6)$ é

$$f(0.6) \approx p_3(0.6) = 4.1.$$

b) Tem de se acrescentar à tabela anterior o ponto mais próximo do ponto interpolador que ainda não tenha sido usado no cálculo do polinómio, por exemplo, $x = 1$. Poder-se-ia usar também o ponto $x = 0.2$, já que está à mesma distância do ponto interpolador.

x_i	f_i	$dd1$	$dd2$	$dd3$	$dd4$
0.3	2				
		0			
0.4	2		50		
		10		-150	
0.5	3		-25		297.619047
		0		58.333333	
0.8	3		10		
		5			
1.0	4				

O erro de truncatura é dado por

$$\begin{aligned}|e_3| &\leq |(x - x_0)(x - x_1)(x - x_2)(x - x_3))| \times |dd4| \\ &= |(0.6 - 0.3)(0.6 - 0.4)(0.6 - 0.5)(0.6 - 0.8))| \times 297.619047 = 0.357143\end{aligned}$$

5.5 Exercícios

1. Dada a tabela de valores de uma função $f(x)$

x_i	0.0	0.1	0.2	0.3	0.4	0.5	0.8	1.0
$f(x_i)$	0	1	1	2	2	3	3	4

- a) Pretende-se aproximar $f(0.6)$ usando um polinómio de grau 3. Use a fórmula interpoladora de Newton baseada em diferenças divididas.
- b) Estime o erro de truncatura cometido na alínea anterior.
- c) Estime $f(0.6)$ usando todos os pontos da tabela.
2. A tabela seguinte apresenta a população dos Estados Unidos da América (em milhões) de 1940 e 1980.

ano	1940	1950	1960	1970	1980
população	132.165	151.326	179.323	203.302	226.542

- a) Construa o polinómio interpolador de Newton de grau 4 para estimar a população no ano 1965.
- b) A população em 1930 foi 123.203. Qual a precisão do valor calculado na alínea a)?
3. Os registos efectuados numa linha de montagem são os seguintes:

nº de unidades	1	3	4	6	7	10
horas necessárias	2	3	4	5	6	10

- a) Tendo sido recebidos pedidos para a montagem de 2 unidades e 8 unidades, use interpolação cúbica para estimar o tempo (em horas) necessário para satisfazer cada pedido.
- b) Calcule uma estimativa do erro de truncatura cometido na alínea anterior para cada um dos pedidos.

4. Considere a seguinte tabela de uma função polinomial

x	-1	0	1	2	3	4
$p(x)$	-1	-3	-1	5	15	29

Sem recorrer à expressão analítica de $p(x)$:

- a) mostre que $p(x)$ é um polinómio interpolador de grau 2.
 b) determine $p(10)$.

5. Considere a tabela de valores da função $f(x)$

x_i	0	1	3	4
$f(x_i)$	a	2	4	b

Determine a e b por forma a que o polinómio interpolador de Newton que aproxima f seja de grau 3, com coeficiente do termo de maior grau igual à unidade e coeficiente do termo de menor grau igual a zero. Escreva o polinómio.

6. Considere a seguinte tabela da função $f(x)$.

x	-2	-1	0	1	2
$f(x)$	a	2	1	0	4

Determine $a \in \mathbb{R}$ de modo a que $f(x)$ seja um polinómio de grau 3.

5.6 Soluções

1. a) $f(0.6) \approx p_3(0.6) = 4.1$.
b) $|e_T| \leq 0.3571$.
c) $f(0.6) \approx p_7(0.6) = 7.2$.
2. a) $f(1965) \approx p_4(1965) = 191.990625$.
b) $|e_T| \leq 0.28125$.
3. a) $f(2) \approx p_4(2) = 2.1999$, $f(8) \approx p_4(8) = 7.2224$.
b) $|e_T(x=2)| \leq 0.2$, $|e_T(x=8)| \leq 0.2224$.
4. a) As diferenças divididas de grau dois são todas iguais entre si e diferentes de zero.
b) $p(10) = 197$.
5. $a = 0$, $b = 16$ e $p_3(x) = 2x - 0.3333x(x-1) + x(x-1)(x-3)$.
6. $a = -2$.

5.7 MATLAB

5.7.1 polyfit

A função `polyfit` ajusta um polinómio de grau n a um conjunto de dados, no sentido dos mínimos quadrados. No entanto, pode ser usado para determinar um polinómio interpolador, sendo que, neste caso, terão de ser fornecidos exatamente $n+1$ pontos, que devem ser escolhidos de forma adequada, como foi descrito nas secções anteriores deste capítulo.

A sintaxe desta função é

$$\mathbf{p} = \text{polyfit}(\mathbf{x}, \mathbf{y}, \mathbf{n})$$

em que \mathbf{p} devolve os coeficientes do polinómio em potências descendentes. \mathbf{x} e \mathbf{y} são dois vetores que contêm os dados que se pretendem aproximar e \mathbf{n} é o grau do polinómio.

5.7.2 polyval

Para se fazer a estimativa obtida pelo polinómio interpolador calculado num ponto, ou num conjunto de pontos, usa-se a função `polyval`, cuja sintaxe é

$$\mathbf{Y} = \text{polyval}(\mathbf{p}, \mathbf{X})$$

em que \mathbf{p} é o vetor que contém os coeficientes do polinómio obtidos através da função `polyfit` e \mathbf{X} é o valor ou um vetor contendo um conjunto de valores onde se pretende fazer a estimativa usando o polinómio calculado.

Exemplo 5.3 *É dada uma tabela para aproximar os dados a um polinómio interpolador de grau dois:*

x_i	1	2	4	6	7	8
f_i	10	5	8	4	3	7

Pretende estimar-se o valor de f no ponto $x = 5$.

A primeira coisa a fazer é escolher os $n + 1 = 3$ pontos adequados a esta estimativa. Usando as regras descritas, os pontos escolhidos são 4, 6 e 7. Estes pontos são introduzidos no MATLAB como vetor x , assim como os respetivos valores da função, que serão colocados no vetor y .

Usando o comando `polyfit`, de acordo com a sintaxe definida


```
>>p2=polyfit(x,y,2)
```

o MATLAB devolve o vetor p

```
p2 =
```

```
    0.3333    -5.3333    24.0000
```

com o qual se pode construir o polinómio de grau dois

$$p_2(x) = 0.3333x^2 - 5.3333x + 24.0000.$$

Para se obter a estimativa no ponto interpolador $x = 5$, usa-se o comando polyval:

```
>> polyval(p2,5)
```

5.7.3 Exercícios para resolver em MATLAB

1. Dada a tabela de valores de uma função $f(x)$,

x_i	5.0	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
f_i	0.0639	0.0800	0.0988	0.1203	0.1442	0.1714	0.2010	0.2330	0.2673	0.3036	0.3414

Para aproximar $f(5.44)$,

- apresente o polinómio interpolador de grau 2;
 - com base no polinómio anterior, estime $f(5.44)$;
 - apresente o polinómio interpolador de grau 5;
 - com base no polinómio anterior, estime $f(5.44)$;
 - apresente o polinómio interpolador de grau 10;
 - com base no polinómio anterior, estime $f(5.44)$.
2. Considere a tabela seguinte de 12 valores de $f(x)$.

x_i	0.00	0.30	0.50	0.70	0.90	1.00	1.20	1.50	1.60	1.75	2.00	2.10
f_i	0.0000	0.2955	0.4794	0.6442	0.7833	0.8415	0.9320	0.9975	0.9996	0.9840	0.9093	0.8632

Para aproximar $f(1.57)$

- apresente o polinómio interpolador usando quatro pontos;
- com base no polinómio anterior, estime $f(1.57)$;
- apresente o polinómio interpolador usando seis pontos;
- com base no polinómio anterior, estime $f(1.57)$;
- apresente o polinómio interpolador usando 12 pontos;
- com base no polinómio anterior, estime $f(1.57)$;
- represente graficamente os pontos e o polinómio de grau 11.

3. A velocidade de ascensão de um foguetão, $v(t)$, é conhecida para diferentes tempos conforme a seguinte tabela. Esta velocidade pode ser estimada a partir de um polinómio interpolador de grau três.

$t(\text{s})$	0	5	10	15	20	30
$v(t)(\text{m/s})$	0	106.8	227.04	362.78	517.35	901.67

- a) Calcule o polinómio e estime a velocidade do foguetão para $t = 8$ s.
- b) Represente graficamente os pontos o polinómio calculado.
4. Considere um reservatório de água com 2.1 m de altura. No início, o reservatório está cheio de água. Num certo instante, abre-se a a válvula e o reservatório começa a ser esvaziado. A altura (em metros) de água do reservatório, t horas depois de este ter começado a ser esvaziado, é dada por $h(t)$, de acordo com a tabela

instante, t_i	0	1	4	7	8	10	14
altura de água, $h(t)_i$	2.1	2.0	1.8	1.5	1.4	1.1	0

Pretende estimar-se a altura de água no reservatório ao fim de 5h.

- a) apresente o polinómio interpolador de grau 2;
- b) com base no polinómio anterior, estime $f(5)$;
- c) apresente o polinómio interpolador de grau 5;
- d) com base no polinómio anterior, estime $f(5)$;
- e) apresente o polinómio interpolador de grau 6;
- f) com base no polinómio anterior, estime $f(5)$.

Capítulo 6

Interpolação segmentada - 'spline'

Uma 'spline' é uma régua de madeira utilizada para traçar curvas suaves entre dois pontos dados.

Esta técnica surgiu nos anos 40 do século XX na engenharia náutica, para a elaboração de trajetórias de grandes navios. Estas devem ser curvas suaves que passam pelos vários pontos de paragem. É muito usada também na indústria naval para apurar a forma dos cascos a partir de esboços grosseiros e na área da robótica para a definição de trajetórias de movimentos de robôs. Na informática, as 'splines' são a base da gráfica computacional.

A grande vantagem de se usarem 'splines', comparativamente aos polinómios interpoladores, é evitar o ruído que surge quando se usam muitos pontos, pelo facto de o polinómio obtido ser de grau elevado (Figura 6.1).

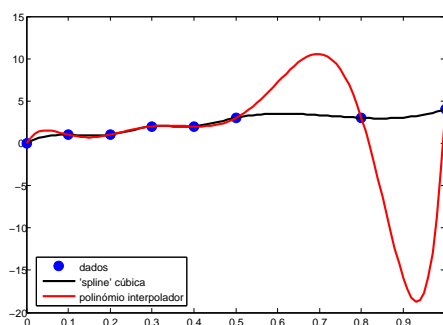


Figura 6.1: 'Spline' cúbica, $s_3(x)$, e polinómio interpolador de grau sete, $p_7(x)$, obtidos para um conjunto de oito pontos.

6.1 Definição

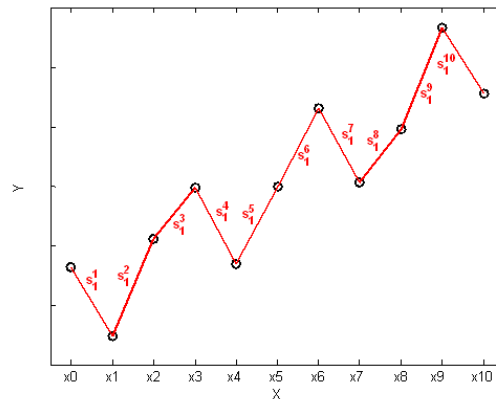
Uma 'spline' é uma função segmentada (definida por segmentos), isto é, é formada por vários polinómios ligados uns aos outros de uma forma contínua e suave.

Dado um conjunto de $n + 1$ pontos, $x_0, x_1, \dots, x_{n-1}, x_n$, chamam-se nós interiores aos pontos x_1, \dots, x_{n-1} , sendo x_0 e x_n os nós exteriores ou fronteiras.

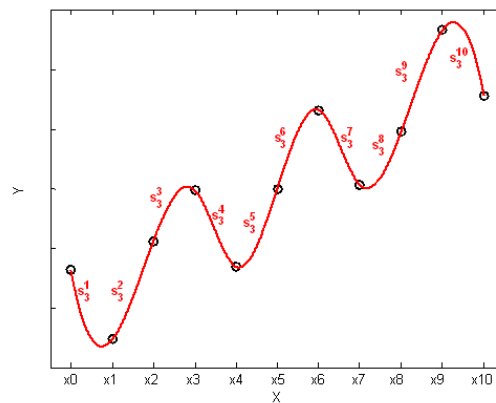
Uma função $s_k(x)$, com k inteiro e não negativo, chama-se 'spline' de grau k se possuir as seguintes propriedades:

- $s_k(x)$ é uma função continuamente diferenciável até à ordem $k - 1$;
- $s_k^i(x)$ é um polinómio de grau k , em cada segmento i , para $x \in [x_{i-1}, x_i]$, $1 \leq i \leq n$.

Exemplo 6.1 'Spline' linear, $s_1(x)$



Exemplo 6.2 'Spline' cúbica, $s_3(x)$



6.2 'Spline' linear

6.2.1 Definição

Uma 'spline' linear é formada pela ligação de polinómios de grau um, em que o segmento i é definido por $[x_{i-1}, x_i]$. Em cada um destes segmentos, o polinómio de grau um obtém-se através de

$$s_1^i(x) = f_{i-1} + \frac{f_i - f_{i-1}}{x_i - x_{i-1}}(x - x_{i-1}),$$

com $i = 1, 2, \dots, n$ e em que $f_i \equiv f(x_i)$.

6.2.2 Limite superior do erro de truncatura

Seja $f(x)$ contínua, com derivadas contínuas até à segunda ordem. Sejam os pontos do intervalo $[a, b]$ tais que $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$. Seja ainda $s_1(x)$ a 'spline' linear composta pelos polinómios de grau um $s_1^i(x)$, $i = 1, 2, \dots, n$, para aproximar $f(x)$ em $[a, b]$. O limite superior do erro de truncatura em valor absoluto cometido por esta aproximação é dado por

$$|f(x) - s_1(x)| \leq \frac{1}{8}h^2 M_2.$$

h é o espaçamento máximo entre os pontos que foram usados para construir a 'spline',

$$h = \max_{0 \leq i \leq n-1} (x_{i+1} - x_i),$$

e M_2 é valor absoluto do majorante da segunda derivada de $f(x)$ em $[a, b]$,

$$\max_{\xi \in [a, b]} |f''(\xi)| \leq M_2.$$

Se $f(x)$ não for dada por uma expressão matemática, substitui-se M_2 pela diferença dividida de segunda ordem de maior módulo em valor absoluto, multiplicada por 2!.

6.3 'Spline' cúbica

6.3.1 Definição

Uma 'spline' cúbica $s_3(x)$ é formada pela ligação suave dos polinómios de grau três,

$$s_3(x) = \begin{cases} s_3^1(x) & x \in [x_0, x_1] \text{ (para o segmento 1)} \\ s_3^2(x) & x \in [x_1, x_2] \text{ (para o segmento 2)} \\ \vdots & \vdots \\ s_3^n(x) & x \in [x_{n-1}, x_n] \text{ (para o segmento } n\text{)}. \end{cases}$$

A forma deste polinómio em cada segmento i é dada por

$$\begin{aligned} s_3^i(x) = & \frac{M_{i-1}}{6(x_i - x_{i-1})}(x_i - x)^3 + \frac{M_i}{6(x_i - x_{i-1})}(x - x_{i-1})^3 \\ & + \left[\frac{f_{i-1}}{x_i - x_{i-1}} - \frac{M_{i-1}(x_i - x_{i-1})}{6} \right] (x_i - x) \\ & + \left[\frac{f_i}{x_i - x_{i-1}} - \frac{M_i(x_i - x_{i-1})}{6} \right] (x - x_{i-1}), \end{aligned} \quad (6.1)$$

em que $i = 1, 2, \dots, n$, $f_i \equiv f(x_i)$ e $M_i \equiv M(x_i)$ representa a curvatura da 'spline' em x_i , isto é, o valor da segunda derivada da 'spline' no nó x_i .

Para que a ligação entre os vários segmentos seja suave, tem de haver continuidade nos nós e tem de se manter a curvatura. Isto significa que em cada nó interior x_i , $i = 1, 2, \dots, n-1$ tem de se verificar

- $s_3^i(x_i) = s_3^{i+1}(x_i)$,
- $s_3^{i'}(x_i) = s_3^{i+1'}(x_i)$,
- $s_3^{i''}(x_i) = s_3^{i+1''}(x_i)$.

Pode verificar-se em (6.1) que $s_3^i(x)$ depende de M_{i-1} e M_i , sendo que no primeiro segmento ($i = 1$), é necessário conhecer M_0 e M_1 , no segundo segmento ($i = 2$) é necessário conhecer M_1 e M_2 , e assim sucessivamente, sendo que para $i = n$ é necessário conhecer M_{n-1} e M_n .

Por se exigir continuidade da primeira derivada nos $n-1$ nós interiores, para cada um

destes nós resulta a equação

$$\begin{aligned} (x_i - x_{i-1})M_{i-1} + 2(x_{i+1} - x_{i-1})M_i + (x_{i+1} - x_i)M_{i+1} = \\ = \frac{6}{x_{i+1} - x_i}(f_{i+1} - f_i) - \frac{6}{x_i - x_{i-1}}(f_i - f_{i-1}), \end{aligned} \quad (6.2)$$

$i = 1, 2, \dots, n-1$. Estas $n-1$ equações nas $n+1$ incógnitas $M_0, M_1, M_2, M_{n-1}, M_n$ definem um sistema linear tridiagonal, que deve ser resolvido por EGPP. No entanto, este sistema tem duas incógnitas a mais que o número de equações. Para que o sistema seja possível e determinado, é necessário adicionar mais duas equações ao sistema. Estas vão depender do tipo de 'spline' cúbica.

6.3.2 'Spline' cúbica natural

Se a 'spline' cúbica for natural, a curvatura da 'spline' nos extremos é nula, o que significa que

$$s_3^{1''}(x_0) = 0 \text{ e } s_3^{n''}(x_n) = 0,$$

ou seja

$$M_0 = 0 \text{ e } M_n = 0.$$

Assim, para uma 'spline' cúbica natural, o sistema definido por (6.2) tem $n-1$ equações nas $n-1$ incógnitas M_1, M_2, \dots, M_{n-1} .

6.3.3 'Spline' cúbica completa

Para uma 'spline' cúbica completa, a curvatura nos extremos é não nula, pelo que é necessário acrescentar ao sistema resultante de (6.2) mais duas equações, que dizem respeito ao nó da fronteira inferior x_0 (6.3) e ao nó da fronteira superior x_n (6.4).

$$2(x_1 - x_0)M_0 + (x_1 - x_0)M_1 = \frac{6}{x_1 - x_0}(f_1 - f_0) - 6f'(x_0), \quad (6.3)$$

$$2(x_n - x_{n-1})M_n + (x_n - x_{n-1})M_{n-1} = 6f'(x_n) - \frac{6}{x_n - x_{n-1}}(f_n - f_{n-1}). \quad (6.4)$$

O cálculo destas equações envolve o cálculo das derivadas nos extremos, f'_0 e f'_n . Se $f(x)$ for dada por uma expressão, calcula-se $f'(x)$ e assim $f'_0 \equiv f'(x_0)$ e $f'_n \equiv f'(x_n)$. Se a expressão de

$f(x)$ não for conhecida, as primeiras derivadas $f'(x_0)$ e $f'(x_n)$ têm de ser estimadas recorrendo às diferenças divididas de primeira ordem, ou seja,

$$f'_0 = \frac{f(A) - f_0}{A - x_0} \quad \text{e} \quad f'_n = \frac{f_n - f(B)}{x_n - B}$$

com $x_0 < A$ e $B < x_n$. Quanto mais próximo A estiver de x_0 , melhor é a aproximação a $f'(x_0)$. De igual modo, quanto mais próximo B estiver de x_n , melhor é a aproximação a $f'(x_n)$. Em termos práticos, isto significa que A e B são o segundo e o penúltimo pontos, respetivamente, do conjunto de pontos dado. Ao reservar A e B para calcular a aproximação às primeiras derivadas, não é aconselhável incluir os pares $(A, f(A))$ e $(B, f(B))$ na construção da 'spline'.

A 'spline' cúbica completa tem curvatura nos extremos, iniciando-se e terminando de forma suave, ao passo que a 'spline' cúbica natural não tem curvatura nos extremos, iniciando-se e terminando de forma abrupta (Figura 6.2).

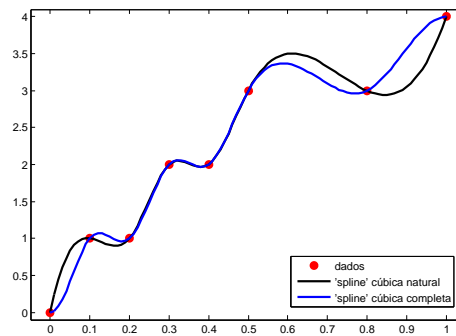


Figura 6.2: 'Spline' cúbica natural e 'spline' cúbica completa para um conjunto de oito pontos.

6.3.4 Limite superior do erro de truncatura

Seja $f(x)$ contínua, com derivadas contínuas até à quarta ordem. Sejam os pontos do intervalo $[a, b]$ tais que $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$. Seja ainda $s_3(x)$ a 'spline' cúbica composta pelos polinómios de grau três $s_3^i(x)$, $i = 1, 2, \dots, n$, para aproximar $f(x)$ em $[a, b]$. O limite superior do erro de truncatura em valor absoluto cometido por esta aproximação é dado por

$$|f(x) - s_3(x)| \leq \frac{5}{384} h^4 M_4$$

e

$$|f'(x) - s'_3(x)| \leq \frac{1}{24} h^3 M_4.$$

h é o espaçamento máximo entre os pontos que foram usados para construir a 'spline',

$$h = \max_{0 \leq i \leq n-1} (x_{i+1} - x_i),$$

e M_4 é o valor absoluto do majorante da quarta derivada de $f(x)$ em $[a, b]$,

$$\max_{\xi \in [a, b]} |f^{iv}(\xi)| \leq M_4.$$

Se $f(x)$ não for dada por uma expressão matemática, substitui-se M_4 pela diferença dividida de quarta ordem de maior módulo em valor absoluto, multiplicada por $4!$.

Exemplo 6.3 Num certo campeonato regional de futebol há 7 equipas. No fim da temporada, o número de pontos ganhos e o número de golos sofridos por 6 das equipas estão representados na tabela

Equipa	F.C.Sol	F.C.Lá	S.C.Gato	Nova F.C.	Vila F.C.	F.C.Chão
Nº de pontos, x_i	10	12	18	27	30	34
Nº de golos, $f(x_i)$	20	18	15	9	12	10

- a) Use uma 'spline' cúbica completa para descrever a relação entre o número de pontos e o número de golos sofridos pelas equipas no campeonato. Sabendo que a 7ª equipa terminou o campeonato com 29 pontos, estime o número de golos que terá sofrido.
- b) Calcule uma estimativa do erro de truncatura cometido na alínea anterior.

Resolução:

- a) Como a 'spline' cúbica é completa e não é conhecida uma expressão analítica que represente a função, retiram-se o segundo ($A = 12$) e penúltimo ($B = 30$) pontos da tabela, já que serão usados como pontos auxiliares para calcular uma aproximação por diferenças divididas às derivadas nos extremos. Assim,

$$f'_0 = \frac{20 - 18}{10 - 12} = -1 \quad e \quad f'_n = \frac{12 - 10}{30 - 34} = -0.5.$$

Restam 4 pontos $\Rightarrow n = 3 \Rightarrow 2$ pontos interiores.

- $i = 0$

$$\begin{aligned} 2(x_1 - x_0)M_0 + (x_1 - x_0)M_1 &= \frac{6}{x_1 - x_0}(f_1 - f_0) - 6f'_0 \Leftrightarrow \\ &\Leftrightarrow 16M_0 + 8M_1 = 2.25 \end{aligned}$$

- $i = 1$

$$(x_1 - x_0)M_0 + 2(x_2 - x_0)M_1 + (x_2 - x_1)M_2 = \frac{6}{x_2 - x_1}(f_2 - f_1) - \frac{6}{x_1 - x_0}(f_1 - f_0) \Leftrightarrow$$

$$\Leftrightarrow 8M_0 + 34M_1 + 9M_2 = -0.25$$

- $i = 2$

$$(x_2 - x_1)M_1 + 2(x_3 - x_1)M_2 + (x_3 - x_2)M_3 = \frac{6}{x_3 - x_2}(f_3 - f_2) - \frac{6}{x_2 - x_1}(f_2 - f_1) \Leftrightarrow$$

$$\Leftrightarrow 9M_1 + 32M_2 + 7M_3 = 4.8571$$

- $i = 3$

$$2(x_3 - x_2)M_3 + (x_3 - x_2)M_2 = 6f'_3 - \frac{6}{x_3 - x_2}(f_3 - f_2) \Leftrightarrow$$

$$\Leftrightarrow 7M_2 + 14M_3 = -3.8571$$

O sistema resultante deve resolver-se por EGPP. A sua solução é

$$\left(\begin{array}{cccc|c} 16 & 8 & 0 & 0 & 2.25 \\ 8 & 34 & 9 & 0 & -0.25 \\ 0 & 9 & 32 & 7 & 4.8571 \\ 0 & 0 & 7 & 14 & -3.8571 \end{array} \right) \rightarrow \begin{cases} M_0 = 0.2054 \\ M_1 = -0.1295 \\ M_2 = 0.2790 \\ M_3 = -0.4150 \end{cases}$$

O ponto $x = 29$ está no terceiro segmento (entre $x = 27$ e $x = 34$), logo,

$$s_3^3(x) = \frac{M_2}{6(x_3 - x_2)}(x_3 - x)^3 + \frac{M_3}{6(x_3 - x_2)}(x - x_2)^3 + \left(\frac{f_2}{x_3 - x_2} - \frac{M_2(x_3 - x_2)}{6} \right) (x_3 - x)$$

$$+ \left(\frac{f_3}{x_3 - x_2} - \frac{M_3(x_3 - x_2)}{6} \right) (x - x_2)$$

$$s_3^3(x) = 0.0066(34 - x)^3 - 0.0099(x - 27)^3 + 0.9602(34 - x) + 1.9128(x - 27)$$

$$f(29) \approx s_3^3(29) = 9.3779 \approx 9 \text{ golos.}$$

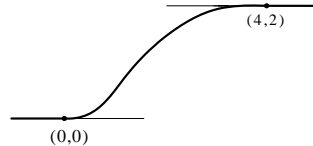
b) Como a função não é conhecida, tem de se construir a tabela das diferenças divididas para calcular uma aproximação ao majorante da quarta derivada.

x_i	f_i	$dd1$	$dd2$	$dd3$	$dd4$
10	20				
		-1			
12	18		0.0625		
		-0.5		-0.0043	
18	15		-0.0111		0.0006
		-0.6667		-0.0083	
27	9		0.1389		-0.0014
		1		-0.0221	
30	12		-0.2143		
		-0.5			
34	10				

$$|f(x) - s_3(x)| \leq \frac{5}{384} h^4 M_4 = \frac{5}{384} \times 9^4 \times 0.0014 \times 4! = 2.8335.$$

6.4 Exercícios

1. Pretende-se construir um desvio entre duas linhas de caminho de ferro paralelas. O desvio deve corresponder a um polinómio de grau três que une os pontos $(0,0)$ e $(4,2)$, como mostra a figura



Com base nos quatro pontos da tabela

x_i	-1	0	4	5
$f_i = f(x_i)$	0.4375	0	2	1.5625

construa uma 'spline' cúbica natural para definir a trajetória do desvio e calcular $f(2)$.

2. A resistência de um certo fio de metal, $f(x)$, varia com o diâmetro desse fio, x . Foram medidas as resistências de 6 fios de diversos diâmetros:

x_i	1.5	2.0	2.2	3.0	3.8	4.0
$f(x_i)$	4.9	3.3	3.0	2.0	1.75	1.5

Como se pretende estimar a resistência de um fio de diâmetro 1.75, use uma 'spline' cúbica natural para calcular esta aproximação.

3. Num certo campeonato regional de futebol há 7 equipas. No fim da temporada, o número de pontos ganhos e o número de golos sofridos por 6 das equipas estão representados na tabela

Equipa	F.C.Sol	F.C.Lá	S.C.Gato	Nova F.C.	Vila F.C.	F.C.Chão
Nº de pontos, x_i	10	12	18	27	30	34
Nº de golos, $f(x_i)$	20	18	15	9	12	10

- a) Use uma 'spline' cúbica completa para descrever a relação entre o número de pontos e o número de golos sofridos pelas equipas no campeonato. Sabendo que a 7ª equipa terminou o campeonato com 29 pontos, estime o número de golos que terá sofrido.
- b) Calcule uma estimativa do erro de truncatura cometido na alínea anterior.

4. Um braço de um robô deve passar nos instantes t_0, t_1, t_2, t_3, t_4 e t_5 por posições pré-definidas $\theta(t_0), \theta(t_1), \theta(t_2), \theta(t_3), \theta(t_4)$ e $\theta(t_5)$, onde $\theta(t)$ é o ângulo (em radianos) que o braço do robô faz com o eixo dos X's.

t_i	1	2	3	4	5	6
$\theta_i = \theta(t_i)$	1	1.25	1.75	2.25	3	3.15

- a) Com base nos dados da tabela, aproxime a trajetória do robô por uma 'spline' cúbica completa. Indique também uma aproximação da posição do robô no instante $t = 1.5$.
- b) Calcule uma aproximação à velocidade do robô no instante $t = 1.5$
- c) Calcule um limite superior do erro de truncatura que se comete quando se usa a derivada da 'spline' calculada para aproximar a velocidade do robô.
5. Considere a função $f(x)$ definida por

x	-2	0	1	2
$f(x)$	-8	0	1	8

Sabendo que $s_3^{1''}(-2) = 12$ e $s_3^{n''}(2) = 20$ estime o valor de $f(-1)$ através de uma 'spline' cúbica.

6. A seguinte função segmentada $s_3(x)$ no intervalo $[0, 3]$, poderá representar uma spline cúbica? Justifique.

$$s_3(x) = \begin{cases} s_3^1(x) = 3x^3 - x^2 + x - 2, & 0 \leq x \leq 1 \\ s_3^2(x) = 2x^3 + 2x - 3, & 1 \leq x \leq 3 \end{cases}$$

7. A tabela seguinte representa o crescimento de uma cultura de bactérias num líquido ao longo de 20 dias.

dia	0	4	12	18	20
nº de bactérias $\times 10^{-6}$	a	78.6	91.7	b	112.6

- a) Pretende-se determinar uma 'spline' cúbica completa para estimar o número de bactérias no líquido. Apresente o sistema das equações lineares que precisaria para calcular os valores dos M 's, em função das constantes a e b .
- b) Para $a = 67.4$ e $b = 107.0$, estime o número de bactérias presentes no líquido ao fim de 10 dias, usando uma 'spline' cúbica completa.

6.5 Soluções

1. $s_3^2(x) = 0.0390625(4-x)^3 - 0.0390625x^3 - 0.625(4-x) + 1.125x,$

$$f(2) \approx s_3^2(2) = 1.$$

2. $s_3^1(x) = 2.6830(x-1.5)^3 + 9.8(2-x) + 6.1292(x-1.5),$

$$f(1.75) \approx s_3^1(1.75) = 4.0242.$$

3. a) $s_3^3(x) = 0.0066(34-x)^3 - 0.0099(x-27)^3 + 0.9602(34-x) + 1.9128(x-27),$

$$f(29) \approx s_3^3(29) = 9.3779 \approx 9 \text{ golos.}$$

b) $|f(x) - s_3(x)| \leq 2.8335.$

4. a) $s_3^1(x) = 0.0134(3-x)^3 - 0.0044(x-1)^3 + 0.4464(3-x) + 0.8573(x-1),$

$$\theta(1.5) \approx s_3^1(1.5) = 1.1429.$$

b) $s_3^{1'}(x) = -0.0402(3-x)^2 - 0.0133(x-1)^2 + 0.4109,$

$$v(1.5) \approx s_3^{1'}(1.5) = 0.3171.$$

c) $|f'(x) - s_3'(x)| \leq 0.3672.$

5. $s_3^1(x) = (-x)^3 - 0.6667(x+2)^3 + 8x + 2.6667(x+2),$

$$f(-1) \approx s_3^1(-1) = -7.$$

6. A função dada não é 'spline' cúbica ($s_3^{1''}(1) \neq s_3^{2''}(1)$).

7. a)
$$\begin{cases} 24M_0 + 12M_1 & = a - 72.05 \\ 12M_0 + 40M_1 + 8M_2 & = 0.5a - 30.175 \\ 8M_1 + 16M_2 & = -3b + 322.125 \end{cases}$$

b) $s_3^1(x) = -0.0039(12-x)^3 + 0.0024x^3 + 6.1804(12-x) + 7.2892x,$

$$f(10) \approx s_3^1(10) = 87.6692.$$

6.6 MATLAB

6.6.1 spline

Para se obter uma spline cúbica em MATLAB, usa-se o comando `spline`. A sintaxe é

$$PP = \text{spline}(x,y)$$

em que PP vai conter uma estrutura que em `coefs` terá uma matriz de quatro colunas com os coeficientes que permitem escrever cada um dos segmentos da spline. Cada linha da matriz corresponde a um segmento, sendo que cada coluna representa os coeficientes do polinómio de grau três por ordem decrescente, na forma

$$c_1(x - x_0)^3 + c_2(x - x_0)^2 + c_3(x - x_0) + c_4,$$

sendo x_0 o nó inicial do segmento em questão.

Não existe nenhum comando em MATLAB para o cálculo de splines cúbicas naturais. No entanto, podem calcular-se splines cúbicas completas, indicando o declive (primeira derivada) nos extremos. A sintaxe, neste caso, passa a ser

$$PP = \text{spline}(x,[dd0 \text{ y } ddn])$$

em que `dd0` e `ddn` são as derivadas no ponto inicial e no ponto final, respetivamente.

Para se determinar a aproximação a um valor ou conjunto de valores, usa-se o comando `spline`, mas indicando como terceiro parâmetro de entrada o valor do ponto (ou vetor de pontos) interpolador:

$$YY = \text{spline}(x,y,XX)$$

ou

$$YY = \text{spline}(x,[dd0 \text{ y } ddn],XX)$$

Em alternativa pode usar-se o comando `ppval`:

$$YY = \text{ppval}(PP,XX)$$

Exemplo 6.4 *Usando a mesma tabela do Exemplo 5.3, pretende calcular-se uma spline cúbica sem fornecer derivadas e uma spline cúbica completa para determinar uma aproximação ao ponto $x = 5$.*

Para a spline cúbica, introduzem-se os pontos na forma de vetores x e y e executam-se os comandos

```
>> PP=spline(x,y);
```

```
>> PP.coefs
```

e o MATLAB devolve a matriz:

```
ans =
```

```
-0.8910    5.7308   -9.8397   10.0000
-0.8910    3.0577   -1.0513    5.0000
 0.5224   -2.2885    0.4872    8.0000
 0.5513    0.8462   -2.3974    4.0000
 0.5513    2.5000    0.9487    3.0000
```

Uma vez que o ponto interpolador $x = 5$ está no terceiro segmento, os coeficientes para este segmento estão na terceira linha desta matriz, e o polinómio pretendido é

$$s_3^3(x) = 0.5224(x - 4)^3 - 2.2885(x - 4)^2 + 0.4872(x - 4) + 8.0000.$$

Para se determinar a aproximação obtida em $x = 5$ através desta spline, usa-se o comando

```
>>YY=spline(x,y,5)
```

ou

```
>>YY=ppval(PP,5)
```

No caso da spline cúbica completa, a diferença é que terão de ser fornecidas as derivadas nos extremos, que neste caso se podem calcular aproximadamente por diferenças divididas. Sendo assim, o segundo e penúltimo pontos da tabela serão usados como pontos auxiliares no cálculo das derivadas e devem ser retirados da tabela. Os comandos para resolver este problema podem ser, por exemplo (depois de se introduzirem os vetores x e y como no caso anterior)

```
>>np=6 %número de pontos
```

```
>>dd0=(y(2)-y(1))/(x(2)-x(1))
```

```
>>ddn=(y(np)-y(np-1))/(x(np)-x(np-1))
```

```
>>x=x([1,3:np-2,np])
```

```
>>y=y([1,3:np-2,np])
```

```
>>PP = spline(x,[dd0 y ddn])
```

O polinómio correspondente ao segmento de interesse (que agora passa a ser o segundo) é escrito da mesma forma.

Para se determinar a aproximação por esta spline no ponto interpolador $x = 5$,

```
>>YY = spline(x,[dd0 y ddn],5)
```

ou

```
>>YY=ppval(PP,5)
```

6.6.2 Exercícios para resolver em MATLAB

1. Considerando a função
- $f(x)$
- dada pela tabela

x_i	5.0	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
f_i	0.0639	0.0800	0.0988	0.1203	0.1442	0.1714	0.2010	0.2331	0.2673	0.3036	0.3414

qual o valor aproximado da função no ponto $x = 5.45$

- a) usando uma 'spline' cúbica sem considerar derivadas nos extremos?

$$s_3(x) =$$

$$f(5.45) \approx s_3(5.45) =$$

- b) usando uma 'spline' cúbica completa?

$$s_3(x) =$$

$$f(5.45) \approx s_3(5.45) =$$

2. De uma tabela de logaritmos obteve-se o seguinte quadro de valores.

x_i	1	1.5	2	3	3.5
$\ln(x_i)$	0	0.4055	0.6931	1.0986	1.2528

- a) Usando uma função 'spline' cúbica sem usar derivadas nos extremos, calcule uma aproximação a
- $\ln(2.5)$
- .

$$s_3(x) =$$

$$f(2.5) \approx s_3(2.5) =$$

- b) Repita a alínea anterior, mas agora usando uma 'spline' cúbica completa.

$$s_3(x) =$$

$$f(2.5) \approx s_3(2.5) =$$

3. Foram registados os consumos de combustível $f(x_i)$, de um automóvel a arrancar em determinados instantes, x_i (em segundos).

x_i	0	0.1	0.2	0.3	0.4	0.5	0.6	3.6	6.6	9.6	9.8	10
f_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.6	0.6	0.6	0.7	0.8

- a) Usando uma função 'spline' cúbica sem usar derivadas nos extremos, calcule o consumo no instante de tempo $x_i = 5$ s.

$$s_3(x) =$$

$$f(5) \approx s_3(5) =$$

- b) Repita a alínea anterior, mas agora usando uma 'spline' cúbica completa.

$$s_3(x) =$$

$$f(5) \approx s_3(5) =$$

Capítulo 7

Integração numérica

7.1 Forma geral do problema

Pretende-se calcular uma aproximação ao integral definido

$$\int_a^b f(x) dx,$$

em que $f(x)$ é a função integranda definida em $[a, b]$ e os limites a e b são finitos.

Este tipo de aproximação é usado quando a primitiva de f não pode vir expressa em termos de funções elementares ou quando a função integranda, ainda que conhecida, é demasiado complicada. Aplica-se ainda quando a função integranda é conhecida apenas para um conjunto discreto de pontos.

Se $p_n(x)$ for uma aproximação polinomial a $f(x)$, então

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

é aproximado por

$$\int_a^b p_n(x) dx.$$

Se $e_n(x)$ é o erro da aproximação polinomial, $e_n(x) = f(x) - p_n(x)$, então

$$I = \int_a^b p_n(x) dx + \int_a^b e_n(x) dx,$$

ou seja,

$$I = \int_a^b p_n(x) dx + \{\text{erro de integração}\}.$$

7.2 Fórmulas simples de Newton-Cotes

O polinómio de Lagrange de grau menor ou igual a n é dado por

$$p_n(x) = \sum_{j=0}^n L_j(x) f(x_j)$$

com $f(x_j) = p_n(x_j)$, $j = 0, 1, \dots, n$, para $n + 1$ pontos em $[a, b]$. Se se usar este polinómio $p_n(x)$ para aproximar a função integranda $f(x)$,

$$\int_a^b f(x) dx \approx \int_a^b p_n(x) dx = \int_a^b \left(\sum_{j=0}^n L_j(x) f(x_j) \right) dx = \sum_{j=0}^n \underbrace{\left(\int_a^b L_j(x) dx \right)}_{\omega_j} f(x_j) = \sum_{j=0}^n \omega_j f(x_j) \quad (7.1)$$

Os coeficientes ω_j dependem da escolha dos pontos x_j , $j = 0, 1, \dots, n$ através dos polinómios $L_j(x)$ e são independentes dos valores de $f(x)$.

$$\omega_j = \int_a^b L_j(x) dx = \int_a^b \frac{(x - x_0) \cdots (x - x_{j-1})(x - x_{j+1}) \cdots (x - x_n)}{(x_j - x_0) \cdots (x_j - x_{j-1})(x_j - x_{j+1}) \cdots (x_j - x_n)} dx.$$

Para um conjunto de pontos igualmente espaçados em $[a, b]$,

$$x_j = a + jh, \quad j = 0, 1, \dots, n, \quad h = \frac{b - a}{n}.$$

É a partir dos polinómios de Lagrange que são deduzidas as fórmulas ou regras de Newton-Cotes.

7.2.1 Regra do retângulo

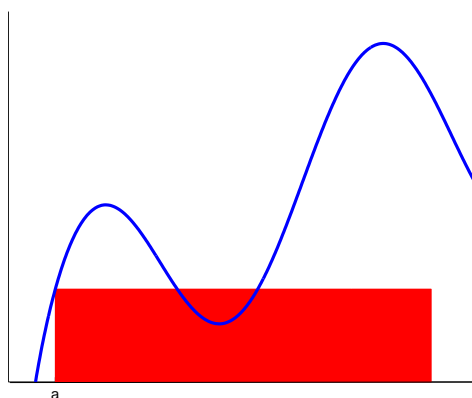


Figura 7.1: Representação gráfica da regra do retângulo.

Na regra do retângulo (Figura 7.1) é usado apenas um ponto dos extremos do intervalo $[a, b]$, logo, o polinómio para aproximar a função integranda é de grau zero.

Por exemplo, usando o ponto a ,

$$n = 0, x_0 = a, L_0 = 1, \omega_0 = \int_a^b L_0(x)dx = b - a.$$

De (7.1) vem

$$\int_a^b f(x)dx \approx (b - a)f(a).$$

7.2.2 Regra do ponto médio

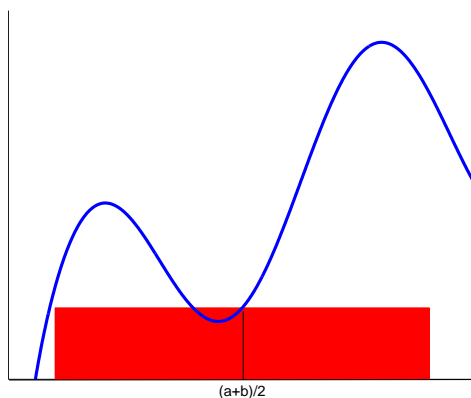


Figura 7.2: Representação gráfica da regra do ponto médio.

Na regra do ponto médio (Figura 7.2) é usado apenas o ponto médio do intervalo $[a, b]$, logo, o polinómio para aproximar a função integranda é também, neste caso, de grau zero.

$$n = 0, x_0 = \frac{a+b}{2}, L_0(x) = 1, \omega_0 = \int_a^b L_0(x)dx = b - a.$$

De (7.1) vem

$$\int_a^b f(x)dx \approx (b - a)f\left(\frac{a+b}{2}\right).$$

7.2.3 Regra do trapézio

Na regra do trapézio (Figura 7.3) são usados os dois pontos extremos do intervalo $[a, b]$, logo, o polinómio para aproximar a função integranda é de grau um.

$$n = 1, x_0 = a, x_1 = b, L_0(x) = \frac{x-b}{a-b}, \omega_0 = \int_a^b L_0(x)dx = \frac{b-a}{2},$$

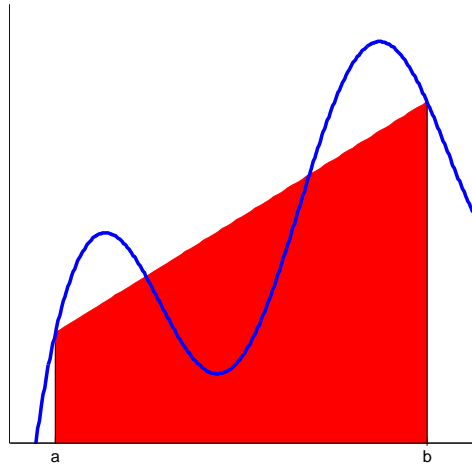


Figura 7.3: Representação gráfica da regra do trapézio.

$$L_1(x) = \frac{x-a}{b-a}, \quad \omega_1 = \int_a^b L_1(x) dx = \frac{b-a}{2}.$$

De (7.1) vem

$$\int_a^b f(x) dx \approx \frac{b-a}{2} [f(a) + f(b)].$$

7.2.4 Regra de Simpson

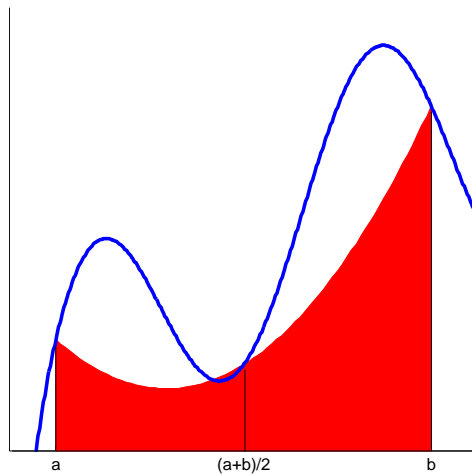


Figura 7.4: Representação gráfica da regra de Simpson.

Na regra de Simpson (Figura 7.4) são usados três pontos do intervalo $[a, b]$ - os dois pontos extremos e o ponto médio -, logo, o polinômio para aproximar a função integranda é de grau dois.

$$n = 2, x_0 = a, x_1 = \frac{a+b}{2}, x_2 = b, \omega_0 = \omega_2 = \int_a^b \frac{(x - \frac{a+b}{2})(x-b)}{(a - \frac{a+b}{2})(a-b)} dx = \frac{b-a}{6},$$

$$\omega_1 = \int_a^b \frac{(x-a)(x-b)}{(\frac{a+b}{2} - a)(\frac{a+b}{2} - b)} dx = \frac{4(b-a)}{6}.$$

De (7.1) vem

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right].$$

7.2.5 Regra dos três oitavos

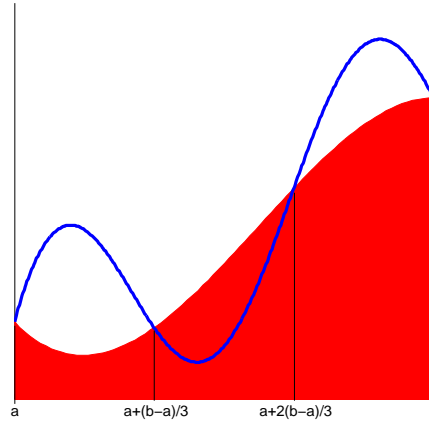


Figura 7.5: Representação gráfica da regra dos três oitavos.

Na regra dos três oitavos (Figura 7.5) são usados os quatro pontos do intervalo $[a, b]$, logo, o polinômio para aproximar a função integranda é de grau três.

$$n = 3, x_0 = a, x_1 = a + \frac{b-a}{3}, x_2 = a + 2\frac{b-a}{3}, x_3 = b, \omega_0 = \omega_3 = \frac{b-a}{8}, \omega_1 = \omega_2 = \frac{3(b-a)}{8}.$$

De (7.1) vem

$$\int_a^b f(x) dx \approx \frac{b-a}{8} \left[f(a) + 3f\left(\frac{2a+b}{3}\right) + 3f\left(\frac{a+2b}{3}\right) + f(b) \right].$$

7.2.6 Erros de truncatura

Como já se pôde perceber anteriormente, as fórmulas de integração numérica, em concreto, as fórmulas de Newton-Cotes, têm um erro de truncatura associado, resultante da discretização do integral.

Se a derivada de ordem $n + 1$ de $f(x)$ for contínua em $[a, b]$ e se os x_j , $j = 0, 1, \dots, n$ pertencem ao intervalo $[a, b]$, então existe um ponto $\eta \in [a, b]$ tal que o erro de aproximar $f(x)$ pelo polinômio $p_n(x)$ é

$$e_n(x) = \prod_{j=0}^n (x - x_j) \frac{f^{(n+1)}(\eta)}{(n+1)!}, \quad \eta \in [a, b]$$

e o erro de integração é dado por

$$E(x) = \frac{1}{(n+1)!} \int_a^b \prod_{j=0}^n (x - x_j) f^{(n+1)}(\eta) dx, \quad \eta \in [a, b]. \quad (7.2)$$

Esta fórmula pode ser simplificada, usando o teorema do valor médio (Teorema 7.2.1).

Teorema 7.2.1 *Teorema do valor médio*

Se duas funções $f(x)$ e $g(x)$ são contínuas e se, além disso, g não muda de sinal no intervalo $[a, b]$, então existe um ponto $\xi \in [a, b]$ tal que

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

Aplicando o Teorema 7.2.1 a (7.2) vem

$$E(x) = \frac{f^{(n+1)}(\eta)}{(n+1)!} \int_a^b \prod_{j=0}^n (x - x_j) dx, \quad \eta \in [a, b].$$

No entanto, nem sempre é possível aplicar o teorema do valor médio. Nas regras do ponto médio e de Simpson, $g(x)$ muda de sinal em $[a, b]$, no ponto $\frac{a+b}{2}$. Neste caso, os erros são deduzidos a partir da forma geral do resto (7.2). Assim, tem-se

- Erro de truncatura da regra do retângulo

$$e_R = \frac{(b-a)^2}{2} f'(\eta), \quad \eta \in [a, b];$$

- Erro de truncatura da regra do ponto médio

$$e_M = \frac{(b-a)^3}{24} f''(\eta), \quad \eta \in [a, b];$$

- Erro de truncatura da regra do trapézio

$$e_T = -\frac{(b-a)^3}{12} f''(\eta), \quad \eta \in [a, b];$$

- Erro de truncatura da regra de Simpson

$$e_S = -\frac{(b-a)^5}{32} \frac{1}{90} f^{(iv)}(\eta), \eta \in [a, b];$$

- Erro de truncatura da regra dos três oitavos

$$e_{3/8} = -\frac{(b-a)^5}{6480} f^{(iv)}(\eta), \eta \in [a, b].$$

O comportamento do erro depende da derivada de ordem $n+1$ da função, dependendo também do valor selecionado para n . Se o resultado obtido por uma fórmula simples de Newton-Cotes não é satisfatório, isto é, se o erro de truncatura for muito grande, é possível aumentar o valor de n , o que equivale a aumentar o número de pontos e o grau do polinômio que aproxima $f(x)$. No entanto, nem sempre é verdade que quanto maior for n maior seja a precisão do resultado numérico, uma vez que para polinômios de grau elevado os erros da aproximação podem ser também grandes. A forma mais evidente de reduzir o erro de integração seria diminuir o valor de $(b-a)$, mas este intervalo é fixo! A solução será dividir o intervalo fixo $[a, b]$ em subintervalos, que podem ser tão pequenos quanto o necessário, e em cada um desses subintervalos, $f(x)$ é aproximada por um polinômio de grau baixo. Surgem assim as **fórmulas compostas**.

7.3 Fórmulas compostas

Tome-se como exemplo o intervalo $[a, b]$ subdividido em 6 subintervalos.

Para $a < x_1 < x_2 < \dots < x_5 < b$:

$$\int_a^b f(x)dx = \int_a^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \int_{x_2}^{x_3} f(x)dx + \int_{x_3}^{x_4} f(x)dx + \int_{x_4}^{x_5} f(x)dx + \int_{x_5}^b f(x)dx$$

7.3.1 Fórmula composta do trapézio

A fórmula composta do trapézio pode deduzir-se da correspondente regra simples dividindo o intervalo $[a, b]$ em n subintervalos iguais de comprimento $h = \frac{b-a}{n}$ (Figura 7.6),

$$\int_a^b f(x)dx = \sum_{j=0}^{n-1} \left\{ \int_{x_j}^{x_{j+1}} f(x)dx \right\}$$

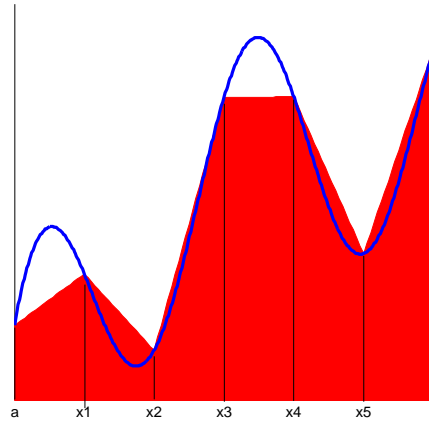


Figura 7.6: Representação gráfica da fórmula composta do trapézio para $n = 6$ subintervalos.

com $x_j = a + jh$ e $j = 0, 1, \dots, n$. Usando a regra do trapézio em cada subintervalo $[x_j, x_{j+1}]$:

$$\int_{x_j}^{x_{j+1}} f(x)dx \approx \frac{x_{j+1} - x_j}{2} [f(x_j) + f(x_{j+1})] = \frac{1}{2}h [f_j + f_{j+1}],$$

com erro

$$e_{T_j} = -\frac{h^3}{12}f''(\eta_j), \text{ com } \eta_j \in [x_j, x_{j+1}].$$

Somando as contribuições dos n subintervalos:

$$\begin{aligned} \int_a^b f(x)dx &= \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \dots + \int_{x_{n-2}}^{x_{n-1}} f(x)dx + \int_{x_{n-1}}^{x_n} f(x)dx \\ &\approx \sum_{j=0}^{n-1} \left\{ \frac{1}{2}h [f_j + f_{j+1}] \right\} = \frac{1}{2}h [f_0 + f_1 + f_1 + f_2 \\ &+ f_2 + f_3 + \dots + f_{n-2} + f_{n-1} + f_{n-1} + f_n]. \end{aligned}$$

Assim,

$$T(h) = \frac{h}{2} [f_0 + 2f_1 + 2f_2 + \dots + 2f_{n-1} + f_n].$$

Somando também os erros:

$$\begin{aligned}
 e_{CT} &= \sum_{j=0}^{n-1} \left\{ -\frac{h^3}{12} f''(\eta_j) \right\} \\
 &= -\frac{h^3}{12} \{f''(\eta_0) + f''(\eta_1) + \cdots + f''(\eta_{n-1})\} \\
 &= -\frac{h^2}{12} \frac{(b-a)}{n} n f''(\eta) \\
 &= -\frac{h^2}{12} (b-a) f''(\eta) \\
 e_{CT} &= -\frac{h^2}{12} (b-a) f''(\eta), \quad \eta \in [a, b].
 \end{aligned}$$

7.3.2 Fórmula composta de Simpson

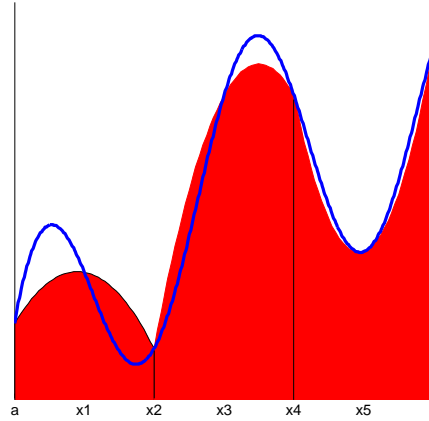


Figura 7.7: Representação gráfica da fórmula composta de Simpson para $n = 6$ subintervalos com $m = 3$ contribuições.

A fórmula composta de Simpson pode deduzir-se da correspondente regra simples. Cada contribuição necessita de dois subintervalos. Assim, $n = 2m$. O espaçamento entre pontos é $h = \frac{b-a}{n} = \frac{b-a}{2m}$ e existem m contribuições. Em cada par de subintervalos usa-se a regra de Simpson (Figura 7.7):

$$\int_{x_{j-1}}^{x_{j+1}} f(x) dx \approx \frac{x_{j+1} - x_{j-1}}{6} [f_{j-1} + 4f_j + f_{j+1}] = \frac{2h}{6} [f_{j-1} + 4f_j + f_{j+1}]$$

com erro

$$e_{S_j} = -\frac{(2h)^5}{32 \times 90} f^{(iv)}(\eta_j), \text{ com } \eta_j \in [x_{j-1}, x_{j+1}].$$

Somando as m contribuições:

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{m \text{ termos}} \int_{x_{j-1}}^{x_{j+1}} f(x) dx \approx \sum_{j=1(2 \text{ em } 2)}^{n-1} \left\{ \frac{h}{3} [f_{j-1} + 4f_j + f_{j+1}] \right\} \\ &= \frac{h}{3} [f_0 + 4f_1 + f_2 + f_2 + 4f_3 + f_4 + f_4 + 4f_5 + f_6 + \cdots + f_{n-2} + 4f_{n-1} + f_n]. \end{aligned}$$

Assim,

$$S(h) = \frac{h}{3} [f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \cdots + 2f_{n-2} + 4f_{n-1} + f_n].$$

Somando também os erros:

$$\begin{aligned} e_{CS} &= \sum_{m \text{ termos}} \left\{ -\frac{2^5 h^5}{32 \times 90} f^{(iv)}(\eta_j) \right\} \\ &= -\frac{h^5}{90} \{ f^{(iv)}(\eta_1) + f^{(iv)}(\eta_3) + f^{(iv)}(\eta_5) + \cdots + f^{(iv)}(\eta_{m-1}) \} \\ &= -\frac{h^4}{90} \frac{(b-a)}{2m} m f^{(iv)}(\eta) \\ &= -\frac{h^4}{180} (b-a) f^{(iv)}(\eta) \\ e_{CS} &= -\frac{h^4}{180} (b-a) f^{(iv)}(\eta), \quad \eta \in [a, b]. \end{aligned}$$

7.3.3 Fórmula composta dos três oitavos

A fórmula composta dos três oitavos pode ser deduzida através da correspondente regra simples, mas, neste caso, cada contribuição necessita de três subintervalos. Assim, $n = 3r$. O espaçamento entre pontos é $h = \frac{b-a}{n} = \frac{b-a}{3r}$ e existem r contribuições. Em cada conjunto de 3 subintervalos usa-se a regra dos três oitavos (Figura 7.8):

$$\int_{x_{j-1}}^{x_{j+2}} f(x) dx \approx \frac{x_{j+2} - x_{j-1}}{8} [f_{j-1} + 3f_j + 3f_{j+1} + f_{j+2}] = \frac{3h}{8} [f_{j-1} + 3f_j + 3f_{j+1} + f_{j+2}],$$

com erro

$$e_{3/8_j} = -\frac{(3h)^5}{6480} f^{(iv)}(\eta_j), \text{ com } \eta_j \in [x_{j-1}, x_{j+2}].$$

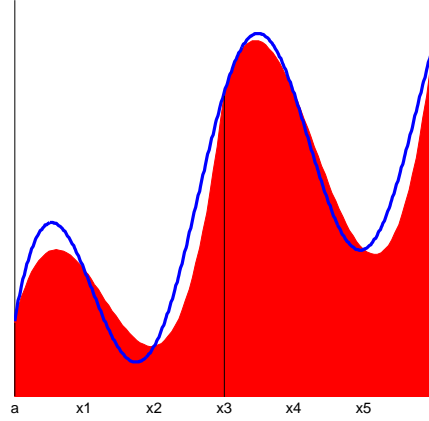


Figura 7.8: Representação gráfica da fórmula composta dos três oitavos para $n = 6$ subintervalos com $r = 2$ contribuições.

Somando todas as r contribuições:

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{r \text{ termos}} \int_{x_{j-1}}^{x_{j+2}} f(x)dx \approx \sum_{j=1(3 \text{ em } 3)}^{n-2} \left\{ \frac{3h}{8} [f_{j-1} + 3f_j + 3f_{j+1} + f_{j+2}] \right\} \\ &= \frac{3h}{8} [f_0 + 3f_1 + 3f_2 + f_3 + f_3 + 3f_4 + 3f_5 + f_6 + \cdots + f_{n-3} + 3f_{n-2} + 3f_{n-1} + f_n]. \end{aligned}$$

Assim,

$$3/8(h) = \frac{3h}{8} [f_0 + 3f_1 + 3f_2 + 2f_3 + 3f_4 + \cdots + 2f_{n-3} + 3f_{n-2} + 3f_{n-1} + f_n].$$

Somando os erros:

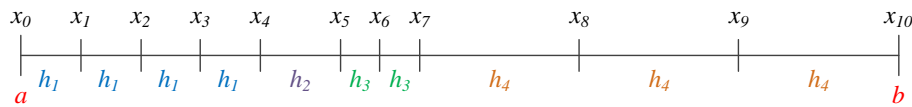
$$\begin{aligned} e_{C3/8} &= \sum_{r \text{ termos}} \left\{ -\frac{(3h)^5}{6480} f^{(iv)}(\eta_j) \right\} \\ &= -\frac{(3h)^5}{6480} \{ f^{(iv)}(\eta_1) + f^{(iv)}(\eta_4) + f^{(iv)}(\eta_7) + \cdots + f^{(iv)}(\eta_{n-1}) \} \\ &= -\frac{81 \times 3 h^4}{6480} \frac{(b-a)}{3r} r f^{(iv)}(\eta) \\ &= -\frac{h^4}{80} (b-a) f^{(iv)}(\eta) \end{aligned}$$

$$e_{C3/8} = -\frac{h^4}{80} (b-a) f^{(iv)}(\eta), \quad \eta \in [a, b].$$

7.4 Aplicação das fórmulas de integração a intervalos de amplitudes diferentes

Como se verificou pelo descrito anteriormente, as fórmulas compostas de integração só podem aplicar-se a intervalos de amplitude constante. Surge então a questão do se deve fazer no caso de haver intervalos cujo espaçamento entre pontos não é constante. A resposta consiste em agrupar subintervalos com amplitudes iguais e aplicar a cada grupo uma fórmula de integração. Atente-se no Exemplo 7.1.

Exemplo 7.1 Considere-se o seguinte conjunto de 11 pontos no intervalo $[a, b]$.



$$\int_a^b f(x)dx = \int_{x_0}^{x_{10}} f(x)dx = \int_{x_0}^{x_4} f(x)dx + \int_{x_4}^{x_5} f(x)dx + \int_{x_5}^{x_7} f(x)dx + \int_{x_7}^{x_{10}} f(x)dx$$

por exemplo

$$\approx \underbrace{S(h_1)}_{n=4} + \underbrace{T(h_2)}_{n=1} + \underbrace{S(h_3)}_{n=2} + \underbrace{3/8(h_4)}_{n=3}$$

7.5 Escolha da melhor fórmula

O integral deve ser sempre calculado de forma a que se garanta que o erro de truncatura cometido é o menor possível. Tendo presente que cada caso tem de ser analisado de acordo com as suas particularidades, podem descrever-se algumas regras gerais, que a seguir se apresentam.

- condição de aplicabilidade

se n (número de subintervalos)

- é par e múltiplo de 3 (ex. 6, 12, ...) pode usar-se Simpson, 3 oitavos e trapézio;
- é par e não é múltiplo de 3 (ex. 4, 8, 10, ...) pode usar-se Simpson e trapézio;
- é múltiplo de 3 e não é par (ex. 9, 15, ...) pode usar-se 3 oitavos e trapézio;
- não é par nem múltiplo de 3 (ex. 5, 11, ...) pode usar-se trapézio.

- entre Simpson e 3 oitavos, Simpson tem sempre um erro menor;
- entre trapézio e Simpson, ou trapézio e 3 oitavos, deve analisar-se:
 - h^2 (trapézio) vs h^4 (Simpson e 3 oitavos);
 - coeficientes em valor absoluto

$$\frac{1}{12} \text{ (trapézio); } \frac{1}{180} \text{ (Simpson); } \frac{1}{80} \text{ (3 oitavos)}$$

- M_2 vs M_4 com

$$\left| f''_{[a,b]} \right| \leq M_2 \text{ (majorante) (trapézio)}$$

$$\left| f^{(iv)}_{[a,b]} \right| \leq M_4 \text{ (majorante) (Simpson e 3 oitavos).}$$

Exemplo 7.2 Na tabela seguinte são apresentados registos pontuais das vendas de um produto que foi lançado no início do ano de 2009. A variável x representa a semana (de 2009).

x_i	1	2	3	4	5	7	9	11	13	15	16	17	18	19
$v(x_i)$	10	9	8	8	8	6	5	5	4	4	4	4	3	1

- a) Calcule a melhor aproximação ao integral $\int_1^{19} v(x)dx$, com base em toda a informação fornecida na tabela sobre $v(x)$.
- b) Estime o erro de truncatura cometido com a aproximação obtida na alínea anterior no intervalo $[5, 15]$.

Resolução:

a)

$$\int_1^{19} v(x)dx = \underbrace{\int_1^5 v(x)dx}_{h=1, n=4, \mathbf{S}} + \underbrace{\int_5^{15} v(x)dx}_{h=2, n=5, \mathbf{T}} + \underbrace{\int_{15}^{19} v(x)dx}_{h=1, n=4, \mathbf{S}}$$

$$\approx \frac{1}{3}(10 + 4 \times 9 + 2 \times 8 + 4 \times 8 + 8) + \frac{2}{2}(8 + 2 \times 6 + 2 \times 5 + 2 \times 5 + 2 \times 4 + 4) + \frac{1}{3}(4 + 4 \times 4 + 2 \times 4 + 4 \times 3 + 1)$$

$$= 32 + 52 + 13.666667 = 99.666667$$

b) Trapézio - [5,15]

x_i	v_i	$dd1$	$dd2$
5	8		
		-1	
7	6		0.125
		-0.5	
9	5		0.125
		0	
11	5		-0.125
		-0.5	
13	4		0.125
		0	
15	4		

$$|e_T| = \frac{2^2}{12}(15 - 5) \times 0.125 \times 2! = 0.833333.$$

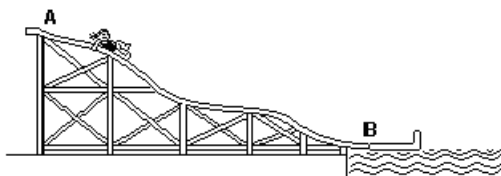
7.6 Exercícios

1. Considere o erro de truncatura da fórmula do retângulo, baseada em a , de Newton-Cotes

$$e_R = \frac{(b-a)^2}{2} f'(\eta), \quad \eta \in [a, b]$$

para aproximar o integral $\int_a^b f(x)dx$. Deduza a fórmula do erro de truncatura da correspondente fórmula composta.

2. A figura mostra uma pessoa que desliza, sem atrito, do alto de um escorrega (ponto A), acoplando-se a um carrinho que se encontra em repouso no ponto B. A partir deste instante, a pessoa e o carrinho movem-se juntos na água até parar.



- a) Sabendo que a velocidade do conjunto pessoa-carrinho imediatamente após o acoplamento é 4 m/s e que a velocidade, v , em cada instante t na água é dada pela tabela seguinte, calcule (usando todos os pontos da tabela) a distância percorrida na água pelo conjunto pessoa-carrinho até parar.

t	0.0	0.3	0.6	0.8	1.0	1.2	1.8	2.4	3.0	3.6	4.2
v	4.0	3.9	3.7	3.5	3.3	2.9	2.5	2.0	1.25	0.75	0.0

- b) Estime o erro de truncatura cometido na alínea anterior.
- c) Seleccione o maior número possível de pontos da tabela por forma a obter um conjunto de pontos igualmente espaçados, e calcule a mesma distância usando uma única fórmula composta de integração no intervalo $[0, 4.2]$.
3. A resposta de um transdutor a uma onda de choque causada por uma explosão é dada pela função $F(t) = 8e^{-t} \frac{I(a)}{\pi}$ para $t \geq a$, em que

$$I(a) = \int_1^2 f(x, a) dx \quad \text{com } f(x, a) = \frac{e^{ax}}{x}.$$

Calcule $I(1)$ usando a fórmula composta do trapézio com erro de truncatura inferior a 0.05.

4. Uma corrida de *dragsters* tem duas fases distintas: na primeira fase, a mais curta, o movimento do carro é perfeitamente não determinístico, dependendo das derrapagens e da forma como o condutor consegue dominar o carro. Na segunda fase, o carro tem um movimento muito rápido, cuja aceleração está perfeitamente definida.



Considere-se a prova do condutor Don Nase de duração 7.5 s. Na primeira fase os valores da aceleração em cada instante encontram-se na tabela:

t_i	0	0.5	1	1.5
$a(t_i)$	0	0.35	0.55	0.9

Na segunda fase da corrida a aceleração é definida pela seguinte expressão:

$$a(t) = 0.5t^2 - 0.15t \quad \text{para } t \in [1.5, 7.5].$$

- Estime a velocidade na primeira fase da corrida, utilizando a fórmula de integração mais adequada.
 - Estime a velocidade na segunda fase da corrida, utilizando a fórmula composta do trapézio com erro de truncatura em valor absoluto inferior a 0.3.
 - Estime o erro de truncatura cometido na alínea a).
5. Considere a seguinte função dada pela tabela

x_i	1	1.15	1.3	1.45	1.6	1.75	1.9
$f(x_i)$	a	16.8	19.4	22	b	27.6	30.7

e seja $I = \int_1^{1.9} f(x) dx$. Ao utilizar as fórmulas compostas de Simpson e dos três oitavos foram obtidas as seguintes aproximações a I , respectivamente $S(0.15) = 20.005$ e $3/8(0.15) = 20.030625$. Determine os valores de a e b . Use 6 casas decimais nos cálculos.

6. Na tabela seguinte são apresentados registos pontuais das vendas de um produto que foi lançado no início do ano de 2009. A variável x representa a semana (de 2009).

x_i	1	2	3	4	5	7	9	11	13	15	16	17	18	19
$v(x_i)$	10	9	8	8	8	6	5	5	4	4	4	4	3	1

- a) Calcule a melhor aproximação ao integral $\int_1^{19} v(x)dx$, com base em toda a informação fornecida na tabela sobre $v(x)$.
- b) Estime o erro de truncatura cometido com a aproximação obtida na alínea anterior no intervalo $[5, 15]$.
- c) Seleccione o maior número possível de pontos da tabela para calcular uma aproximação ao integral da alínea a), usando só uma fórmula composta de integração no intervalo $[1, 19]$.
7. Considere a seguinte tabela da função $f(x)$

x_i	0.0	1.0	2.0
$f(x_i)$	0.0000	0.8415	0.9093

- a) Determine um valor aproximado de $I = \int_0^2 f(x)dx$, usando a fórmula composta do trapézio com $h = 1$.
- b) Sabendo que um valor aproximado de I , usando a fórmula composta do trapézio com $h = 0.5$ é $T(0.5) = 1.2667$, determine uma nova aproximação de I , usando a fórmula composta de Simpson com $h = 0.5$.
8. Determine uma aproximação ao valor do integral definido

$$\int_0^1 \left(x^2 + \frac{1}{x+1} \right) dx$$

através da fórmula de Simpson, com um erro de truncatura inferior a 0.0005 em valor absoluto.

9. Admita que, para acções de uma determinada empresa cotada na bolsa de Nova Iorque, o lucro anual por acção, depois de impostos, é representado por x (US \$), uma variável aleatória que tem a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{4}{27}(9x - 6x^2 + x^3), & \text{para } 0 \leq x \leq 3 \\ 0, & \text{para outros valores de } x. \end{cases}$$

- a) Calcule, numericamente, a probabilidade $PROB$ do lucro anual ser um valor menor do que 1 ou maior do que 2.5 ($PROB = P(x \leq 1) + P(x \geq 2.5)$).

Use a fórmula composta do trapézio para calcular essa probabilidade por forma a que o erro total de truncatura seja inferior a 0.02. Assuma que os erros das duas parcelas são iguais.

Nota: $P(a \leq x \leq b) = \int_a^b f(x) dx$.

- b) Relativamente à primeira parcela para o cálculo de $PROB$, se tivesse usado a fórmula composta de Simpson com o mesmo valor de h que usou na alínea anterior, iria obter um erro menor, ou seja uma melhor aproximação ao valor de $P(x \leq 1)$? Justifique a resposta.

7.7 Soluções

1. $e_{RC} = \frac{h}{2}(b-a)f'(\eta)$, $\eta \in [a, b]$.
2. a) $I \approx 9.125$.
b) $|e| \leq 0.0648$.
c) $I \approx 9.06$.
3. $I \approx 3.0687$ ($h = 0.25$).
4. a) $I \approx 0.675$.
b) $I \approx 65.98125$ ($h = 0.75$).
c) $|e_{3/8}| \leq 0.0013$.
5. $a = 15$ e $b = 25$.
6. a) $I \approx 99.6667$.
b) $|e_T| \leq 0.8333$.
c) $I \approx 99.75$.
7. a) $I \approx 1.29615$.
b) $I \approx 1.256883$.
8. $I \approx 1.0265$ ($h = 0.125$).
9. (a) $I \approx 0.4237$ ($h_1 = 0.25$, $h_2 = 0.5$).
(b) Sim.

7.8 MATLAB

7.8.1 trapz

A função `trapz` do MATLAB implementa a fórmula do trapézio para resolver um integral numericamente, com base num conjunto de pontos (x, y) . A sintaxe é

```
>>I=trapz(x,y)
```

I é a aproximação ao integral e x e y são vetores que contêm os dados e têm de ter a mesma dimensão.

Exemplo 7.3 *Com base na tabela do Exemplo 5.3, calcular uma aproximação de $I = \int_1^8 f(x)dx$. Depois de se introduzirem os vetores x e y no MATLAB, o comando para resolver o problema é*

```
>>I=trapz(x,y)
```

7.8.2 integral

A função `integral` do MATLAB implementa quadratura global adaptável. É necessário fornecer a função a integrar (com o comando `inline` ou criando uma `m-file` do tipo função). Para se usar a função `quad` é necessário conhecer-se a expressão analítica da função integranda. Caso seja apenas conhecida uma tabela de pontos, só pode ser usada a função `trapz`. A sintaxe é

```
I=integral('func',a,b,'AbsTol',val1,'RelTol',val2).
```

É devolvida a aproximação ao integral em I . `func` é o nome da função a integrar, a e b são os limites de integração e opcionalmente podem fornecer-se valores para a tolerância absoluta `AbsTol` e/ou para a tolerância relativa `RelTol`. Os valores por defeito são 10^{-10} e 10^{-6} , respetivamente. Quanto maiores os valores de `AbsTol` e `RelTol`, menor é o tempo de computação. No entanto os resultados serão menos exatos.

Exemplo 7.4 *Calcular uma aproximação a $I = \int_0^1 e^{-x^2}(\ln(x))^2$*

Os comandos a introduzir são

```
>>f = @(x) exp(-x.^2).*log(x).^2
```

```
>>I = integral(f,0,1)
```

Em alternativa, a função f pode ser criada numa `m-file`.

7.8.3 Exercícios para resolver em MATLAB

1. Dada a tabela de valores da função $f(x)$

x_i	0.0	0.5	1.0	1.5	2.0	2.5	3.0	4.0	5.0
f_i	-4271	-2522	-499	1795	4358	7187	10279	13633	17247

Calcule a melhor aproximação ao integral

$$\int_{0.0}^{5.0} f(x) dx$$

usando toda a informação da tabela.

2. Calcule uma aproximação ao integral

$$I = \int_0^1 \frac{4}{1+x^2} dx.$$

3. Considere a seguinte tabela de valores da função $f(x)$:

x_i	-1.0	-0.95	-0.8	-0.6	-0.4	-0.2	0.2	0.6	1.0	1.4
f_i	-1.000	-0.05	1.00	0.980	0.95	0.900	0.850	0.200	-0.500	-2.000

- a) Calcule numericamente $\int_{-1.0}^{1.4} f(x) dx$ usando todos os pontos da tabela.
- b) Calcule o mesmo integral usando um valor de h constante, deixando de fora o menor número possível de pontos.
4. A função $F(t)$ surge na determinação da tensão à superfície de um líquido que rodeia uma bolha esférica de gás:

$$F(t) = \int_0^t \frac{P(x)}{Q(x)} dx, \quad \text{para } 0 \leq t \leq 1$$

em que

$$P(x) = 3 + 3x + x^2$$

$$Q(x) = 3 + 6x + 6x^2 + 2x^3.$$

Determine $F(1)$, usando uma tolerância absoluta de 10^{-15} e uma tolerância relativa de 10^{-10} .

Capítulo 8

Aproximação dos mínimos quadrados

8.1 Forma geral do problema

Pretende-se, com esta aproximação, definir um modelo, $M(x; c_i)$, dado por uma expressão matemática, que se ajuste o melhor possível à função dada $f(x)$, definida no intervalo $[a, b]$. Usando a técnica dos mínimos quadrados, pretende-se minimizar a soma dos quadrados dos erros (Figura 8.1). Podem ter-se dois tipos de problema distintos:

- O problema discreto, em que são dados m pontos $x_1 < x_2 < \dots < x_m$ no intervalo $[a, b]$ (Figura 8.1(a)) e o objetivo é

$$\text{minimizar } \sum_{j=1}^m (f(x_j) - M(x_j; c_i))^2.$$

- O problema contínuo, em que é dada uma função $f(x)$ (Figura 8.1(b)) e o objetivo é

$$\text{minimizar } \int_a^b (f(x) - M(x; c_i))^2 dx.$$

Neste capítulo vai abordar-se apenas o problema discreto, ou seja, a função f é dada por um conjunto discreto de valores $(x_1, f_1), (x_2, f_2), \dots, (x_m, f_m)$.

Os modelos de mínimos quadrados, dependendo da sua forma, podem classificar-se em

- Modelo linear e polinomial, que define um problema de mínimos quadrados linear e a sua expressão é dada por um polinómio. Por exemplo,

$$M(x; a_0, a_1, a_2) \equiv p_2(x) = a_0 + a_1x + a_2x^2.$$

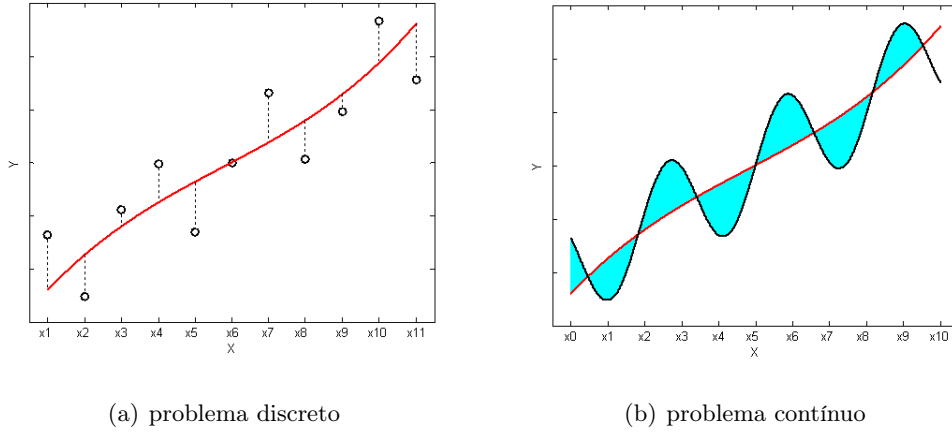


Figura 8.1: Problema dos mínimos quadrados

- Modelo linear e não polinomial, que define um problema de mínimos quadrados linear, mas não é um polinômio. Por exemplo,

$$M(x; c_1, c_2) = c_1 e^{x^2} + c_2 \sin(x).$$

- Modelo não linear, que define um problema de mínimos quadrados não linear. Por exemplo,

$$M(x; c_1, c_2) = \ln(c_1 x^2) + e^{c_2 x}.$$

O modelo não linear não vai ser abordado neste texto.

8.2 Modelo polinomial - polinômios ortogonais

Pretende-se construir um modelo definido por um polinômio completo de grau n , $p_n(x)$. Para que o problema seja bem definido, deve verificar-se a condição única $m \geq n + 1$, sendo m o número de pontos onde a função é conhecida e n o grau do polinômio. No caso particular em que $m = n + 1$ está-se perante o polinômio interpolador, já que o polinômio que passa por um dado conjunto de pontos é único e por isso $\sum_{j=1}^m (f_j - p_1(x_j))^2 = 0$.

Para que o problema seja bem condicionado, isto é, não seja sensível a erros nos dados ou erros de arredondamento nos cálculos, o polinômio $p_n(x)$ deve ser construído usando uma sequência de polinômios ortogonais, $P_0(x), P_1(x), \dots, P_n(x)$, na forma

$$p_n(x) = c_0 P_0(x) + c_1 P_1(x) + c_2 P_2(x) + \dots + c_n P_n(x). \quad (8.1)$$

8.2.1 Polinômios ortogonais

A propriedade que define os polinômios ortogonais é

$$\sum_{i=1}^m P_j(x_i) P_k(x_i) \begin{cases} = 0, & \text{se } j \neq k \\ \neq 0, & \text{se } j = k. \end{cases}$$

A partir dos pontos dados (x_j, f_j) , $j = 1, 2, \dots, m$ determina-se a sequência de polinômios ortogonais $P_0(x), \dots, P_n(x)$ e os coeficientes c_0, \dots, c_n para construir o polinômio completo $p_n(x)$ através de (8.1).

passo 1 Constroem-se os polinômios ortogonais, $P_0(x), P_1(x), P_2(x), \dots, P_n(x)$, da sequência de polinômios ortogonais, usando a relação de recorrência (8.2).

$$P_{i+1}(x) = (x - B_i) P_i(x) - \mathbb{C}_i P_{i-1}(x), \quad \text{para } i = 0, 1, \dots, n-1 \quad (8.2)$$

em que

$$P_{-1}(x) = 0 \text{ e } P_0(x) = 1,$$

$$B_i = \frac{\sum_{j=1}^m x_j P_i^2(x_j)}{\sum_{j=1}^m P_i^2(x_j)}, \quad \text{para todo o } i$$

$$\mathbb{C}_0 = 0 \text{ e } \mathbb{C}_i = \frac{\sum_{j=1}^m P_i^2(x_j)}{\sum_{j=1}^m P_{i-1}^2(x_j)} \quad \text{para } i > 0.$$

passo 2 Calculam-se os coeficientes do polinômio, $c_0, c_1, c_2, \dots, c_n$, usando (8.3).

$$c_i = \frac{\sum_{j=1}^m f_j P_i(x_j)}{\sum_{j=1}^m P_i^2(x_j)}, \quad i = 0, 1, \dots, n. \quad (8.3)$$

passo 3 Forma-se o polinômio pretendido.

$$p_n(x) = c_0 P_0(x) + c_1 P_1(x) + c_2 P_2(x) + \dots + c_n P_n(x).$$

Exemplo 8.1 Construir um polinómio de grau 1, $p_1(x)$ (Figura 8.2)

Se o número de pontos for $m = 2$, o polinómio resultante é o polinómio interpolador, por isso passa nos pontos (Figura 8.2(a)). Se o número de pontos $m > 2$, $p_1(x)$ é o polinómio que melhor se ajusta à “mancha” de pontos (Figura 8.2(b)).

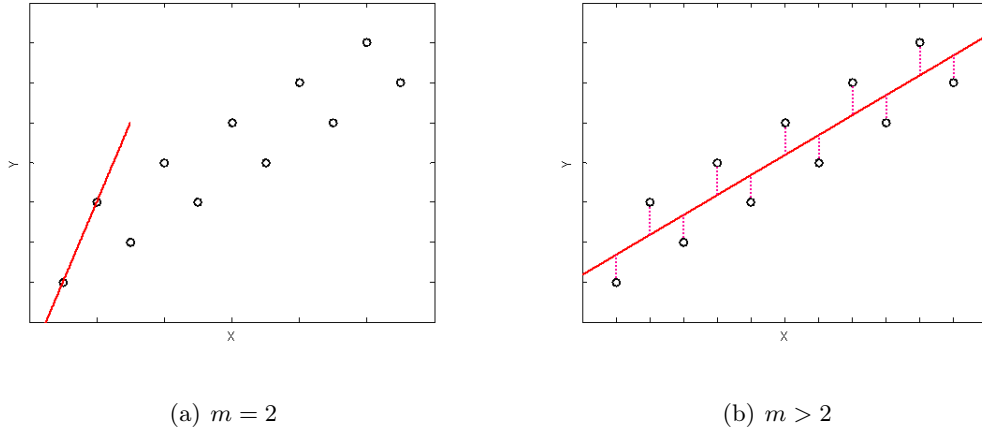


Figura 8.2: Polinómio de grau 1, $p_1(x)$

Exemplo 8.2 Construir um polinómio de grau 1, $p_1(x)$ (Figura 8.3)

Neste caso todos os polinómios $p_1(x)$ são iguais e passam nos pontos de $f(x)$ porque os pontos pertencem a um polinómio de grau 1, o que significa que $\sum_{j=1}^m (f_j - p_1(x_j))^2 = 0$.

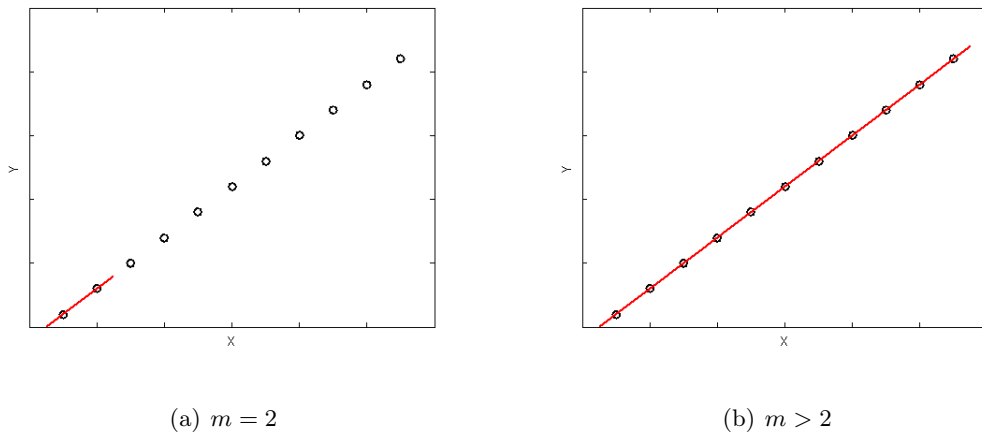


Figura 8.3: Polinómio de grau 1, $p_1(x)$

8.3 Modelo linear não polinomial

O modelo linear não polinomial tem a forma

$$M(x; c_1, \dots, c_n) = c_1 \Phi_1(x) + c_2 \Phi_2(x) + \dots + c_n \Phi_n(x).$$

Este modelo é linear nos coeficientes, c_1, c_2, \dots, c_n e $\Phi_1(x), \Phi_2(x), \dots, \Phi_n(x)$ são funções. No caso deste modelo, a única coisa a determinar são os coeficientes, uma vez que as funções Φ_i , $i = 1, \dots, n$ são dadas. O número de termos na definição do modelo caracteriza a dimensão do problema, n , que é também o número de coeficientes a determinar.

A condição única para que este problema seja bem definido é que o número de pontos onde a função é definida seja maior ou igual ao número de parâmetros do modelo, isto é, $m \geq n$. Se $m = n$, o modelo passa em todos os pontos da função e por isso $\sum_{j=1}^m (f_j - p_1(x_j))^2 = 0$.

O cálculo dos coeficientes c_1, c_2, \dots, c_n é feito a partir do sistema das equações normais.

8.3.1 Sistema das equações normais

Seja o modelo

$$M(x; c_1, \dots, c_n) = c_1 \Phi_1(x) + c_2 \Phi_2(x) + \dots + c_n \Phi_n(x).$$

No sentido dos mínimos quadrados, o objetivo é encontrar o modelo $M(x; c_1, \dots, c_n)$ resultante de

$$\underset{c_1, \dots, c_n}{\text{minimizar}} \ S(c_1, c_2, \dots, c_n) \equiv \sum_{j=1}^m (f_j - M(x_j; c_1, \dots, c_n))^2,$$

isto é,

$$\underset{c_1, \dots, c_n}{\text{minimizar}} \sum_{j=1}^m (f_j - (c_1 \Phi_1(x_j) + c_2 \Phi_2(x_j) + \dots + c_n \Phi_n(x_j)))^2.$$

Como se pretende calcular c_1, c_2, \dots, c_n de forma a que $S(c_1, \dots, c_n)$ seja mínimo, usa-se cálculo diferencial, ou seja, deriva-se $S(c_1, \dots, c_n)$ em ordem a cada um dos coeficientes do modelo e igualam-se essas derivadas parciais a zero.

$$\begin{cases} \frac{\partial S}{\partial c_1} = -2 \sum_{j=1}^m (f_j - c_1 \Phi_1(x_j) - c_2 \Phi_2(x_j) - \dots - c_n \Phi_n(x_j)) \Phi_1(x_j) = 0 \\ \frac{\partial S}{\partial c_2} = -2 \sum_{j=1}^m (f_j - c_1 \Phi_1(x_j) - c_2 \Phi_2(x_j) - \dots - c_n \Phi_n(x_j)) \Phi_2(x_j) = 0 \\ \dots \\ \frac{\partial S}{\partial c_n} = -2 \sum_{j=1}^m (f_j - c_1 \Phi_1(x_j) - c_2 \Phi_2(x_j) - \dots - c_n \Phi_n(x_j)) \Phi_n(x_j) = 0, \end{cases}$$

ou seja

$$\begin{cases} \sum_{j=1}^m f_j \Phi_1(x_j) - \sum_{j=1}^m c_1 \Phi_1(x_j) \Phi_1(x_j) - \cdots - \sum_{j=1}^m c_n \Phi_n(x_j) \Phi_1(x_j) = 0 \\ \sum_{j=1}^m f_j \Phi_2(x_j) - \sum_{j=1}^m c_1 \Phi_1(x_j) \Phi_2(x_j) - \cdots - \sum_{j=1}^m c_n \Phi_n(x_j) \Phi_2(x_j) = 0 \\ \cdots \\ \sum_{j=1}^m f_j \Phi_n(x_j) - \sum_{j=1}^m c_1 \Phi_1(x_j) \Phi_n(x_j) - \cdots - \sum_{j=1}^m c_n \Phi_n(x_j) \Phi_n(x_j) = 0, \end{cases}$$

ou ainda

$$\begin{cases} c_1 \sum_{j=1}^m \Phi_1^2(x_j) + c_2 \sum_{j=1}^m \Phi_2(x_j) \Phi_1(x_j) + \cdots + c_n \sum_{j=1}^m \Phi_n(x_j) \Phi_1(x_j) = \sum_{j=1}^m f_j \Phi_1(x_j) \\ c_1 \sum_{j=1}^m \Phi_1(x_j) \Phi_2(x_j) + c_2 \sum_{j=1}^m \Phi_2^2(x_j) + \cdots + c_n \sum_{j=1}^m \Phi_n(x_j) \Phi_2(x_j) = \sum_{j=1}^m f_j \Phi_2(x_j) \\ \cdots \\ c_1 \sum_{j=1}^m \Phi_1(x_j) \Phi_n(x_j) + c_2 \sum_{j=1}^m \Phi_2(x_j) \Phi_n(x_j) + \cdots + c_n \sum_{j=1}^m \Phi_n^2(x_j) = \sum_{j=1}^m f_j \Phi_n(x_j). \end{cases}$$

Este sistema $n \times n$ é linear nos coeficientes a determinar, c_1, \dots, c_n . Na forma matricial,

$$\begin{pmatrix} \sum_{j=1}^m \Phi_1^2(x_j) & \sum_{j=1}^m \Phi_2(x_j) \Phi_1(x_j) & \cdots & \sum_{j=1}^m \Phi_n(x_j) \Phi_1(x_j) \\ \sum_{j=1}^m \Phi_1(x_j) \Phi_2(x_j) & \sum_{j=1}^m \Phi_2^2(x_j) & \cdots & \sum_{j=1}^m \Phi_n(x_j) \Phi_2(x_j) \\ \cdots & \cdots & \ddots & \cdots \\ \sum_{j=1}^m \Phi_1(x_j) \Phi_n(x_j) & \sum_{j=1}^m \Phi_2(x_j) \Phi_n(x_j) & \cdots & \sum_{j=1}^m \Phi_n^2(x_j) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m f_j \Phi_1(x_j) \\ \sum_{j=1}^m f_j \Phi_2(x_j) \\ \cdots \\ \sum_{j=1}^m f_j \Phi_n(x_j) \end{pmatrix}.$$

A resolução do sistema linear das equações normais fornece os coeficientes pretendidos c_1, \dots, c_n e deve ser feita por um método direto e estável, por exemplo, a eliminação de Gauss com pivotagem parcial.

Para calcular o modelo na forma

$$M(x; c_1, \dots, c_n) = c_1 \Phi_1(x) + c_2 \Phi_2(x) + \cdots + c_n \Phi_n(x)$$

devem seguir-se os passos seguintes.

passo 1 Identificar

a) n , ou seja, o número de termos ou coeficientes, que caracterizam a dimensão do sistema.

b) as n funções $\Phi_1(x), \Phi_2(x), \dots, \Phi_n(x)$.

passo 2 Formar o sistema das equações normais nas n equações e n incógnitas c_1, \dots, c_n , na forma matricial

$$\begin{pmatrix} \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix} = \begin{pmatrix} \dots \\ \dots \\ \dots \\ \dots \end{pmatrix}$$

passo 3 Resolver este sistema por EGPP.

passo 4 Formar o modelo pretendido

$$M(x; c_1, \dots, c_n) = c_1 \Phi_1(x) + c_2 \Phi_2(x) + \dots + c_n \Phi_n(x).$$

Exemplo 8.3 A resistência de um certo fio (de uma certa substância), $f(x)$, varia com o diâmetro desse fio, x . A partir de uma experiência registaram-se os seguintes valores:

x_j	1.5	2.0	3.0	4.0
f_j	4.9	3.3	2.0	1.5

Foram sugeridos os seguintes modelos para ajustar os valores de $f(x)$, no sentido dos mínimos quadrados:

i. uma reta

ii. o modelo linear $M(x, c_1, c_2) = \frac{c_1}{x} + c_2 x$

a) Calcule a reta.

b) Calcule o modelo $M(x)$.

c) Qual dos modelos escolheria? Justifique a sua resposta.

Resolução:

$$a) \ p_1(x) = c_0 P_0(x) + c_1 P_1(x)$$

	x_i	f_i	$P_1(x_i)$	$P_1(x_i)^2$	$f_i P_1(x_i)$
	1.5	4.9	-1.125	1.265625	-5.5125
	2	3.3	-0.625	0.390625	-2.0625
	3	2	0.375	0.140625	0.75
	4	1.5	1.375	1.890625	2.0625
Σ	10.5	11.7		3.6875	-4.7625

$$\bullet \ P_0(x) = 1, \quad C_0 = 0, \quad P_{-1}(x) = 0$$

$$\bullet \ P_1(x) = x - B_0$$

$$B_0 = \frac{\sum x P_0(x)^2}{\sum P_0(x)^2} = \frac{10.5}{4} = 2.625$$

$$P_1(x) = x - 2.625$$

$$\bullet \ c_0 = \frac{\sum f P_0(x)}{\sum P_0(x)^2} = \frac{11.7}{4} = 2.925$$

$$\bullet \ c_1 = \frac{\sum f P_1(x)}{\sum P_1(x)^2} = \frac{-4.7625}{3.6875} = -1.291525$$

$$p_1(x) = 2.925 - 1.291525(x - 2.625)$$

$$b) \ M(x; c_1, c_2) = \frac{c_1}{x} + c_2 x$$

$$\begin{cases} \Phi_1(x) = \frac{1}{x} \\ \Phi_2(x) = x \end{cases}$$

	x_i	f_i	$\Phi_1(x_i)$	$\Phi_2(x_i)$	$\Phi_1(x_i)^2$	$\Phi_2(x_i)^2$	$\Phi_1(x_i)\Phi_2(x_i)$	$f_i \Phi_1(x_i)$	$f_i \Phi_2(x_i)$
	1.5	4.9	0.666667	1.5	0.444444	2.25	1	3.266667	7.35
	2	3.3	0.5	2	0.25	4	1	1.65	6.6
	3	2	0.333333	3	0.111111	9	1	0.666667	6
	4	1.5	0.25	4	0.0625	16	1	0.375	6
Σ					0.868055	31.25	4	5.958334	25.95

$$\left(\begin{array}{cc|c} \sum_i \Phi_1(x_i)^2 & \sum_i \Phi_1(x_i)\Phi_2(x_i) & \sum_i f_i\Phi_1(x_i) \\ \sum_i \Phi_1(x_i)\Phi_2(x_i) & \sum_i \Phi_2(x_i)^2 & \sum_i f_i\Phi_2(x_i) \end{array} \right) \longrightarrow$$

$$\left(\begin{array}{cc|c} 0.868055 & 4 & 5.958334 \\ 4 & 31.25 & 25.95 \end{array} \right) \longrightarrow \begin{cases} c_1 = 7.405414 \\ c_2 = -0.117493 \end{cases}$$

$$M(x) = \frac{7.405414}{x} - 0.117493x$$

c) Cálculo da soma dos quadrados dos resíduos

	x_i	f_i	$p_1(x_i)$	$M(x_i)$	$(f_i - p_1(x_i))^2$	$(f_i - M(x_i))^2$
	1.5	4.9	4.377966	4.760703	0.27252	0.019404
	2	3.3	3.732203	3.467721	0.1868	0.02813
	3	2	2.440678	2.115992	0.194197	0.013454
	4	1.5	1.149153	1.381382	0.123094	0.01407
Σ					0.776611	0.075058

O modelo $M(x)$ ajusta-se melhor no sentido dos mínimos quadrados porque a soma dos quadrados dos resíduos é menor que para o modelo $p_1(x)$ ($0.075058 < 0.776611$).

8.4 Aproximação dos mínimos quadrados

1. Um carro inicia a sua marcha num dia frio de inverno e um aparelho mede o consumo de gasolina verificado no instante em que percorreu x Km. Os resultados obtidos foram:

x (Km)	0	1.25	2.5	3.75	5	6.25
$f(x)$ (l Km ⁻¹)	0.260	0.208	0.172	0.145	0.126	0.113

Construa um modelo quadrático, para descrever o consumo de gasolina em função da distância percorrida, usando a técnica dos mínimos quadrados.

2. A tabela seguinte contém os registos efetuados dos valores médios da radiação solar numa região de Portugal:

mês (x_i)	J(1)	F(2)	M(3)	A(4)	M(5)	J(6)	J(7)	A(8)	S(9)	O(10)	N(11)	D(12)
Radiação	122	-	188	-	-	270	-	-	-	160	-	120

Ajuste o modelo

$$M(x) = c_1x + c_2\sin(x)$$

aos valores da tabela, no sentido dos mínimos quadrados, e use o modelo encontrado para prever a radiação média no mês de agosto.

3. A resistência de um certo fio (de uma certa substância), $f(x)$, varia com o diâmetro desse fio, x . A partir de uma experiência registaram-se os seguintes valores:

x_j	1.5	2.0	3.0	4.0
f_j	4.9	3.3	2.0	1.5

Foram sugeridos os seguintes modelos para ajustar os valores de $f(x)$, no sentido dos mínimos quadrados:

i. uma reta

ii. o modelo linear $M(x, c_1, c_2) = \frac{c_1}{x} + c_2x$

a) Calcule a reta.

b) Calcule o modelo $M(x)$.

c) Qual dos modelos escolheria? Justifique a sua resposta.

4. Um sistema simples de comunicações pode ser representado por um transmissor e um recetor. O transmissor recebe um símbolo, m , e modula o sinal a transmitir, $s_m(t)$, num canal com ruído. O recetor recebe o sinal modulado com o ruído adicionado, $y(t)$, e prevê qual foi o símbolo transmitido. Neste sistema simples, suponha que o transmissor apenas transmite dois sinais

$$s_1(t) = 0.2\alpha_1 \sin(20\pi t) + 0.2\beta_1 \sin(22\pi t)$$

$$s_2(t) = 0.2\alpha_2 \sin(20\pi t) + 0.2\beta_2 \cos(20\pi t)$$

- a) Transmitindo o primeiro sinal ($s_1(t)$) e fazendo uma análise ao transmissor, observaram-se os seguintes valores:

t_i	0.11	0.52	0.79
s_{1i}	-3.1127	0.0625	3.0351

Determine os valores de α_1 e β_1 , no sentido dos mínimos quadrados.

- b) Suponha que $\alpha_1 = -10$, $\beta_1 = -10$, $\alpha_2 = 10$ e $\beta_2 = 10$. Sabendo que o recetor recebeu o sinal indicado na tabela seguinte, determine qual foi o sinal transmitido (isto é, aquele que se ajusta melhor ao sinal recebido, no sentido dos mínimos quadrados).

t_i	0.1	0.45	0.63
$y(t_i)$	1.9963	-2.0100	1.2742

5. Uma companhia de gás sugeriu um modelo do tipo

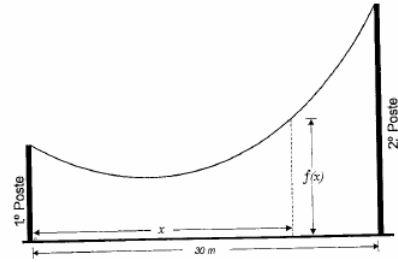
$$M(x; c_1, c_2) = c_1 x^2 + c_2 \frac{1}{x}$$

para estimar o consumo de gás em qualquer altura do ano. No sentido dos mínimos quadrados e considerando a amostra de 6 pontos,

mês	1	3	4	6	9	12
consumo de gás	20.0	7.5	6.5	7.0	10.0	A

- a) comece por apresentar o sistema de equações lineares que deve construir para calcular os parâmetros c_1 e c_2 , em função de A .
- b) Considerando $A = 15.0$, apresente o modelo sugerido.

6. Um fio está suspenso entre dois postes. A distância entre os postes é de 30 metros. A distância do fio ao solo $f(x)$, em metros, depende de x como mostra a figura. A tabela mostra 5 valores conhecidos de f .



x_i	0	8	12	16	20
$f(x_i)$	15.43	10.2	10.2	11.86	15.43

- a) Calcule a parábola que melhor se ajusta aos valores de $f(x_i)$ no sentido dos mínimos quadrados e determine a distância do fio ao solo quando $x = 10$.
- b) A partir da parábola da alínea anterior, verifique se $x = 10$ é o ponto em que a distância do fio ao solo é mínima.
- c) Determine os coeficientes c_1 e c_2 do modelo

$$M(x; c_1, c_2) = c_1 e^{1-0.1x} + c_2 e^{0.1x-1}$$

que melhor se ajusta à função $f(x)$ de acordo com

$$\min_{c_1, c_2} \sum_{i=1}^5 (f(x_i) - M(x_i; c_1, c_2))^2.$$

7. Pretende-se ajustar o modelo linear

$$M(x; c_1, c_2, c_3) = c_1 e^{-x} + c_2 x + c_3$$

à função $f(x)$ dada pela tabela

x_i	-1	0	1	2
$f(x_i)$	1.4	0	0.75	2.3

no sentido dos mínimos quadrados. Determine os coeficientes do modelo apresentado. Apresente uma estimativa para $f(0.5)$.

8. Considere as seguintes observações relativas à função f

x_i	-3	0	2	5
f_i	-10	a	0	b

Determine a e b sabendo que a aproximação polinomial de grau 1 dos mínimos quadrados é $p_1(x) = -4 + 2x$.

8.5 Soluções

1. $p_2(x) = 0.170667 - 0.02304(x - 3.125) + 0.003006 [(x - 3.125)^2 - 4.557292]$.
2. $M(x) = 24.720381x + 148.315675 \sin(x)$; $f(8) \approx M(8) = 344.500384$.
3. a) $p_1(x) = 2.925 - 1.291525(x - 2.625)$.
 b) $M(x) = \frac{7.405414}{x} - 0.117493x$.
 c) $M(x)$.
4. a) $\begin{cases} \alpha_1 &= -9.9997 \\ \beta_1 &= -10.0001 \end{cases}$.
 b) s_2 .
5. a) $\left(\begin{array}{cc|c} 28931 & 35 & 1253.5 + 144A \\ 35 & 1.220679 & 26.402778 + \frac{A}{12} \end{array} \right)$.
 b) $M(x) = 0.093837x^2 + 19.963194\frac{1}{x}$.
6. a) $p_2(x) = 12.624 - 0.019358(x - 11.2) + 0.054579[(x - 8.475676)(x - 11.2) - 47.36]$;
 $f(10) \approx p_2(10) = 9.962539$.
 b) x é mínimo.
 c) $\begin{cases} c_1 &= 2.846971 \\ c_2 &= 6.077745 \end{cases}$.
7. $\begin{cases} c_1 &= 1.984748 \\ c_2 &= 2.008409 \\ c_3 &= -1.986356 \end{cases} ; f(0.5) \approx M(0.5) = 0.221659$.
8. $\begin{cases} a &= -4 \\ b &= 6 \end{cases}$

8.6 MATLAB

8.6.1 polyfit e polyval

À semelhança do que se fez para aproximar polinómios interpoladores, pode usar-se a função `polyfit` para estimar modelos polinomiais no sentido dos mínimos quadrados. Neste caso usam-se todos os pontos disponíveis. A função pode devolver, como segundo parâmetro de saída, a norma do resíduo, pelo que para se obter o resíduo este valor terá de ser elevado ao quadrado.

`[p,r] = polyfit(x,y,n)`

`p` devolve os coeficientes do polinómio em potências descendentes e `r` uma estrutura que contém no seu último termo a norma do resíduo (`normr`). `x` e `y` são dois vetores que contém os dados que se pretendem aproximar e `n` é o grau do polinómio.

Para se fazer a estimativa obtida pelo polinómio interpolador calculado num ponto, ou num conjunto de pontos, usa-se a função `polyval`, tal como foi descrito na Subsecção 5.7.2.

8.6.2 lsqcurvefit

A função `lsqcurvefit` resolve qualquer problema de mínimos quadrados. Como tal, é adequada para problemas não polinomiais. A sua sintaxe é

`[c,S] = lsqcurvefit('mq',c0,x,y).`

`c` é o vetor que contém os parâmetros do modelo e `S` é a soma dos quadrados dos resíduos. `mq` é uma `m-file` que contém o modelo que se pretende estimar, `c0` é o vetor com os valores iniciais dos parâmetros (caso não seja dito nada deve considerar-se um vetor de uns) e `x` e `y` são os vetores que contém a tabela de dados a aproximar.

Exemplo 8.4 *Pretende aproximar-se a tabela*

x_i	1	3	4	5	6
y_i	0.1	0.2	1	0.5	3

a um polinómio de grau dois e a um modelo do tipo $M(x;c) = c_1 \ln(x) + \frac{c_2}{x}$, no sentido dos mínimos quadrados, para estimar o valor em $x = 2$. Qual dos modelos é melhor, no sentido dos mínimos quadrados?

Para se determinar o polinómio de grau dois usa-se a função `polyfit`, depois de se introduzir a tabela sob a forma de dois vetores `x` e `y`:

```
>> [p,r]=polyfit(x,y,2)
```

```
p =
```

```
    0.1832    -0.7973     0.8021
```

```
r =
```

```
    R: [3x3 double]
```

```
    df: 2
```

```
    normr: 1.0895
```

O modelo pretendido é $p_2(x) = 0.1832x^2 - 0.7973x + 0.8021$ e a soma dos quadrados dos resíduos é $S = r^2 = 1.1870$. O valor em $x = 2$ é calculado com a função `polyval`:

```
>> polyval(p,2)
```

```
ans =
```

```
-0.0598
```

Para se determinar o modelo, começa-se por construir a `mfile`.

```
function M=MQex(c,x)
```

```
M=c(1)*log(x)+c(2)./x;
```

Os comandos são

```
>> [c,S]=lsqcurvefit('MQex',[1 1],x,y)
```

```
c =
```

```
    0.8928    -0.1400
```

```
S =
```

```
    3.4906
```

O modelo pretendido é $M(x) = 0.8928 \ln(x) - \frac{0.1400}{x}$ e a soma dos quadrados dos resíduos é $S = 3.4906$. Para se calcular a estimativa em $x = 2$ recorre-se à `m-file` que contém o modelo:

```
>> MQex(c,2)
```

```
ans =
```

```
    0.5489
```

Comparando os a soma dos quadrados dos resíduos, verifica-se que o modelo polinomial $(p_2(x))$ aproxima melhor os dados, no sentido dos mínimos quadrados, que o modelo $M(x)$.

8.6.3 Exercícios para resolver em MATLAB

1. Considere a seguinte tabela:

x_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4
f_i	1.000	1.221	1.492	1.882	2.226	2.718	3.320	4.056

Com base nos mínimos quadrados:

- Escreva um polinómio de grau 3.
 - qual a aproximação para o ponto $x = 0.5$, usando o polinómio da alínea anterior?
 - qual o resíduo do erro?
2. A docente responsável pela UC de MN&ONL registou, para 8 alunos, os resultados obtidos num teste e a respetiva classificação final obtida.

teste	1.2	1.75	1.1	2.0	0.5	0.8	1.0	1.5
classificação final	16	18	16	19	10	11	14	16

- Determine, no sentido dos mínimos quadrados, a reta que melhor aproxima os dados da tabela.
 - Qual o resíduo do erro obtido?
 - Qual será a classificação previsível para um aluno que tenha neste teste uma classificação de 1.6?
3. Considerem-se as seguintes funções de aproximação

um polinómio de grau 3 ($p_3(x)$)

$$M(x) = c_1 + c_2 \cos(x) + c_3 \sin(x)$$

$$N(x) = c_1 e^x + c_2 \frac{1}{x}$$

$$O(x) = c_1 + c_2 x + \frac{c_3}{x}$$

$$Q(x) = c_1 x + c_2 e^x$$

- Calcule os coeficientes dos vários modelos (e construa-os) que melhor se ajustam à função $f(x)$ dada pela tabela seguinte, no sentido dos mínimos quadrados.

x_i	-1.00	-0.95	-0.85	-0.80	0.20	0.50	0.90
f_i	-1.00	-0.05	0.90	1.00	0.90	0.50	-0.30

- b) Estime $f(0.6)$ para cada um deles.
- c) Indique o resíduo para cada um dos modelos.
- d) Qual dos modelos é melhor, no sentido dos mínimos quadrados? Justifique.

Capítulo 9

Otimização não linear sem restrições

A otimização surge no processo de tomada de decisão para se atingir o melhor resultado possível. Dificilmente se consegue imaginar uma área de estudo em que os princípios da otimização não estejam presentes. Na realidade, mesmo no dia a dia, tudo se tenta otimizar - pretende-se sempre o mínimo ou o máximo de algo. É, pois, um dos objetivos dos profissionais das áreas das Ciências de Gestão e Engenharia, estando também presente noutras áreas aplicadas, tais como a economia, as finanças, a medicina ou a estatística.

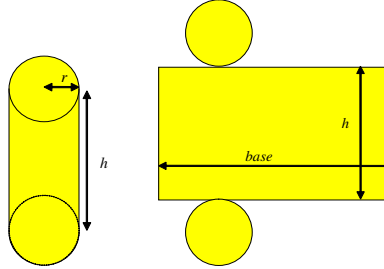
A otimização está relacionada com a maximização ou minimização de modelos matemáticos, e a função que se pretende otimizar é denominada *função objetivo*. No entanto, por vezes as variáveis de decisão estão sujeitas a determinadas condições, designadas por *restrições*.

Os problemas de otimização podem ser classificados de acordo com as características da função objetivo e das restrições, em duas grandes áreas: a Otimização Linear e a Otimização Não Linear (ONL). No primeiro caso, todas as funções (função objetivo e restrições) envolvidas no problema são lineares. Na ONL, pelo menos uma das funções, entre a função objetivo e as restrições, é não linear nas variáveis. Ainda na otimização não linear, há alguns casos particulares mais fáceis de resolver:

- problemas sem restrições nas variáveis;
- problemas quadráticos;
- problemas convexos;
- problemas de mínimos quadrados lineares.

Exemplo 9.1 *Problema com duas variáveis e uma restrição*

Tendo como objetivo fabricar latas cilíndricas com um volume de 1000 cm^3 e tapá-las em ambas as extremidades, qual deverá ser o raio da base e a altura da lata de modo a minimizar a quantidade de placa metálica, em termos de área superficial?



$$\begin{aligned}
 \text{Área Total} &= \text{Área}_{\text{retângulo}} + 2 \times \text{Área}_{\text{círculo}} \\
 &= \text{base} \times h + 2(\pi r^2) \\
 &= \text{Perímetro}_{\text{círculo}} \times h + 2\pi r^2 \\
 &= 2\pi r h + 2\pi r^2
 \end{aligned}$$

$$\text{Volume} = \pi r^2 \times h$$

$$1000 = \pi r^2 \times h$$

Formulação do problema:

$$\begin{aligned}
 &\text{minimizar} \quad A(r, h) \equiv 2\pi r h + 2\pi r^2 \\
 &\text{sujeito a} \quad \pi r^2 h = 1000
 \end{aligned}$$

Exemplo 9.2 *Problema com três variáveis e uma restrição*

O produto de três números positivos é igual a A (dado). Determine esses números por forma que a sua soma seja máxima.

$$\begin{aligned}
 &\text{maximizar} \quad x_1 + x_2 + x_3 \\
 &\text{sujeito a} \quad x_1 x_2 x_3 = A
 \end{aligned} \tag{9.1}$$

9.1 Forma geral do problema

Um problema de otimização sem restrições, em termos gerais, pode ser definido da seguinte forma:

$$\min_{x \in \mathbb{R}^n} f(x) \tag{9.2}$$

em que $f(x)$ é a função que se pretende minimizar.

Num problema de otimização, pretende-se minimizar ou maximizar um objetivo $f(x)$ - função objetivo do problema.

Num problema de minimização pretende-se calcular um ponto $x^* \in \mathbb{R}^n$, denominado minimizante, que verifique $f(x^*) \leq f(x)$ para todo o $x \in \mathbb{R}^n$.

Num problema de maximização pretende-se calcular $x^* \in \mathbb{R}^n$, denominado maximizante, que verifique $f(x^*) \geq f(x)$ para todo o $x \in \mathbb{R}^n$.

9.2 Classificação de mínimos e máximos

Seja $V(x, \delta)$ uma vizinhança de x^* de raio δ ($\delta > 0$). x^* é

- minimizante local forte se $\exists \delta > 0$:

- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) < f(x), \forall x \in V(x^*, \delta); x \neq x^*$

- minimizante local fraco se $\exists \delta > 0$:

- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) \leq f(x), \forall x \in V(x^*, \delta); x \neq x^*$

- maximizante local forte se $\exists \delta > 0$:

- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) > f(x) \forall x \in V(x^*, \delta); x \neq x^*$

- maximizante local fraco se $\exists \delta > 0$:

- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) \geq f(x) \forall x \in V(x^*, \delta); x \neq x^*$

- minimizante global forte se $f(x^*) < f(x)$, para todo o x que pertence ao domínio de $f(x)$, onde a função é definida.

- minimizante global fraco se $f(x^*) \leq f(x)$, para todo o x que pertence ao domínio de $f(x)$, onde a função é definida.
- maximizante global forte se $f(x^*) > f(x)$, para todo o x que pertence ao domínio de $f(x)$, onde a função é definida.
- maximizante global fraco se $f(x^*) \geq f(x)$, para todo o x que pertence ao domínio de $f(x)$, onde a função é definida.

Todo o ótimo global é local. No entanto, um ótimo local pode não ser global.

Exemplo 9.3 *Alguns mínimos e máximos de uma função unidimensional*

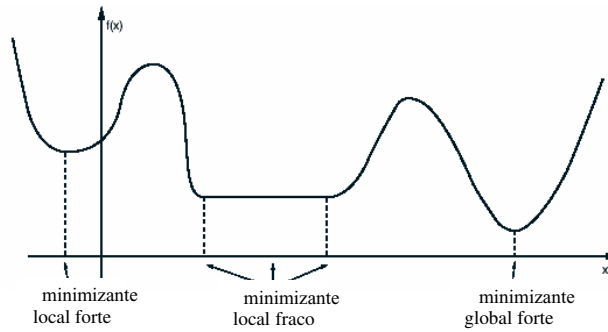


Figura 9.1: Representação gráfica de alguns minimizantes

Exemplo 9.4 *Classificação de mínimos e máximos em algumas funções bi-dimensionais*

9.3 Mínimos *versus* máximos

Em geral, os métodos de otimização estão formulados para o cálculo de mínimos. No entanto, podem também ser usados no cálculo de máximos, já que estes podem ser facilmente relacionados com os mínimos da seguinte forma (Figura 9.3):

$$\max f(x) = -\min(-f(x))$$

$$x^* = \underbrace{\arg \max (f(x))}_{\text{maximizante}} = \underbrace{\arg \min (-f(x))}_{\text{minimizante}}$$

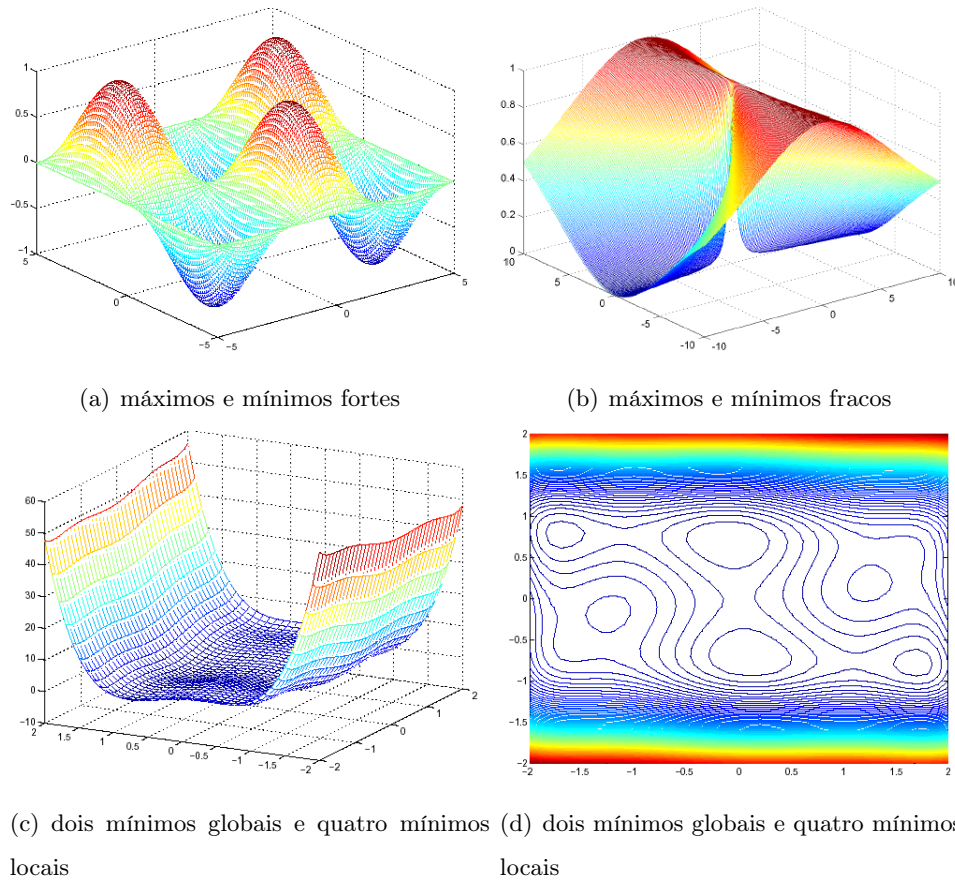


Figura 9.2: Classificação de mínimos e máximos

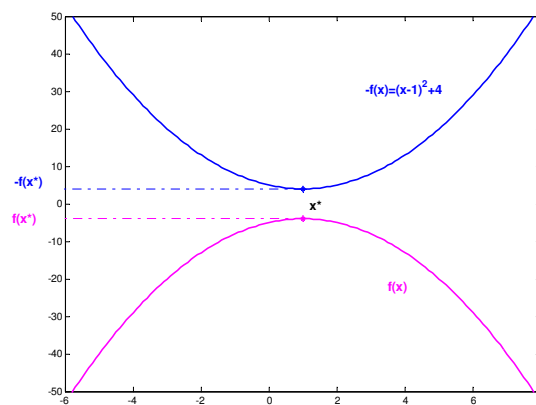


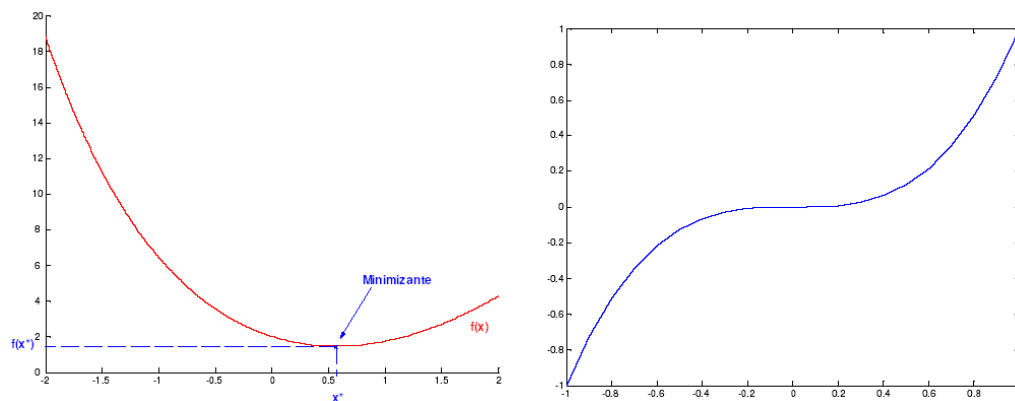
Figura 9.3: Relação entre mínimos e máximos de funções.

Capítulo 10

Otimização unidimensional

Se $n = 1$ (9.2), então está-se perante um problema unidimensional, isto é, com uma só variável, o que significa que x é escalar.

Exemplo 10.5 *Problemas unidimensionais*



(a) $f(x) = x^2 + 2e^{-x}$

(b) $f(x) = x^3$

Figura 10.1: (a) problema com um mínimo; (b) problema sem mínimos (nem máximos)

10.1 Condições de otimalidade

Para aplicar as condições de otimalidade de primeira e segunda ordem, assume-se $f(x)$ continuamente diferenciável até à segunda ordem.

Condição necessária e suficiente de primeira ordem

Se x^* é uma solução do problema (9.2), com $n = 1$, então $f'(x) = 0$. Isto é, a equação $f'(x) = 0$ define os pontos estacionários da função objetivo $f(x)$. Os pontos estacionários podem ser minimizantes, maximizantes ou pontos de inflexão.

Exemplo 10.6 *Determinar os pontos estacionários da função $f(x) = x^2 + 2e^{-x}$.*

$$f'(x) \equiv 2x - 2e^{-x} = 0.$$

A solução desta equação não linear, obtida pelo método iterativo de Newton ou da secante, é única: $x = 0.5671$. Ou seja, $x = 0.567143$ é um ponto estacionário de $f(x)$.

Condição necessária de segunda ordem

Se x^* é uma solução do problema (9.2) para $n = 1$, que satisfaz a condição de primeira ordem, então a condição necessária para que x^* seja minimizante é $f''(x) \geq 0$ e a condição necessária para que x^* seja maximizante é $f''(x) \leq 0$.

Condição suficiente de segunda ordem

Se x^* é solução do problema (9.2) e se $f''(x) > 0$, então x^* é um minimizante local forte de $f(x)$. Se x^* é solução do problema (9.2) e se $f''(x) < 0$, então x^* é um maximizante local forte de $f(x)$.

10.2 Exercícios

1. Dada a função $f(x) = x^3 - 6x^2 + 9x + 4$ calcule os seus pontos estacionários e classifique-os.

10.3 Soluções

1. $x^* = 1$ é maximizante; $x = 3$ é minimizante.

Capítulo 11

Método de DSC

Em termos gerais, os métodos de resolução de problemas unidimensionais podem dividir-se em três grandes grupos:

- Métodos de procura ou pesquisa direta
- Métodos de aproximação
- Métodos mistos

Os métodos mistos combinam as técnicas de procura e de aproximação.

O método de Davies, Swann e Campey (DSC) é um método iterativo que só usa informação da função objetivo f e destina-se a problemas de otimização unidimensional. Trata-se de um método misto, isto é, tem uma fase de procura seguida de uma fase de aproximação baseada em interpolação quadrática.

Fase de procura

Procuram-se, em cada iteração, três pontos igualmente espaçados que definem um intervalo que contém o minimizante da função. Esta procura baseia-se apenas nos valores da função objetivo em diversos pontos.

A procura inicia-se com uma aproximação inicial x_1 e uma perturbação $\delta > 0$. A partir de x_1 e no sentido positivo, calcula-se uma sequência de pontos x_2, x_3, x_4, \dots distanciados uns

dos outros de δ , 2δ , 4δ , $8\delta, \dots$. Assim,

$$x_1$$

$$x_2 = x_1 + \delta$$

$$x_3 = x_2 + 2\delta$$

$$\dots$$

$$x_k = x_{k-1} + 2^{k-2}\delta$$

até que no ponto x_k se tenha $f(x_k) > f(x_{k-1})$. Nesta altura tem-se $\dots < x_{k-2} < x_{k-1} < x_k$,

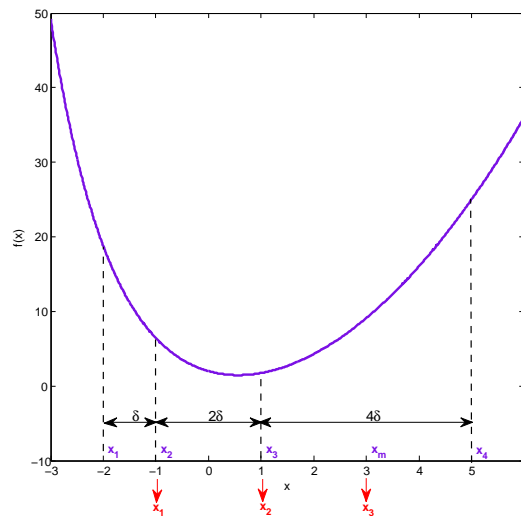


Figura 11.1: Procura do método de DSC para $x_1 = -2$ e $\delta = 1$.

em que $f(x_{k-2}) \geq f(x_{k-1})$ e $f(x_{k-1}) < f(x_k)$. A distância entre x_k e x_{k-1} é duas vezes a distância entre x_{k-1} e x_{k-2} . Para que os pontos estejam igualmente espaçados, calcula-se o ponto médio do último intervalo, $x_m = \frac{x_k + x_{k-1}}{2}$. Fica-se, assim, com quatro pontos igualmente espaçados: $x_{k-2} < x_{k-1} < x_m < x_k$. Para a aproximação quadrática, é necessário seleccionar três dos quatro pontos encontrados na fase de procura. Para isso, comparam-se os valores de $f(x)$ nos dois pontos interiores do intervalo. Se $f(x_{k-1}) \leq f(x_m)$, então escolhem-se os pontos x_{k-2} , x_{k-1} e x_m , caso contrário escolhem-se os pontos x_{k-1} , x_m e x_k . Ver Figura 11.1.

Quando a partir de x_1 o valor de $f(x_2) > f(x_1)$, com $x_2 = x_1 + \delta$, a procura deve voltar-se para o sentido negativo, a começar novamente por x_1 . Neste caso, o próximo ponto na procura é $x_{-1} = x_1 - \delta$. Se $f(x_{-1}) > f(x_1)$, significa que o intervalo definido por $[x_{-1}, x_2]$, com x_1 como ponto médio, contém o minimizante da quadrática que passa pelos três pontos agora

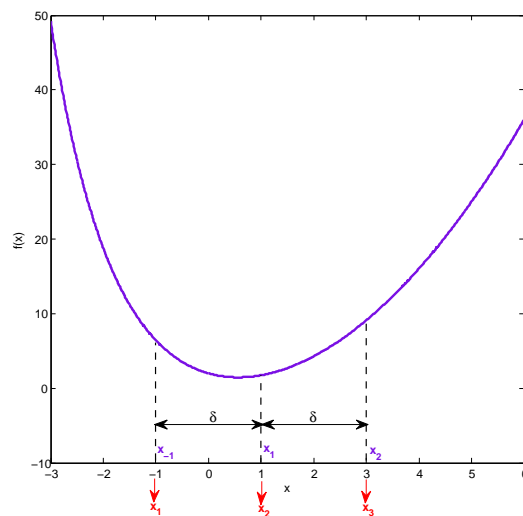


Figura 11.2: Procura do método de DSC para $x_1 = 1$ e $\delta = 2$.

calculados. Ver Figura 11.2.

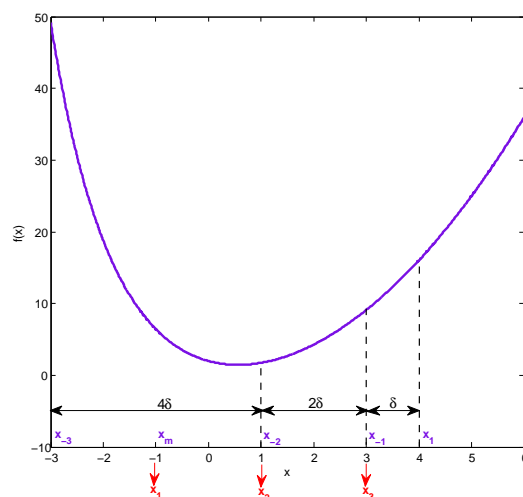


Figura 11.3: Procura do método de DSC para $x_1 = 4$ e $\delta = 1$.

No entanto, se $f(x_{-1}) < f(x_1)$, significa que a procura deve continuar no sentido negativo até que $f(x_{-k}) > f(x_{-(k-1)})$, isto é, procede-se da seguinte forma:

$$x_{-2} = x_{-1} - 2\delta$$

...

$$x_{-k} = x_{-(k-1)} - 2^{k-1}\delta$$

até que no ponto x_{-k} se tenha $f(x_{-k}) > f(x_{-(k-1)})$.

Nesta altura, tem-se

$$x_{-k} < x_{-(k-1)} < x_{-(k-2)} < \dots$$

em que $f(x_{-(k-2)}) \geq f(x_{-(k-1)})$ e $f(x_{-(k-1)}) < f(x_{-k})$, e a distância entre x_{-k} e $x_{-(k-1)}$ é duas vezes a distância entre $x_{-(k-1)}$ e $x_{-(k-2)}$. Para que os pontos estejam igualmente espaçados, calcula-se o ponto médio do último intervalo, $x_m = \frac{x_{-k} + x_{-(k-1)}}{2}$. Fica-se, assim, com quatro pontos igualmente espaçados: $x_{-k} < x_m < x_{-(k-1)} < x_{-(k-2)}$. Para a aproximação quadrática, é necessário seleccionar três dos quatro pontos encontrados na fase de procura. Para isso, comparam-se os valores de $f(x)$ nos dois pontos interiores do intervalo. Se $f(x_m) \leq f(x_{-(k-1)})$, então escolhem-se os pontos x_{-k} , x_m e $x_{-(k-1)}$, caso contrário escolhem-se os pontos x_m , $x_{-(k-1)}$ e $x_{-(k-2)}$. Ver Figura 11.3.

Fase de aproximação

Depois de concluída a fase de procura, entra-se na fase de aproximação, em que se aproxima a função no intervalo obtido por uma função quadrática (polinómio de grau dois) e usa-se o seu minimizante como aproximação ao minimizante da função. Esta quadrática passa pelos três pontos seleccionados.

O minimizante da quadrática, $x^*(q)$, que passa por estes três pontos, que passam a denominar-se $\mathbf{x}_1 < \mathbf{x}_2 < \mathbf{x}_3$, com $\Delta = (\mathbf{x}_2 - \mathbf{x}_1) = (\mathbf{x}_3 - \mathbf{x}_2)$, determina-se por (Figura 11.4)

$$x^*(q) = \mathbf{x}_2 + \Delta \frac{f(\mathbf{x}_1) - f(\mathbf{x}_3)}{2(f(\mathbf{x}_3) - 2f(\mathbf{x}_2) + f(\mathbf{x}_1))}$$

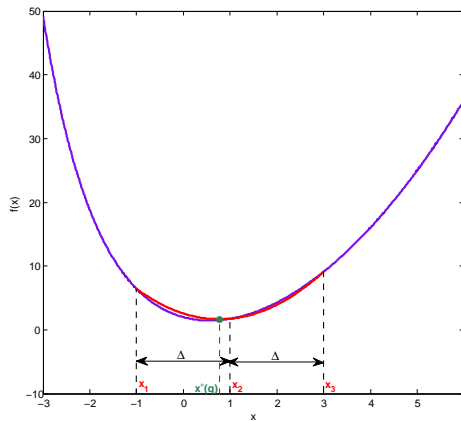


Figura 11.4: Aproximação do método de DSC, com $\Delta = 2$.

Algoritmo 11.1 Método de Davies, Swann e Campey

ler: x_1, δ, M e ε

repetir

$$x_2 \leftarrow x_1 + \delta$$

se $f(x_2) \leq f(x_1)$ **então**

$$k \leftarrow 2$$

repetir

$$k \leftarrow k + 1$$

$$x_k \leftarrow x_{k-1} + 2^{k-2} \delta$$

até $f(x_k) > f(x_{k-1})$

$$x_m \leftarrow \frac{x_k + x_{k-1}}{2}$$

se $f(x_{k-1}) \leq f(x_m)$ **então**

$$\mathbf{x}_1 \leftarrow x_{k-2}, \mathbf{x}_2 \leftarrow x_{k-1}, \mathbf{x}_3 \leftarrow x_m$$

senão

$$\mathbf{x}_1 \leftarrow x_{k-1}, \mathbf{x}_2 \leftarrow x_m, \mathbf{x}_3 \leftarrow x_k$$

fim se

senão

$$x_{-1} = x_1 - \delta$$

se $f(x_{-1}) < f(x_1)$ **então**

$$k \leftarrow 1$$

repetir

$$k \leftarrow k + 1$$

$$x_{-k} \leftarrow x_{-(k-1)} - 2^{k-1} \delta$$

até $f(x_{-k}) > f(x_{-(k-1)})$

$$x_m \leftarrow \frac{x_{-k} + x_{-(k-1)}}{2}$$

se $f(x_m) < f(x_{-(k-1)})$ **então**

$$\mathbf{x}_1 \leftarrow x_{-k}, \mathbf{x}_2 \leftarrow x_m, \mathbf{x}_3 \leftarrow x_{-(k-1)}$$

senão

$$\mathbf{x}_1 \leftarrow x_m, \mathbf{x}_2 \leftarrow x_{-(k-1)}, \mathbf{x}_3 \leftarrow x_{-(k-2)}$$

fim se

senão

$$\mathbf{x}_3 \leftarrow x_2, \mathbf{x}_2 \leftarrow x_1, \mathbf{x}_1 \leftarrow x_{-1}$$

fim se

fim se

$$\Delta \leftarrow (\mathbf{x}_2 - \mathbf{x}_1) = (\mathbf{x}_3 - \mathbf{x}_2)$$

$$x^*(q) \leftarrow \mathbf{x}_2 + \Delta \frac{f(\mathbf{x}_1) - f(\mathbf{x}_3)}{2(f(\mathbf{x}_3) - 2f(\mathbf{x}_2) + f(\mathbf{x}_1))}$$

$$x_1 \leftarrow x^*(q)$$

$$\delta = M\delta$$

até $\Delta \leq \varepsilon$

$$x_{\min} \leftarrow x^*(q)$$

Paragem do método de DSC

O critério de paragem do método iterativo de DSC consiste em a distância entre os pontos que foram usados para construir a função quadrática não exceder uma certa quantidade positiva e próxima de zero, ou seja

$$(\mathbf{x}_2 - \mathbf{x}_1) = (\mathbf{x}_3 - \mathbf{x}_2) = \Delta \leq \varepsilon, \text{ com } \varepsilon > 0 \text{ e } \varepsilon \approx 0.$$

Se o critério de paragem for verificado, o processo iterativo termina, sendo $x^*(q)$ a melhor aproximação calculada à solução. Se o critério de paragem não se verificar, o processo repete-se e o minimizante da quadrática $x^*(q)$ passa a ser o ponto inicial, x_1 , da próxima iteração. A perturbação δ também deve ser atualizada através de $\delta = M\delta$, com $M < 1$.

O Algoritmo 11.1 descreve o método de DSC.

Exemplo 11.7 *Uma empresa precisa de usar x_1 horas de equipamento ao preço (unitário) de 6 unidades monetárias (u.m.) e x_2 horas de mão-de-obra ao preço (unitário) de 4 u.m. para colocar no mercado um certo número fixo de produtos. As horas utilizadas de equipamento e mão-de-obra verificam a relação*

$$x_1^2 + x_1x_2 = 2500.$$

Calcule x_1 e x_2 de modo a minimizar os custos da empresa.

- a) *Comece por formular esta situação como um problema de otimização sem restrições de uma só variável (por exemplo, em função de x_1).*
- b) *Resolva o problema resultante usando o método DSC (baseado em interpolação quadrática). Na implementação do DSC inicie o processo iterativo com a aproximação inicial $x_1 = 50$. Use $\delta = 5$, $\varepsilon = 0.05$ e $M = 0.1$.*

Com a aproximação calculada identifique os valores obtidos para as duas variáveis e o custo mínimo.

Resolução:

- a) *Formular problema sem restrições*

$$\begin{array}{ll} \min & 6x_1 + 4x_2 \\ \text{s.a.} & x_1^2 + x_1x_2 = 2500 \end{array} \Rightarrow x_2 = \frac{2500 - x_1^2}{x_1}$$

$$\min \quad 6x_1 + 4 \times \frac{2500 - x_1^2}{x_1}$$

b) Iniciar o algoritmo DSC: $x_1 = 50, \delta = 5, M = 0.1, \varepsilon = 0.05$

• 1ª iteração

$$\begin{cases} x_1 = 50 \\ f(x_1) = 300 \quad \text{procurar para a direita} \end{cases}$$

$$\begin{cases} x_2 = 50 + \delta = 50 + 5 = 55 \\ f(x_2) = 291.818182 \quad \downarrow \quad \text{continuar} \end{cases}$$

$$\begin{cases} x_3 = 55 + 2 \times \delta = 55 + 2 \times 5 = 65 \\ f(x_3) = 283.846154 \quad \downarrow \quad \text{continuar} \end{cases}$$

$$\begin{cases} x_4 = 65 + 4 \times \delta = 65 + 4 \times 5 = 85 \\ f(x_4) = 287.647059 \quad \uparrow \quad \text{parar e calcular ponto médio} \end{cases}$$

$$\begin{cases} x_m = \frac{65 + 85}{2} = 75 \\ f(x_m) = 283.333333 \end{cases}$$

$$\begin{array}{ccccccc} & & 55 & & \overbrace{65 \quad 75} & & 85 \\ & & & & \underbrace{\quad \quad} & & \\ & & & & 283.846154 & & 283.333333 \end{array}$$

Como $f(x_m) < f(x_1)$ escolher 3 pontos igualmente espaçados: $f(x_3), f(x_m), f(x_4)$

$$\left. \begin{array}{ll} \mathbf{x}_1 \leftarrow 65 & f(\mathbf{x}_1) = 283.846154 \\ \mathbf{x}_2 \leftarrow 75 & f(\mathbf{x}_2) = 283.333333 \\ \mathbf{x}_3 \leftarrow 85 & f(\mathbf{x}_3) = 287.647059 \end{array} \right\} \quad \Delta = 10$$

$$x^*(q) = \mathbf{x}_2 + \Delta \frac{f(\mathbf{x}_1) - f(\mathbf{x}_3)}{2(f(\mathbf{x}_3) - 2f(\mathbf{x}_2) + f(\mathbf{x}_1))} = 71.062501 \quad f(x^*(q)) = 282.846196$$

• Critério de Paragem

$$\Delta \leq \varepsilon \Leftrightarrow 10 \leq 0.05 \quad (\text{falso})$$

$$\delta = M\delta = 5 \times 0.1 = 0.5$$

- **2ª iteração**

$$\begin{cases} x_1 = 71.062501 \\ f(x_1) = 282.846196 \quad \underline{\text{procurar para a direita}} \end{cases}$$

$$\begin{cases} x_2 = 71.062501 + \delta = 71.062501 + 0.5 = 71.562501 \\ f(x_2) = 282.862991 \quad \uparrow \quad \underline{\text{procurar para a esquerda}} \end{cases}$$

$$\begin{cases} x_{-1} = 71.062501 - \delta = 71.062501 - 0.5 = 70.562501 \\ f(x_{-1}) = 282.843335 \quad \downarrow \quad \underline{\text{continuar}} \end{cases}$$

$$\begin{cases} x_{-2} = 70.062501 - 2 \times \delta = 70.062501 - 1 = 69.562501 \\ f(x_{-2}) = 282.880615 \quad \uparrow \quad \underline{\text{parar e calcular ponto médio}} \end{cases}$$

$$\begin{cases} x_m = \frac{70.562501 + 69.562501}{2} = 70.062501 \\ f(x_m) = 282.854706 \end{cases}$$

$$\begin{array}{ccccccc} 69.562501 & & \underbrace{70.062501} & & \underbrace{70.562501} & & 71.062501 \\ & & 282.854706 & & 282.843335 & & \end{array}$$

Como $f(x_m) < f(x_{-1})$ escolher 3 pontos igualmente espaçados: $f(x_{-1}), f(x_m), f(x_1)$

$$\left. \begin{array}{ll} \mathbf{x}_1 \leftarrow 70.062501 & f(\mathbf{x}_1) = 282.854706 \\ \mathbf{x}_2 \leftarrow 70.562501 & f(\mathbf{x}_2) = 282.843335 \\ \mathbf{x}_3 \leftarrow 71.062501 & f(\mathbf{x}_3) = 282.846196 \end{array} \right\} \quad \Delta = 0.5$$

$$x^*(q) = \mathbf{x}_2 + \Delta \frac{f(\mathbf{x}_1) - f(\mathbf{x}_3)}{2(f(\mathbf{x}_3) - 2f(\mathbf{x}_2) + f(\mathbf{x}_1))} = 70.711988 \quad f(x^*(q)) = 282.842713$$

- Critério de Paragem

$$\Delta \leq \varepsilon \Leftrightarrow 0.5 \leq 0.05 \quad (\text{falso})$$

$$\delta = M\delta = 0.5 \times 0.1 = 0.05$$

- **3ª iteração**

$$\begin{cases} x_1 = 70.711988 \\ f(x_1) = 282.842713 \quad \underline{\text{procurar para a direita}} \end{cases}$$

$$\begin{cases} x_2 = 70.711988 + \delta = 70.711988 + 0.05 = 70.761988 \\ f(x_2) = 282.842787 \quad \uparrow \quad \text{procurar para a esquerda} \end{cases}$$

$$\begin{cases} x_{-1} = 70.711988 - \delta = 70.711988 - 0.05 = 70.661988 \\ f(x_{-1}) = 282.842780 \quad \uparrow \quad \text{ordenar pontos} \end{cases}$$

$$\left. \begin{array}{ll} \mathbf{x}_1 \leftarrow 70.661988 & f(\mathbf{x}_1) = 282.842780 \\ \mathbf{x}_2 \leftarrow 70.711988 & f(\mathbf{x}_2) = 282.842713 \\ \mathbf{x}_3 \leftarrow 70.761988 & f(\mathbf{x}_3) = 282.842787 \end{array} \right\} \quad \Delta = 0.5$$

$$x^*(q) = \mathbf{x}_2 + \Delta \frac{f(\mathbf{x}_1) - f(\mathbf{x}_3)}{2(f(\mathbf{x}_3) - 2f(\mathbf{x}_2) + f(\mathbf{x}_1))} = 70.710747 \quad f(x^*(q)) = 282.842713$$

- Critério de Paragem

$$\Delta \leq \varepsilon \Leftrightarrow 0.05 \leq 0.05 \quad (\text{verdadeiro})$$

$x_1 \approx 70.710747, x_2 \approx -35.355442$ e o custo mínimo ≈ 282.842713

11.1 Exercícios

1. Na cidade de Ulam Bator surgiu uma epidemia de gripe asiática. A evolução da doença foi descrita pela fórmula

$$P(t) = e^{0.4t - 0.01t^2}$$

onde $P(t)$ representa a percentagem de pessoas doentes e t é o tempo em dias.

Usando o método DSC (baseado em interpolação quadrática), calcule o pior momento da epidemia identificando a percentagem de doentes nesse momento. Inicie o processo iterativo com $t_1 = 30$ dias. Considere ainda $\delta = 2$, $M = 0.05$ e $\varepsilon = 0.1$ (duas iterações). Use 4 casas decimais nos cálculos.

2. Uma empresa precisa de usar x_1 horas de equipamento ao preço (unitário) de 6 unidades monetárias (u.m.) e x_2 horas de mão-de-obra ao preço (unitário) de 4 u.m. para colocar no mercado um certo número fixo de produtos. As horas utilizadas de equipamento e mão-de-obra verificam a relação

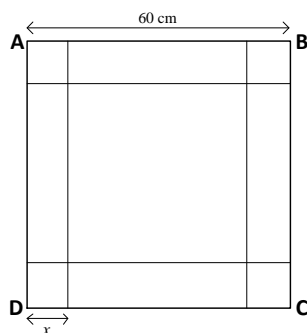
$$x_1^2 + x_1x_2 = 2500.$$

Calcule x_1 e x_2 de modo a minimizar os custos da empresa.

- a) Comece por formular esta situação como um problema de otimização sem restrições de uma só variável (por exemplo, em função de x_1).
- b) Resolva o problema resultante usando o método DSC (baseado em interpolação quadrática). Na implementação do DSC inicie o processo iterativo com a aproximação inicial $x_1 = 50$. Use $\delta = 5$, $\varepsilon = 0.05$ e $M = 0.1$.

Com a aproximação calculada identifique os valores obtidos para as duas variáveis e o custo mínimo.

3. $[ABCD]$ representa uma cartolina quadrada de lado 60 cm. Pretende-se montar uma caixa de volume máximo cortando em cada canto um quadrado de lado x , como mostra a figura.

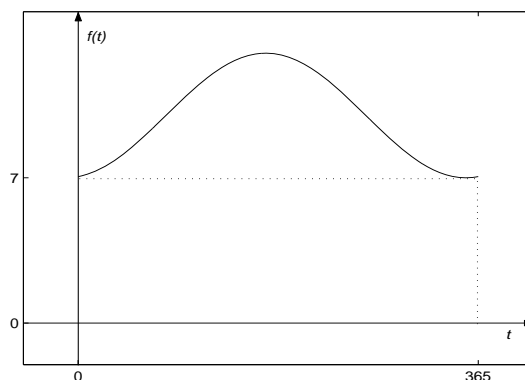


Usando o método DSC (baseado em interpolação quadrática), calcule x . Use duas casas decimais nos cálculos e inicie o processo iterativo com $x_1 = 5$. Considere ainda $\delta = 1$, $M = 0.5$ e $\varepsilon = 0.5$ (duas iterações).

4. A função

$$f(t) = 10 + 3 \sin\left(\frac{2\pi}{365}(t - 80)\right)$$

dá o número de horas com luz do dia numa certa região do país.



O dia 1 de Janeiro corresponde a $t = 0$. Determine o dia do ano (t) em que o número de horas com luz do dia é máximo, usando o método DSC (baseado em interpolação quadrática). Use 2 casas decimais nos cálculos, $\pi = 3.14$ e inicie o processo iterativo com $t_1 = 200$. Considere ainda $\delta = 10$, $M = 0.1$ e $\varepsilon = 2$ (duas iterações). Use radianos nos cálculos.

11.2 Soluções

1. 2 iterações; $P_{\max} = 51.5982\%$; $t_{\max} = 20$ dias.
2. a) $\min 6x_1 + 4\frac{2500-x_1^2}{x_1}$.
b) 3 iterações; $x_1 \approx 70.7107$; $x_2 \approx -35.3554$; custo mínimo ≈ 282.8427 .
3. 2 iterações; $x_{\max} \approx 10.00$; $v_{\max} \approx 16000$.
4. 2 iterações; $t_{\max} \approx 171.74$; $f_{\max} \approx 13.00$.

Capítulo 12

Otimização multidimensional sem restrições

Um problema multidimensional caracteriza-se por envolver mais que uma variável, isto é, $n > 1$ (9.2). Importa, nestes problemas, distinguir os problemas com descontinuidades na função objetivo, uma vez que estes são, em geral, mais difíceis de resolver. Há, no entanto, métodos específicos para esta classe de problemas, conhecida por otimização sem derivadas. Quando os problemas são diferenciáveis, podem ser usados métodos que utilizem informação das derivadas – gradiente e/ou matriz Hessiana – conhecidos por métodos do gradiente.

Exemplo 12.1 *Problemas multidimensionais ($n = 2$) sem restrições*

$$\min_{x \in \mathbb{R}^2} f(x) \equiv (x_1 - 2)^2 + (x_2 - 1)^2 \quad (\text{Figuras 12.1(a) e 12.1(b)})$$

$$\max_{x \in \mathbb{R}^2} f(x) \equiv 2(-x_1^2 - x_2^2 + 1) + x_1 \quad (\text{Figuras 12.1(c) e 12.1(d)})$$

Exemplo 12.2 *Problemas multidimensionais ($n = 2$) sem restrições*

$$\min_{x \in \mathbb{R}^2} f(x) \equiv 3x_1^2 - x_2^2 + x_1^3 \quad (\text{Figuras 12.2(a) e 12.2(b)})$$

$$\min_{x \in \mathbb{R}^2} f(x) \equiv 3x_1^2 - 4x_1x_2 - 4x_2^2 \quad (\text{Figuras 12.2(c) e 12.2(d)})$$

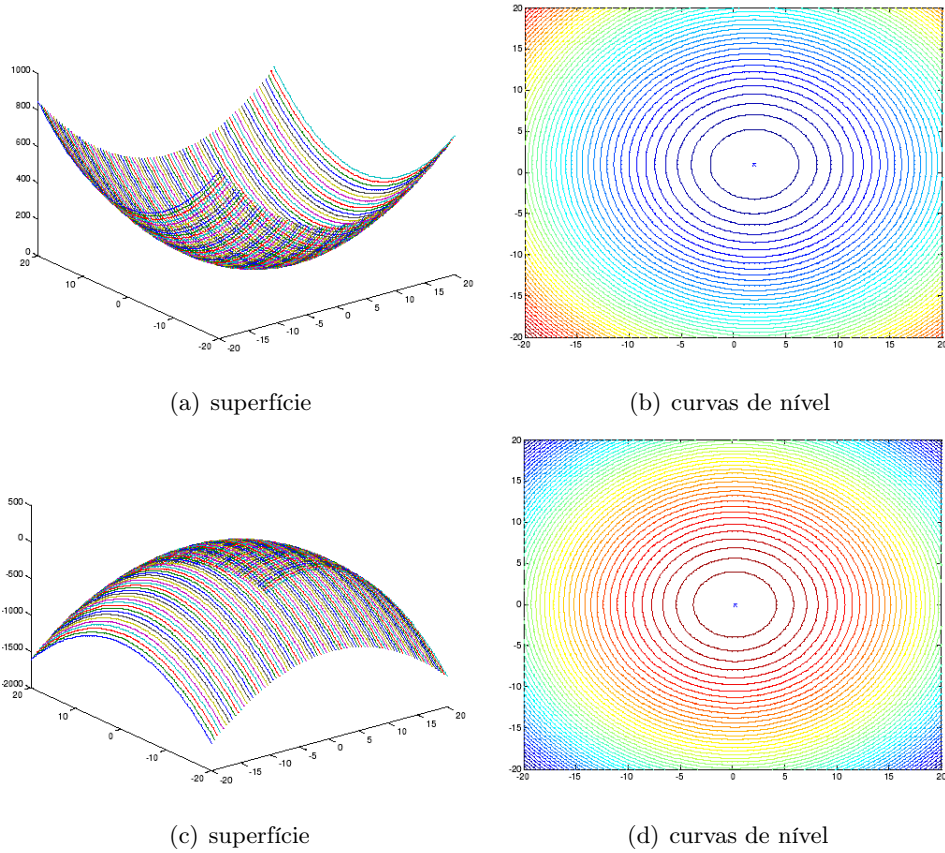


Figura 12.1: Problemas com um ótimo

12.1 Notação

A notação mais usada no âmbito da otimização tem a ver com a definição das primeiras e segundas derivadas da função objetivo.

Seja $f(x)$, $x \in \mathbb{R}^n$ a função objetivo e $x = (x_1, x_2, \dots, x_n)^T$ um vetor com n componentes.

Vetor gradiente de $f(x)$

O vetor gradiente, $\nabla f(x)$, da função $f(x)$, contém as primeiras derivadas parciais de $f(x)$ e é um vetor de dimensão n . Cada i componente do gradiente, $i = 1, \dots, n$, é dada por $\frac{\partial f}{\partial x_i}$.

Assim

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

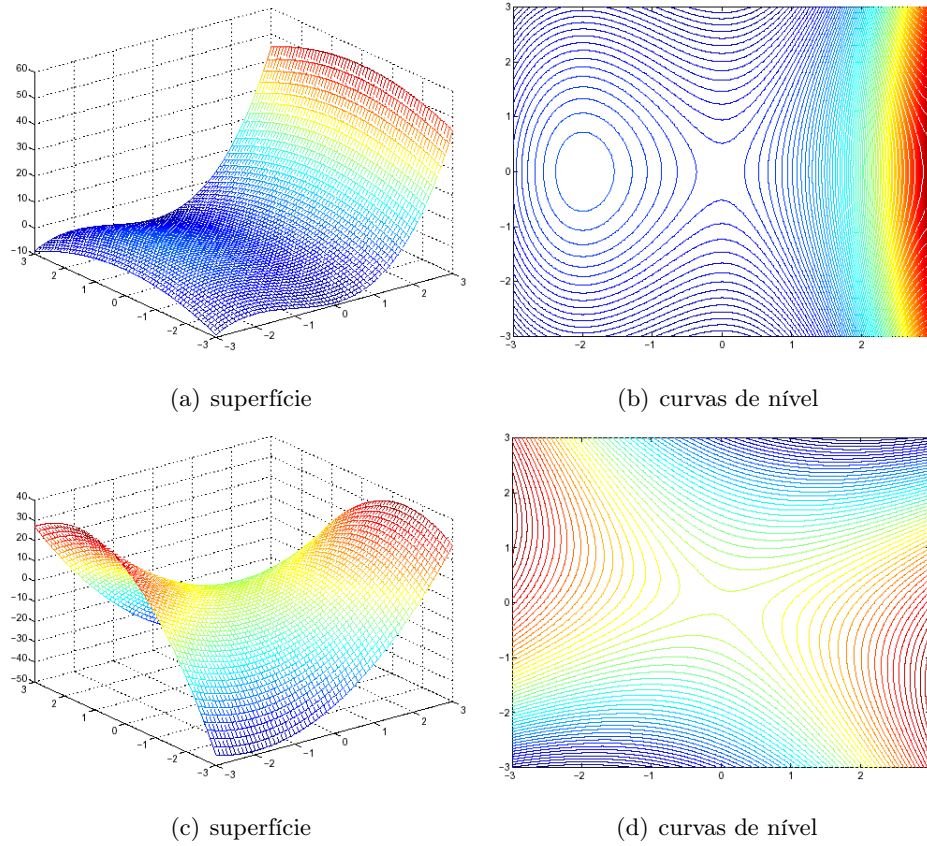


Figura 12.2: Problemas com um ponto sela

Matriz Hessiana de $f(x)$

A matriz Hessiana, $\nabla^2 f(x)$, da função $f(x)$, contém as segundas derivadas parciais de $f(x)$ (ou as primeiras derivadas parciais de cada uma das funções do vetor gradiente) e é uma matriz de dimensão $n \times n$. As ij componentes da Hessiana, $i = 1, \dots, n$, $j = \dots, n$ são dadas por $\frac{\partial^2 f}{\partial x_i \partial x_j}$, para $i \neq j$ e $\frac{\partial^2 f}{\partial x_i^2}$, para $i = j$. Assim

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

12.2 Condições de otimalidade

Para aplicar as condições de otimalidade de primeira e segunda ordem, assume-se $f(x)$ continuamente diferenciável até à segunda ordem.

Condição necessária e suficiente de primeira ordem

Se x^* é uma solução do problema (9.2), com $n > 1$, então $\nabla f(x^*) = 0$. Isto é, a equação $\nabla f(x^*) = 0$ define os pontos estacionários da função objectivo $f(x)$. Os pontos estacionários podem ser minimizantes (Figuras 12.1(a) e 12.1(a)), maximizantes (Figuras 12.1(c) e 12.1(d)) ou pontos sela (Exemplo 12.2).

Condição necessária de segunda ordem

Se x^* é uma solução do problema (9.2) para $n > 1$, que satisfaz a condição de primeira ordem, então a condição necessária para que x^* seja minimizante é que $\nabla^2 f(x^*)$ seja semi-definida positiva e a condição necessária para que x^* seja maximizante é que $\nabla^2 f(x^*)$ seja semi-definida negativa.

Condição suficiente de segunda ordem

Se x^* é solução do problema (9.2) e $\nabla f(x^*) = 0$, e $\nabla^2 f(x^*) \neq$ matriz nula:

- Se $\nabla^2 f(x^*)$ é definida positiva então x^* é minimizante local forte de $f(x)$;
- Se $\nabla^2 f(x^*)$ é definida negativa então x^* é maximizante local forte de $f(x)$;
- Se $\nabla^2 f(x^*)$ é semi-definida positiva então x^* é minimizante ou ponto sela de $f(x)$;
- Se $\nabla^2 f(x^*)$ é semi-definida negativa então x^* é maximizante ou ponto sela de $f(x)$;
- Se $\nabla^2 f(x^*)$ é indefinida então x^* é ponto sela de $f(x)$.

Uma matriz diz-se

- definida positiva se todos os determinantes das submatrizes principais dessa matriz são positivos,

- definida negativa se os determinantes das submatrizes principais dessa matriz são alternadamente negativos e positivos, sendo o determinante da primeira submatriz negativo,
- semi-definida positiva se os determinantes das submatrizes principais dessa matriz são positivos ou iguais a zero,
- semi-definida negativa se os determinantes das submatrizes principais dessa matriz são alternadamente negativos e positivos, sendo o determinante da primeira submatriz negativo, ou iguais a zero,
- indefinida nos restantes casos.

Exemplo 12.3 Dada a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x_1, x_2) = x_1^2(1 - x_1)^2 + x_1x_2$$

verifique se tem maximizantes, minimizantes e/ou pontos sela.

Resolução:

$$f(x_1, x_2) = x_1^2(1 - x_1)^2 + x_1x_2$$

$$\nabla f(x) = \begin{pmatrix} 2x_1(1 - x_1)^2 - 2x_1^2(1 - x_1) + x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 2x_1(1 - x_1)(1 - x_1 - x_1) + x_2 \\ x_1 \end{pmatrix} =$$

$$\begin{pmatrix} (2x_1 - 2x_1^2)(1 - 2x_1) + x_2 \\ x_1 \end{pmatrix}$$

$$\nabla f(x) = 0 \Leftrightarrow \begin{cases} x_2 = 0 \\ x_1 = 0 \end{cases}$$

$$\nabla^2 f(x) = \begin{pmatrix} (2 - 4x_1)(1 - 2x_1) - 2(2x_1 - 2x_1^2) & 1 \\ 1 & 0 \end{pmatrix}$$

$$x^* = (0, 0)$$

$$\nabla^2 f(x^*) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\det(2) = 2 > 0 \quad \det(\nabla^2 f(x^*)) = -1 < 0 \text{ logo a matriz é indefinida} \Rightarrow x^* \text{ é ponto sela.}$$

12.3 Exercícios

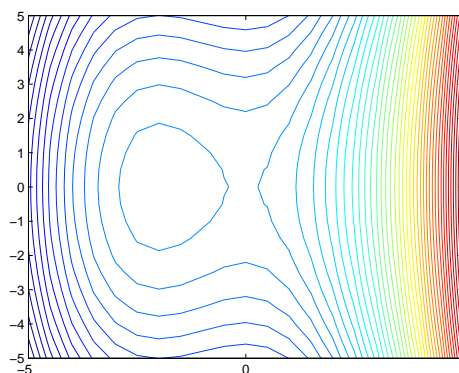
1. Dada a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x_1, x_2) = x_1^2(1 - x_1)^2 + x_1x_2$$

verifique se tem maximizantes, minimizantes e/ou pontos sela.

2. Considere a função

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



Mostre que a função dada tem um máximo local em $(-2, 0)$, tem um ponto sela em $(0, 0)$; e não tem mínimos.

3. Dada a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por

$$f(x_1, x_2, x_3) = 5x_1^2 + 2x_2^2 + x_3^4 - 32x_3 + 6x_1x_2 + 5x_2$$

verifique que ela tem apenas um ponto estacionário. Classifique-o.

4. Mostre que qualquer ponto da linha $x_2 - 2x_1 = 0$ é um mínimo de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x_1, x_2) = 4x_1^2 - 4x_1x_2 + x_2^2.$$

12.4 Soluções

1. $x^* = (0, 0)$ é ponto sela.
2. $x^* = (-2, 0)$ é maximizante; $x = (0, 0)^*$ é ponto sela.
3. $x^* = (7.5, -12.5, 2)$ é minimizante.
4. Os pontos da linha são minimizantes ou pontos sela.

Capítulo 13

Métodos do gradiente

Os métodos do gradiente, tal como o nome indica, usam informação, para além dos valores da função objetivo, das primeiras e/ou segundas derivadas da função (vetor gradiente e matriz Hessiana). Por este motivo, só podem ser usados na resolução de problemas diferenciáveis. Têm a vantagem de convergir mais rapidamente que os métodos de procura direta, que não usam informação sobre as derivadas. Têm, no entanto, a desvantagem do esforço computacional exigido para o cálculo das derivadas.

Estes métodos, que são iterativos, geram uma sucessão de aproximações x^k à solução

$$x^{k+1} = x^k + \alpha^k d^k,$$

em que d^k é a direção de procura ou passo e α^k é o comprimento do passo. d^k é um vetor e α^k é um escalar. A equação iterativa para o cálculo da direção de procura é diferente de acordo com o método que se está a usar.

O Algoritmo 13.2 descreve um método do gradiente.

13.1 Técnicas de globalização

Os métodos do gradiente, quando convergem, convergem para um ponto estacionário, ou seja, um ponto que anula o vetor gradiente de f : $\nabla f(x^*) = 0$. Uma vez que estes métodos exibem convergência local, deve implementar-se uma técnica de globalização de forma a garantir que o método converge, qualquer que seja a aproximação inicial x^1 . Significa que x^1 pode estar fora da região de convergência do método. Além disso, garante-se que o ponto estacionário para o qual o método converge é um minimizante.

Há várias técnicas de globalização, entre as quais:

- Procura unidimensional (*line search*)
 - procura unidimensional exata
 - procura unidimensional aproximada
- Região de confiança (*trust region*)
- Filtro

13.2 Procura unidimensional aproximada - critério de Armijo

Na procura unidimensional aproximada pretende-se calcular α^k , o comprimento do passo, dados x^k e d^k , que origina uma redução significativa do valor de f na nova aproximação. Para garantir essa redução pode usar-se a condição de Armijo (13.1).

$$f(x^k + \alpha^k d^k) \leq f(x^k) + \mu \alpha^k \nabla f(x^k)^T d^k, \quad (13.1)$$

com $0 < \mu < \frac{1}{2}$.

Se a direção d^k usada for descendente para f , ou seja, $\nabla f(x^k)^T d^k < 0$, existe um valor de $\alpha^k \in (0, 1]$ que verifica esta condição.

Descreve-se no Algoritmo 13.3 o cálculo α^k usando o critério de Armijo.

13.3 Método de Newton

O método de Newton é um método do gradiente que usa informação sobre as primeiras derivadas da função f – o vetor gradiente – e as segundas derivadas da função f – a matriz Hessiana. Trata-se de um método iterativo, à semelhança dos outros métodos do gradiente, e baseia-se, em cada iteração, numa aproximação local de $f(x)$ a uma função quadrática. Derivando esta função quadrática em ordem a d e igualando o vetor gradiente resultante a zero, define-se a condição de primeira ordem para o mínimo da quadrática, isto é, $\nabla q(d) = 0$. Assim, obtém-se

$$\begin{aligned} \nabla f(x^k) + \nabla^2 f(x^k) d_N^k &= 0 \Leftrightarrow \\ \nabla^2 f(x^k) d_N^k &= -\nabla f(x^k). \end{aligned} \quad (13.2)$$

Algoritmo 13.2 Método do gradiente

ler: x^1 e ε $k \leftarrow 0$ **repetir** $k \leftarrow k + 1$ calcular d^k calcular α^k $x^{k+1} \leftarrow x^k + \alpha^k d^k$ **até** $\|\nabla f(x^{k+1})\|_2 \leq \varepsilon$ $x^* \leftarrow x^{k+1}$ $f(x^*) \leftarrow f(x^{k+1})$

Algoritmo 13.3 Critério de Armijo

ler: x^k , d^k , $f(x^k)$, $\nabla f(x^k)$ e μ $\alpha \leftarrow 2$ **repetir** $\alpha \leftarrow 0.5 \times \alpha$ $x^{\text{aux}} \leftarrow x^k + \alpha d^k$ **até** $f(x^{\text{aux}}) \leq f(x^k) + \mu \alpha \nabla f(x^k)^T d^k$ $\alpha^k \leftarrow \alpha$

A solução do sistema linear (13.2), d^k , é a direção de procura. Deve ser usado um método direto e estável, por exemplo, o método de eliminação de Gauss com pivotagem parcial (EGPP), para resolver este sistema.

A nova aproximação $x^k + d_N^k$ não é necessariamente o minimizante de $f(x)$ e por isso o processo deve ser repetido.

13.3.1 Propriedades do método de Newton

O método de Newton tem convergência local, isto é, a convergência para a solução só é garantida se a aproximação inicial, x^1 , estiver na vizinhança da solução. Exibe convergência quadrática, ou seja,

$$\|x^{k+1} - x^*\| \leq \gamma \|x^k - x^*\|^2, \gamma > 0.$$

O método de Newton possui a propriedade da terminação quadrática, isto é, se $f(x)$, com $x \in \mathbb{R}^n$, for uma função quadrática e convexa, o método de Newton necessita no máximo de n iterações para encontrar a solução exata do problema.

Apesar do método de Newton ter boas propriedades de convergência, tem algumas limitações e desvantagens, que a seguir se descrevem.

13.3.2 Limitações do método de Newton

Em qualquer dos casos a seguir descritos, não é possível obter-se uma direção descendente através do método de Newton.

Direção Newton ascendente

A direção d_N^k , solução do sistema Newton (13.2), pode ser ascendente para f em x^k (Figura 13.1), ou seja,

$$\nabla f(x^k)^T d_N^k > 0.$$

Direção Newton ortogonal ao gradiente

A direção Newton d_N^k , solução do sistema Newton (13.2), pode ser ortogonal ao gradiente em x^k (Figura 13.2), ou seja,

$$\nabla f(x^k)^T d_N^k = 0.$$

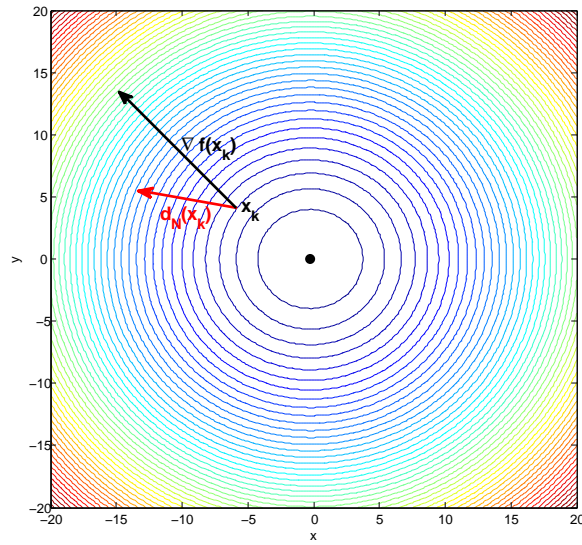


Figura 13.1: Direção ascendente.

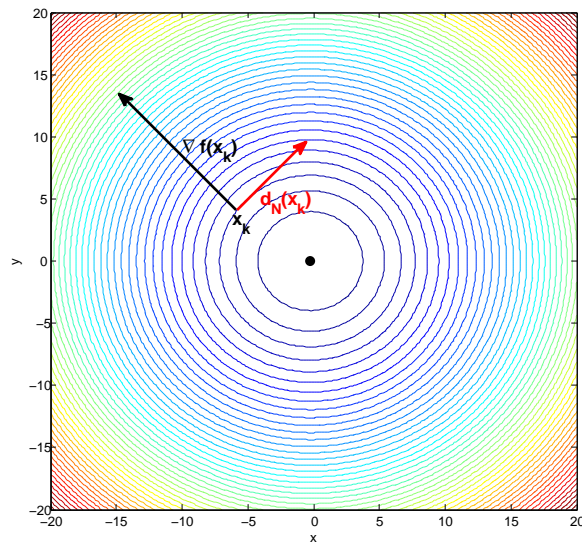


Figura 13.2: Direção ortogonal.

Direção Newton muito grande

A direção d_N^k , solução do sistema Newton (13.2), ainda que seja descendente, isto é, $\nabla f(x^k)^T d_N^k < 0$, pode ser muito grande (Figura 13.3) e por isso não se verifica

$$f(x^k + d_N^k) < f(x^k).$$

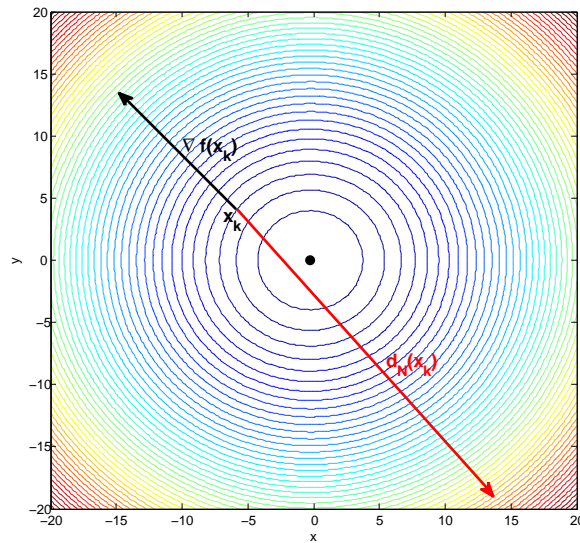


Figura 13.3: Direção muito grande.

Matriz Hessiana em x^k singular

A matriz dos coeficientes do sistema Newton (13.2) $\nabla^2 f(x^k)$ pode ser singular, o que significa que o sistema Newton não tem solução ou tem uma infinidade de soluções, ou seja,

$$\nexists d_N^k \text{ única.}$$

13.3.3 Desvantagens do método de Newton

A maior desvantagem do método de Newton é o facto de exigir o cálculo de segundas derivadas, por vezes muito complexo e dispendioso. Quando a expressão de $f(x)$ é complicada, estas tornam-se difíceis de calcular. Por outro lado, quando a dimensão do problema, n , for muito grande, o esforço de cálculo também será grande. Há ainda a questão de a convergência ser local e, por esse facto, estar condicionada à escolha do valor inicial.

Para estas limitações serem ultrapassadas, podem implementar-se algumas soluções, como a seguir se descreve, dando origem ao método de segurança de Newton.

13.4 Método de segurança de Newton

Para que o método de Newton resulte, é necessário ultrapassar as suas limitações de modo a garantir que a direção obtida é descendente para a função f . Para isso, implementa-se o método de segurança de Newton que consiste em qualquer iteração k :

- quando $\nabla^2 f(x^k)$ é singular, usar-se $d_{SN}^k = -\nabla f(x^k)$, em que $-\nabla f(x^k)$ é a direção de descida máxima e é descendente para f .
- quando d_N^k é ortogonal ao gradiente, isto é, $|\nabla f(x^k)^T d_N^k| \leq \eta$ (com $\eta > 0$ e $\eta \approx 0$), usar-se $d_{SN}^k = -\nabla f(x^k)$.
- quando d_N^k é ascendente, isto é, $\nabla f(x^k)^T d_N^k > \eta$, usar-se $d_{SN}^k = -d_N^k$.
- quando d_N^k é descendente, isto é, $\nabla f(x^k)^T d_N^k < \eta$, usar-se $d_{SN}^k = d_N^k$.

A direção d_{SN}^k assim obtida, é descendente para todo o k .

No Algoritmo 13.4 encontra-se descrito o método de segurança de Newton.

Algoritmo 13.4 Método de segurança de Newton

ler: x^k e η

resolver o sistema linear Newton $\nabla^2 f(x^k)d_N^k = -\nabla f(x^k)$ por EGPP

se $\exists d_N^k$ (o sistema linear tem solução única) **então**

se $|\nabla f(x^k)^T d_N^k| \leq \eta$ **então**

$$d_{SN}^k = -\nabla f(x^k)$$

senão

se $\nabla f(x^k)^T d_N^k > \eta$ **então**

$$d_{SN}^k = -d_N^k$$

senão

$$d_{SN}^k = d_N^k$$

fim se

fim se

senão

$$d_{SN}^k = -\nabla f(x^k)$$

fim se

O método de segurança de Newton mantém as mesmas propriedades de convergência que o método de Newton.

Exemplo 13.4 Considere a função

$$f(x_1, x_2) = -\sin(x_1 - 1) - x_2^4.$$

Implemente, no máximo, duas iterações do método de segurança de Newton para determinar o máximo da função $f(x_1, x_2)$. Considere $\eta = 10^{-6}$, $\mu = 10^{-6}$, $\varepsilon = 1$ e $x^{(1)} = (1, 1)^T$.

Resolução:

$$\max \bar{f}(x_1, x_2) = -\sin(x_1 - 1) - x_2^4$$

$$\min f(x_1, x_2) = \sin(x_1 - 1) + x_2^4$$

$$\nabla f(x) = \begin{pmatrix} \cos(x_1 - 1) \\ 4x_2^3 \end{pmatrix}$$

$$\nabla^2 f(x) = \begin{pmatrix} -\sin(x_1 - 1) & 0 \\ 0 & 12x_2^2 \end{pmatrix}$$

Iniciar o algoritmo de Segurança de Newton: $x^1 = (1, 1)$, $\eta = 10^{-6}$, $\mu = 10^{-6}$, $\varepsilon = 1$

• **1ª iteração**

$$x^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \nabla f(x^1) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \quad \nabla^2 f(x^1) = \begin{pmatrix} 0 & 0 \\ 0 & 12 \end{pmatrix}$$

Cálculo da direção d_N^1

$$\begin{pmatrix} 0 & 0 & | & -1 \\ 0 & 12 & | & -4 \end{pmatrix} \rightarrow \text{sistema impossível} \Rightarrow d_{SN}^1 = -\nabla f(x^1) = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$$

Cálculo de α

$$\alpha = 1$$

$$x^{\text{aux}} = x^1 + \alpha d_{SN}^1 = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$$

$$\begin{cases} f(x^1) = 1 \\ f(x^{\text{aux}}) = 80.158562 \end{cases} \quad \uparrow$$

$$\alpha = 0.5 \times 1 = 0.5$$

$$x^{\text{aux}} = x^1 + \alpha d_{SN}^1 = \begin{pmatrix} 0.5 \\ -1 \end{pmatrix}$$

$$\begin{cases} f(x^1) = 1 \\ f(x^{\text{aux}}) = 0.520574 \quad \downarrow \end{cases}$$

Critério de Armijo

$$f(x^{\text{aux}}) \leq f(x^1) + \mu \alpha \nabla f(x^1)^T d_{SN}^1 \Leftrightarrow 0.520574 \leq 1 + 10^{-6} \times 0.5 \times (-17)$$

$\Leftrightarrow 0.520574 \leq 1.0000085$ (verdadeiro) logo a descida é significativa.

$$x^2 = \begin{pmatrix} 0.5 \\ -1 \end{pmatrix}$$

- Critério de Paragem

$$\|\nabla f(x^2)\|_2 = \left\| \begin{pmatrix} 0.877583 \\ -4 \end{pmatrix} \right\|_2 = 4.095138 \leq \varepsilon \quad (\text{falso})$$

- 2ª iteração

$$x^2 = \begin{pmatrix} 0.5 \\ -1 \end{pmatrix} \quad \nabla f(x^2) = \begin{pmatrix} 0.877583 \\ -4 \end{pmatrix} \quad \nabla^2 f(x^2) = \begin{pmatrix} 0.479426 & 0 \\ 0 & 12 \end{pmatrix}$$

Cálculo da direção d_N^2

$$\left(\begin{array}{cc|c} 0.479426 & 0 & -0.877583 \\ 0 & 12 & 4 \end{array} \right) \rightarrow d_N^2 = \begin{pmatrix} -1.830487 \\ 0.333333 \end{pmatrix}$$

O sistema tem solução única.

$$\nabla f(x^2)^T d_N^2 = \begin{pmatrix} 0.877583 & -4 \end{pmatrix} \begin{pmatrix} -1.830487 \\ 0.333333 \end{pmatrix} = -2.939736$$

$|\nabla f(x^2)^T d_N^2| = 2.939736 > 10^{-6}$, logo d_N^2 não é ortogonal ao gradiente.

$\nabla f(x^2)^T d_N^2 = -2.939736 > 10^{-6}$, logo d_N^2 não é ascendente.

$$d_{SN}^2 = d_N^2 = \begin{pmatrix} -1.830487 \\ 0.333333 \end{pmatrix}$$

Cálculo de α

$$\alpha = 1$$

$$x^{\text{aux}} = x^2 + \alpha d_{SN}^2 = \begin{pmatrix} -1.330487 \\ -0.666667 \end{pmatrix}$$

$$\begin{cases} f(x^2) = 0.520574 \\ f(x^{\text{aux}}) = -0.527518 \end{cases} \quad \downarrow$$

Critério de Armijo

$$f(x^{\text{aux}}) \leq f(x^2) + \mu \alpha \nabla f(x^2)^T d_{SN}^2 \Leftrightarrow -0.527518 \leq 0.520574 + 10^{-6} \times 1 \times (-2.939736)$$

(verdadeiro), logo a descida é significativa.

$$x^3 = \begin{pmatrix} -1.330487 \\ -0.666667 \end{pmatrix}$$

- Critério de Paragem

$$\|\nabla f(x^3)\|_2 = \left\| \begin{pmatrix} -0.688697 \\ -1.185187 \end{pmatrix} \right\|_2 = 1.676108 \leq \varepsilon \quad (\text{falso})$$

Como o número máximo de iterações é dois,

$$x_{\max} \approx \begin{pmatrix} -1.330487 \\ -0.666667 \end{pmatrix} \text{ e } f_{\max} \approx 0.527518$$

13.5 Método quasi-Newton

Como foi referido anteriormente, uma das grandes desvantagens do método de Newton é o facto de exigir o cálculo das segundas derivadas de f (13.2). Para se evitar o cálculo das segundas derivadas, pode usar-se uma aproximação à matriz Hessiana,

$$B^k \approx \nabla^2 f(x^k)$$

e assim a direção de procura passa a ser calculada pela resolução do sistema linear por EGPP

$$B^k d_{QN}^k = -\nabla f(x^k).$$

No entanto, o sistema Newton (13.2) pode ser escrito de forma equivalente

$$d_N^k = - \left(\nabla^2 f(x^k) \right)^{-1} \nabla f(x^k),$$

por isso pode usar-se em alternativa, em cada iteração k , uma aproximação à inversa da Hessiana

$$H^k \approx \left(\nabla^2 f(x^k) \right)^{-1}$$

e calcula-se a direção de procura pelo produto desta matriz pelo vetor gradiente, que é um cálculo mais simples que a resolução de um sistema linear. Assim,

$$d_{QN}^k = -H^k \nabla f(x^k). \quad (13.3)$$

Usando este procedimento, evita-se não só o cálculo das segundas derivadas de f para formar a matriz Hessiana, como também a resolução de um sistema linear em cada iteração, que é substituído pelo produto de uma matriz por um vetor.

13.5.1 Características da matriz H

A matriz H deve aproximar o melhor possível a inversa de $\nabla^2 f(x^k)$, isto é, deve verificar-se a condição secante

$$H^k y^{k-1} = s^{k-1},$$

em que y representa a variação no gradiente da iteração $k-1$ para a iteração k e é dado por

$$y^{k-1} = \nabla f(x^k) - \nabla f(x^{k-1}).$$

Por sua vez, s representa a variação verificada em x da iteração $k-1$ para a iteração k e é dado por

$$s^{k-1} = x^k - x^{k-1} = \alpha^{k-1} d_{QN}^{k-1}.$$

A matriz H deve ainda ser, preferencialmente, simétrica, pois $\nabla^2 f(x^k)^{-1}$ também é simétrica e definida positiva e a direção (13.3) é descendente para f em x^k .

As matrizes H^k são geradas através de fórmulas de atualização do tipo

$$H^k = H^{k-1} + E^{k-1}$$

e devem manter-se simétricas e definidas positivas. Para que isso aconteça, a matriz inicial H^1 deve ser também simétrica e definida positiva. A matriz mais usual para iniciar este processo é a matriz identidade, isto é,

$$H^1 = I.$$

A matriz identidade não é necessariamente uma boa aproximação a $\nabla^2 f(x^k)^{-1}$, mas as fórmulas de atualização rapidamente ultrapassam esse problema e melhoram as aproximações.

Existem várias fórmulas de atualização para as matrizes H , no entanto, nem todas conservam a matriz simétrica e definida positiva. As fórmulas (13.4) e (13.5) conservam as aproximações H simétricas e definidas positivas. A condição $y^{k-1T} s^{k-1} > 0$ é necessária e suficiente para que as matrizes se conservem simétricas e definidas positivas e estas fórmulas verificam essa condição.

Fórmula de atualização de Davidon, Fletcher e Powell (DFP)

$$H^k = H^{k-1} - \frac{H^{k-1} y^{k-1} y^{k-1T} H^{k-1}}{y^{k-1T} H^{k-1} y^{k-1}} + \frac{s^{k-1} s^{k-1T}}{s^{k-1T} y^{k-1}} \quad (13.4)$$

Fórmula de atualização de Broyden, Fletcher, Goldfarb e Shanno (BFGS)

$$H^k = \left(I - \frac{s^{k-1} y^{k-1T}}{s^{k-1T} y^{k-1}} \right) H^{k-1} \left(I - \frac{y^{k-1} s^{k-1T}}{s^{k-1T} y^{k-1}} \right) + \frac{s^{k-1} s^{k-1T}}{s^{k-1T} y^{k-1}} \quad (13.5)$$

13.5.2 Propriedades do método quasi-Newton

O método quasi-Newton, tal como o método de Newton, tem convergência local, ou seja, a convergência para a solução só é garantida quando a aproximação inicial x^1 estiver na vizinhança da solução. No entanto, a rapidez de convergência é inferior à do método de Newton, sendo que é superlinear. Significa que se verifica

$$\|x^{k+1} - x^*\| \leq \gamma_k \|x^k - x^*\|$$

com a sucessão $\{\gamma_k\} \rightarrow 0$ quando $k \rightarrow \infty$.

À semelhança do método de Newton, o método quasi-Newton tem a propriedade da terminação quadrática, isto é, o mínimo de uma função quadrática $q(x)$, $x \in \mathbb{R}^n$ obtém-se em n ou menos que n iterações.

Devido aos erros de arredondamento que se cometem nos cálculos ao longo das iterações, a matriz H^k pode deixar de ser definida positiva e assim a direção d_{QN}^k deixa de ser descendente

para f em x^k . Neste caso deve fazer-se $H^k = I$, que é simétrica e definida positiva. Quando isto acontece, $d_{QN}^k = -\nabla f(x^k)$ – direção de descida máxima.

Descreve-se no Algoritmo 13.5 o método quasi-Newton.

Algoritmo 13.5 Método quasi-Newton

ler: x^k

se $k = 1$ **então**

$$H^k = I$$

senão

$$s^{k-1} \leftarrow x^k - x^{k-1}$$

$$y^{k-1} \leftarrow \nabla f(x^k) - \nabla f(x^{k-1})$$

atualizar H^k por (13.4) ou (13.5)

fim se

$$d_{QN}^k \leftarrow -H^k \nabla f(x^k)$$

se $\nabla f(x^k)^T d_{QN}^k \geq 0$ **então**

$$d_{QN}^k \leftarrow -\nabla f(x^k)$$

fim se

Exemplo 13.5 O lucro, em milhares de euros, da colocação de um sistema elétrico é dado por

$$\mathcal{L}(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

em que x_1 e x_2 designam, respectivamente, o custo da mão de obra e do material. Calcule o lucro máximo usando o método quasi-Newton baseado na fórmula DFP, considerando na paragem do processo iterativo $\varepsilon = 0.0001$. Tome a seguinte aproximação inicial $(0, 0)$. No critério de Armijo use $\mu = 0.001$.

Resolução:

$$\max \mathcal{L}(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

$$\min -\mathcal{L}(x_1, x_2) = f(x_1, x_2) = -20x_1 - 26x_2 - 4x_1x_2 + 4x_1^2 + 3x_2^2$$

$$\nabla f(x) = \begin{pmatrix} -20 - 4x_2 + 8x_1 \\ -26 - 4x_1 + 6x_2 \end{pmatrix}$$

Iniciar o algoritmo de quasi-Newton: $x^1 = (0, 0), \mu = 0.001, \varepsilon = 0.001$

• **1ª iteração**

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \nabla f(x^1) = \begin{pmatrix} -20 \\ -26 \end{pmatrix}$$

$$H^1 = I$$

Cálculo da direção d_{QN}^1

$$d_{QN}^1 = -H^1 \nabla f(x^1) = \begin{pmatrix} 20 \\ 26 \end{pmatrix}$$

$$\nabla f(x^1)^T d_{QN}^1 = \begin{pmatrix} -20 & -26 \end{pmatrix} \begin{pmatrix} 20 \\ 26 \end{pmatrix} = -1076 < 0, \text{ logo } d_{QN}^1 \text{ é descendente.}$$

Cálculo de α

$$\alpha = 1$$

$$x^{\text{aux}} = x^1 + \alpha d_{QN}^1 = \begin{pmatrix} 20 \\ 26 \end{pmatrix}$$

$$\begin{cases} f(x^1) = 0 \\ f(x^{\text{aux}}) = 472 \quad \uparrow \end{cases}$$

$$\alpha = 0.5$$

$$x^{\text{aux}} = x^1 + \alpha d_{QN}^1 = \begin{pmatrix} 10 \\ 13 \end{pmatrix}$$

$$\begin{cases} f(x^1) = 0 \\ f(x^{\text{aux}}) = -151 \quad \downarrow \end{cases}$$

Critério de Armijo

$$f(x^{\text{aux}}) \leq f(x^1) + \mu \alpha \nabla f(x^1)^T d_{QN}^1 \Leftrightarrow -151 \leq 0 + 0.001 \times 0.5 \times (-1076)$$

$$\Leftrightarrow -151 \leq -0.538 \text{ (verdadeiro) logo a descida é significativa.}$$

$$x^2 = \begin{pmatrix} 10 \\ 13 \end{pmatrix}$$

- Critério de Paragem

$$\|\nabla f(x^2)\|_2 = \left\| \begin{pmatrix} 8 \\ 12 \end{pmatrix} \right\|_2 = 14.4222 \leq \varepsilon \quad (\text{falso})$$

- 2ª iteração

$$x^2 = \begin{pmatrix} 10 \\ 13 \end{pmatrix} \quad \nabla f(x^2) = \begin{pmatrix} 8 \\ 12 \end{pmatrix}$$

$$H^2 = H^1 - \frac{H^1 y^1 y^{1T} H^1}{y^{1T} H^1 y^1} + \frac{s^1 s^{1T}}{s^{1T} y^1}$$

$$s^1 = x^2 - x^1 = \begin{pmatrix} 10 \\ 13 \end{pmatrix}$$

$$y^1 = \nabla f(x^2) - \nabla f(x^1) = \begin{pmatrix} 28 \\ 38 \end{pmatrix}$$

$$y^1 y^{1T} = \begin{pmatrix} 784 & 1064 \\ 1064 & 1444 \end{pmatrix}$$

$$y^{1T} y^1 = 2228$$

$$s^1 s^{1T} = \begin{pmatrix} 100 & 130 \\ 130 & 169 \end{pmatrix}$$

$$s^{1T} y^1 = 774$$

$$H^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{\begin{pmatrix} 784 & 1064 \\ 1064 & 1444 \end{pmatrix}}{2228} + \frac{\begin{pmatrix} 100 & 130 \\ 130 & 169 \end{pmatrix}}{774} = \begin{pmatrix} 0.7773 & -0.3096 \\ -0.3096 & 0.5702 \end{pmatrix}$$

Cálculo da direção d_{QN}^2

$$d_{QN}^2 = -H^2 \nabla f(x^2) = \begin{pmatrix} -2.5032 \\ -4.3656 \end{pmatrix}$$

$\nabla f(x^2)^T d_{QN}^2 = -72.4128 < 0$, logo d_{QN}^2 é descendente.

Cálculo de α

$$\alpha = 1$$

$$x^{\text{aux}} = x^2 + \alpha d_{QN}^2 = \begin{pmatrix} 7.4968 \\ 8.6344 \end{pmatrix}$$

$$\begin{cases} f(x^2) = -151 \\ f(x^{\text{aux}}) = -184.8852 \end{cases} \quad \downarrow$$

Critério de Armijo

$$f(x^{\text{aux}}) \leq f(x^2) + \mu \alpha \nabla f(x^2)^T d_{QN}^2 \Leftrightarrow -184.8852 \leq -151 + 0.001 \times 1 \times (-72.4128)$$

$$\Leftrightarrow -184.8852 \leq -151.0724 \text{ (verdadeiro) logo a descida é significativa.}$$

$$x^3 = \begin{pmatrix} 7.4968 \\ 8.6344 \end{pmatrix}$$

- Critério de Paragem

$$\|\nabla f(x^3)\|_2 = \left\| \begin{pmatrix} 5.4368 \\ -4.1808 \end{pmatrix} \right\|_2 = 6.8584 \leq \varepsilon \quad (\text{falso})$$

- 3ª iteração

$$x^3 = \begin{pmatrix} 7.4968 \\ 8.6344 \end{pmatrix} \quad \nabla f(x^3) = \begin{pmatrix} 5.4368 \\ -4.1808 \end{pmatrix}$$

$$H^3 = H^2 - \frac{H^2 y^2 y^{2T} H^2}{y^{2T} H^2 y^2} + \frac{s^2 s^{2T}}{s^{2T} y^2}$$

$$s^2 = x^3 - x^2 = \begin{pmatrix} -2.5032 \\ -4.3656 \end{pmatrix}$$

$$y^2 = \nabla f(x^3) - \nabla f(x^2) = \begin{pmatrix} -2.5632 \\ -16.1808 \end{pmatrix}$$

$$H^2 y^2 = \begin{pmatrix} 3.0172 \\ -8.4327 \end{pmatrix}$$

$$y^{2T} H^2 = \begin{pmatrix} 3.0172 & -8.4327 \end{pmatrix}$$

$$H^2 y^2 y^{2T} H^2 = \begin{pmatrix} 9.1035 & -25.4431 \\ -25.4431 & 71.1104 \end{pmatrix}$$

$$y^{2T} H^2 y^2 = 128.7141$$

$$s^2 s^{2T} = \begin{pmatrix} 6.2660 & 10.9280 \\ 10.9280 & 19.0585 \end{pmatrix}$$

$$s^{2T} y^2 = 77.0551$$

$$H^3 = \begin{pmatrix} 0.7773 & -0.3096 \\ -0.3096 & 0.5702 \end{pmatrix} - \frac{\begin{pmatrix} 9.1035 & -25.4431 \\ -25.4431 & 71.1104 \end{pmatrix}}{128.7141} + \frac{\begin{pmatrix} 6.2660 & 10.9280 \\ 10.9280 & 19.0585 \end{pmatrix}}{77.0551} =$$

$$\begin{pmatrix} 0.7879 & 0.0299 \\ 0.0299 & 0.2651 \end{pmatrix}$$

Cálculo da direção d_{QN}^3

$$d_{QN}^3 = -H^3 \nabla f(x^3) = \begin{pmatrix} -4.1586 \\ 0.9458 \end{pmatrix}$$

$\nabla f(x^3)^T d_{QN}^3 = -26.5637 < 0$, logo d_{QN}^3 é descendente.

Cálculo de α

$$\alpha = 1$$

$$x^{\text{aux}} = x^3 + \alpha d_{QN}^3 = \begin{pmatrix} 3.3382 \\ 9.5802 \end{pmatrix}$$

$$\begin{cases} f(x^3) = -184.8852 \\ f(x^{\text{aux}}) = -123.8567 \quad \uparrow \end{cases}$$

$$\alpha = 0.5$$

$$x^{\text{aux}} = x^3 + \alpha d_{QN}^3 = \begin{pmatrix} 5.4175 \\ 9.1073 \end{pmatrix}$$

$$\begin{cases} f(x^3) = -184.8852 \\ f(x^{\text{aux}}) = -176.2690 \quad \uparrow \end{cases}$$

$$\alpha = 0.25$$

$$x^{\text{aux}} = x^3 + \alpha d_{QN}^3 = \begin{pmatrix} 6.4572 \\ 8.8708 \end{pmatrix}$$

$$\begin{cases} f(x^3) = -184.8852 \\ f(x^{\text{aux}}) = -186.0518 \quad \downarrow \end{cases}$$

Critério de Armijo

$$f(x^{\text{aux}}) \leq f(x^3) + \mu \alpha \nabla f(x^3)^T d_{QN}^3 \Leftrightarrow -186.0517 \leq -184.8852 + 0.001 \times 0.25 \times (-26.5637)$$

$$\Leftrightarrow -186.0517 \leq -184.8918 \text{ (verdadeiro) logo a descida é significativa.}$$

$$x^4 = \begin{pmatrix} 6.4572 \\ 8.8709 \end{pmatrix}$$

- Critério de Paragem

$$\|\nabla f(x^3)\|_2 = \left\| \begin{pmatrix} -3.8262 \\ 1.3965 \end{pmatrix} \right\|_2 = 4.0731 \leq \varepsilon \quad (\text{falso})$$

- 4ª iteração

$$x^4 = \begin{pmatrix} 6.4572 \\ 8.8709 \end{pmatrix} \quad \nabla f(x^4) = \begin{pmatrix} -3.8262 \\ 1.3965 \end{pmatrix}$$

$$H^4 = H^3 - \frac{H^3 y^3 y^{3T} H^3}{y^{3T} H^3 y^3} + \frac{s^3 s^{3T}}{s^{3T} y^3}$$

$$s^3 = x^4 - x^3 = \begin{pmatrix} -1.0397 \\ 0.2364 \end{pmatrix}$$

$$y^3 = \nabla f(x^4) - \nabla f(x^3) = \begin{pmatrix} -9.2630 \\ 5.5773 \end{pmatrix}$$

$$H^3 y^3 = \begin{pmatrix} -7.1316 \\ 1.2016 \end{pmatrix}$$

$$y^{3^T} H^3 = \begin{pmatrix} -7.1316 & 1.2016 \end{pmatrix}$$

$$H^3 y^3 y^{3^T} H^3 = \begin{pmatrix} 50.8597 & -8.5693 \\ -8.5693 & 1.4438 \end{pmatrix}$$

$$y^{3^T} H^3 y^3 = 72.7617$$

$$s^3 s^{3^T} = \begin{pmatrix} 1.0809 & -0.2458 \\ -0.2458 & 0.0559 \end{pmatrix}$$

$$s^{3^T} y^3 = 10.9490$$

$$H^4 = \begin{pmatrix} 0.7879 & 0.0299 \\ 0.0299 & 0.2651 \end{pmatrix} - \frac{\begin{pmatrix} 50.8597 & -8.5693 \\ -8.5693 & 1.4438 \end{pmatrix}}{72.7617} + \frac{\begin{pmatrix} 1.0809 & -0.2458 \\ -0.2458 & 0.0559 \end{pmatrix}}{10.9490} = \begin{pmatrix} 0.1876 & 0.1252 \\ 0.1252 & 0.2504 \end{pmatrix}$$

Cálculo da direção d_{QN}^4

$$d_{QN}^4 = -H^4 \nabla f(x^4) = \begin{pmatrix} 0.5430 \\ 0.1294 \end{pmatrix}$$

$\nabla f(x^4)^T d_{QN}^4 = -1.8968 < 0$, logo d_{QN}^4 é descendente.

Cálculo de α

$$\alpha = 1$$

$$x^{\text{aux}} = x^4 + \alpha d_{QN}^4 = \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix}$$

$$\begin{cases} f(x^4) = -186.0517 \\ f(x^{\text{aux}}) = -187.0000 \end{cases} \quad \downarrow$$

Critério de Armijo

$$f(x^{\text{aux}}) \leq f(x^4) + \mu \alpha \nabla f(x^4)^T d_{QN}^4 \Leftrightarrow -187.0000 \leq -186.0517 + 0.001 \times 1 \times (-1.8968)$$

$\Leftrightarrow -187.0000 \leq -186.0536$ (verdadeiro) logo a descida é significativa.

$$x^5 = \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix}$$

- Critério de Paragem

$$\|\nabla f(x^4)\|_2 = \left\| \begin{pmatrix} 0.0000 \\ 0.0008 \end{pmatrix} \right\|_2 = 0.0008 \leq \varepsilon \quad (\textit{verdadeiro})$$

$$x_{\max} \approx \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix} \text{ e } \mathcal{L}_{\max} \approx 187.$$

13.6 Exercícios

1. Considere a função

$$f(x_1, x_2) = -\sin(x_1 - 1) - x_2^4.$$

Implemente, no máximo, duas iterações do método de segurança de Newton para determinar o máximo da função $f(x_1, x_2)$. Considere $\eta = 10^{-6}$, $\mu = 10^{-6}$, $\varepsilon = 1$ e $x^{(1)} = (1, 1)^T$.

2. A soma de três números $(x_1, x_2$ e $x_3)$ positivos é igual a 40. Determine esses números de modo que a soma dos seus quadrados seja mínima.

Use a relação da soma para colocar x_3 em função das outras 2 variáveis. Formule o problema como um problema de otimização sem restrições.

A partir da aproximação inicial $(x_1, x_2)^{(1)} = (10, 10)$, use o método de Segurança de Newton (com $\eta = 0.00001$) para calcular esses números, considerando no critério de paragem $\varepsilon = 0.001$. Na condição de Armijo tome $\mu = 0.001$.

3. Uma empresa fabrica e comercializa dois tipos de computadores portáteis. O custo de fabrico de cada um deles decresce à medida que o número de unidades produzidas aumenta e é dado pelas seguintes relações empíricas:

$$c_1 = 5 + \frac{1500}{x_1} \quad c_2 = 7 + \frac{2500}{x_2},$$

em que x_1 e x_2 são o número de unidades de cada um dos portáteis produzidos. O preço de venda dos computadores é tanto menor quanto maior for o número de unidades produzidas, de acordo com as seguintes relações:

$$p_1 = 15 - 0.001x_1 \quad \text{e} \quad p_2 = 25 - 0.0015x_2.$$

- a) Formule o problema de otimização que consiste em determinar quantas unidades de cada computador a firma deve produzir de modo a maximizar os lucros.
- b) Resolva o problema usando o método de Segurança de Newton (com $\eta = 0.00001$). Considere a seguinte aproximação inicial $(x_1, x_2)^{(1)} = (20, 30)$ e $\varepsilon = 0.001$. Na condição de Armijo tome $\mu = 0.001$.

- c) Com base na aproximação calculada na alínea anterior ao número de computadores produzidos, a empresa terá lucro?
4. Três estações elétricas vão fornecer energia a uma certa região da forma mais económica possível. Os custos individuais de operação de cada uma das estações são dados por

$$\begin{aligned}f_1 &= 0.1 + 0.25x \\f_2 &= 0.08 + 0.12y + 0.00125y^2 \\f_3 &= 0.05 + 0.09z + 0.001z^2 + 0.0001z^3\end{aligned}$$

em que x , y e z são as energias fornecidas pelas três estações (em MWatt). Determine os valores de x , y e z que minimizam o custo total, se a energia a ser fornecida for de 100 MWatt, recorrendo ao método de segurança de Newton.

Como valores iniciais use $(x, y)^{(1)} = (30, 50)$, no critério de paragem considere $\varepsilon = 0.05$ e tome $\eta = 0.0001$. Como estratégia de procura unidimensional utilize o critério de Armijo com $\mu = 0.01$. Use a relação relacionada com a energia a fornecer para eliminar uma das variáveis, por exemplo, $x = 100 - y - z$.

5. Numa situação monopolista, o rendimento de uma empresa face à venda de um produto ou serviço depende do nível de produção z . O rendimento é uma função crescente de z mas tende em direção a uma assíntota assim que o mercado fica saturado.

Considere a seguinte função rendimento

$$R(z) = z^2/(1 + z^2)$$

que depende da produção z dada por $z = x_1^{1/2}x_2^{1/2}$, em que x_1 representa o capital e x_2 o trabalho.

Supondo que a função lucro é dada por

$$\pi(x_1, x_2) = R(z) - 0.04x_1 - 0.06x_2$$

calcule o lucro máximo que a empresa pode ter. Use o método quasi-Newton (com fórmula BFGS). Como aproximação inicial considere o ponto $(2, 1)$. Use na paragem do processo iterativo $\varepsilon = 0.1$. No critério de Armijo use $\mu = 0.001$.

6. Suponha que pretendia representar um número A positivo na forma de um produto de quatro fatores positivos x_1, x_2, x_3 e x_4 . Para $A = 2401$, determine esses fatores de tal forma que a sua soma seja a menor possível.

Formule o problema como um problema de otimização sem restrições em função das 3 variáveis x_1, x_2 e x_3 .

A partir da aproximação inicial $(x_1, x_2, x_3)^{(1)} = (6, 7, 5)$, use o método quasi-Newton (com fórmula DFP), para calcular esses fatores. Na paragem do processo iterativo use $\varepsilon = 0.1$. No critério de Armijo use $\mu = 0.001$.

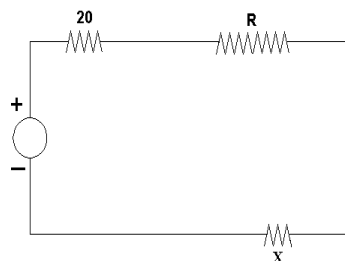
7. O lucro, em milhares de euros, da colocação de um sistema elétrico é dado por

$$\mathcal{L}(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

em que x_1 e x_2 designam, respectivamente, o custo da mão de obra e do material. Calcule o lucro máximo usando o método quasi-Newton baseado na fórmula DFP, considerando na paragem do processo iterativo $\varepsilon = 0.0001$. Tome a seguinte aproximação inicial $(0, 0)$. No critério de Armijo use $\mu = 0.001$.

8. Considere um circuito elétrico em que existem duas resistências variáveis, R e X . O valor médio da energia do circuito é dado por

$$P = \frac{10^4 R}{(R + 20)^2 + X^2}.$$



Determine os valores de R e X para os quais se obtém uma energia de saída máxima. Use o método quasi-Newton (fórmula DFP) e os valores iniciais $(R, X)^{(1)} = (10, 5)$. Considere $\mu = 0.001$ e $\varepsilon = 0.5$.

13.7 Soluções

1. 2 iterações; $x_{\max} \approx \begin{pmatrix} -1.3305 \\ -0.6667 \end{pmatrix}$; $f_{\max} = 0.5275$.
2. a) $\min x_1^2 + x_2^2 + (40 - x_1 - x_2)^2$.
 b) 1 iteração; $x_1 = 13.3333$; $x_2 = 13.3333$; $x_3 = 13.3333$; $f_{\min} = 533.3333$.
3. a) $\min 0.001x_1^2 + 0.0015x_2^2 - 10x_1 - 18x_2 + 4000$.
 b) 1 iteração; $x_{\max} = \begin{pmatrix} 5000 \\ 6000 \end{pmatrix}$; $f_{\max} = 75000$.
 c) Sim, o lucro é positivo.
4. a) $\min 25.23 - 0.13y + 0.00125y^2 - 0.16z + 0.001z^2 + 0.0001z^3$.
 b) 2 iterações; $(x, y, z)_{\min} \approx \begin{pmatrix} 26.8794 \\ 52 \\ 21.1206 \end{pmatrix}$; $f_{\min} \approx 19.8589$.
5. 2 iterações; $x_{\max} \approx \begin{pmatrix} 2.3111 \\ 1.6365 \end{pmatrix}$; $\pi_{\max} \approx 0.6003$.
6. 3 iterações; $x_1 \approx 7.0417$; $x_2 \approx 7.4110$; $x_3 \approx 6.7836$; $x_4 \approx 6.7823$; soma máxima ≈ 28.0186 .
7. $x_{\max} \approx \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix}$; $\mathcal{L} \approx 187$.
8. 3 iterações; $R_{\max} \approx 19.0402$; $X_{\max} \approx 1.1263$; $P_{\max} \approx 124.8206$.

13.8 MATLAB

13.8.1 fminunc

A função `fminunc` encontra o mínimo de uma função multidimensional usando métodos do gradiente. A sua sintaxe é

```
[xmin,fmin,exitflag,output] = fminunc('func',x0,options)
```

em que os parâmetros de saída definem o minimizante em `xmin`, o valor mínimo da função `fmin`, a `exitflag` define a forma como parou o algoritmo (a desejável é 1 - significa que o processo convergiu para um ponto estacionário) e na estrutura `output` encontra-se informação sobre o processo iterativo (número de iterações, número de cálculos da função, algoritmo usado...). Os parâmetros de entrada a fornecer são a função que se pretende minimizar (caso o problema seja de maximização tem de se trocar o sinal da função *a priori*), que deve ser escrita numa `m-file` do tipo função. Opcionalmente, esta `m-file` poderá também conter as primeiras e/ou segundas derivadas da função. `x0` é o valor inicial e na estrutura `options` podem alterar-se alguns dos valores que o MATLAB tem por defeito, nomeadamente definir se é pretendido usar as derivadas fornecidas na `m-file`. Para se saber que opções podem ser alteradas e quais os valores que se encontram por defeito no MATLAB, basta escrever na janela de comandos `fminunc('defaults')`. Para se alterar a estrutura `options` deve usar-se a sintaxe

```
options = optimset('param1',value1,'param2',value2,...)
```

Os valores que são *strings* devem indicar-se entre plicas (`'`).

Exemplo 13.6 *Resolva o problema* Aluffi-Pentini,

$$\min_x f(x) \equiv 0.25x_1^4 - 0.5x_1^2 + 0.1x_1 + 0.5x_2^2,$$

considerando o valor inicial $(-1, 0.5)$,

- usando o método quasi-Newton sem fornecer as primeiras derivadas da função objetivo.*
- usando o método quasi-Newton com procura unidimensional, fornecendo as primeiras derivadas da função objetivo.*
- usando o método de Newton com regiões de confiança, fornecendo as primeiras derivadas da função objetivo.*

d) usando o método de Newton com regiões de confiança, fornecendo as primeiras e segundas derivadas da função objetivo.

a) `function f=exemplo(x)`

```
f=0.25*x(1)^4-0.5*x(1)^2+0.1*x(1)+0.5*x(2)^2;
```

```
>> x1=[-1 0.5];
```

```
>> [x,f,e,o]=fminunc('exemplo',x1)
```

b) `function [f,g]=exemplo(x)`

```
f=0.25*x(1)^4-0.5*x(1)^2+0.1*x(1)+0.5*x(2)^2;
```

```
if nargin>1
```

```
    g=[x(1)^3-x(1)+0.1  
        x(2)];
```

```
end
```

```
>> op=optimset('GradObj','on','LargeScale','off');
```

```
>> x1=[-1 0.5];
```

```
>> [x,f,e,o]=fminunc('exemplo',x1,op)
```

c) `function [f,g]=exemplo(x)`

```
f=0.25*x(1)^4-0.5*x(1)^2+0.1*x(1)+0.5*x(2)^2;
```

```
if nargin>1
```

```
    g=[x(1)^3-x(1)+0.1  
        x(2)];
```

```
end
```

```
>> op=optimset('GradObj','on');
```

```
>> x1=[-1 0.5];
```

```
>> [x,f,e,o]=fminunc('exemplo',x1,op)
```

```
d) function [f,g,h]=exemplo(x)
    f=0.25*x(1)^4-0.5*x(1)^2+0.1*x(1)+0.5*x(2)^2;
    if nargin>1
        g=[x(1)^3-x(1)+0.1
            x(2)];
        if nargin>2
            h=[3*x(1)^2-1 0
                0 1];
        end
    end
end

>> op=optimset('GradObj','on','Hessian','on');
>> x1=[-1 0.5];
>> [x,f,e,o]=fminunc('exemplo',x1,op)
```

13.8.2 Exercícios para resolver em MATLAB

1. Resolva o problema *Aluffi-Pentini*,

$$\min_x f(x) \equiv 0.25x_1^4 - 0.5x_1^2 + 0.1x_1 + 0.5x_2^2,$$

considerando o valor inicial $(-1, 0.5)$,

- usando o método quasi-Newton sem fornecer as primeiras derivadas da função objetivo.
 - usando o método quasi-Newton com procura unidimensional, fornecendo as primeiras derivadas da função objetivo.
 - usando o método de Newton com regiões de confiança, fornecendo as primeiras derivadas da função objetivo.
 - usando o método de Newton com regiões de confiança, fornecendo as primeiras e segundas derivadas da função objetivo.
2. No planeamento da produção de dois produtos, uma determinada companhia espera obter lucros iguais a P :

$$P(x_1, x_2) = \alpha_1(1 - e^{-\beta_1 x_1}) + \alpha_2(1 - e^{-\beta_2 x_2}) + \alpha_3(1 - e^{-\beta_3 x_1 x_2}) - x_1 - x_2,$$

em que x_1 é a quantia gasta para produzir e promover o produto 1, x_2 é a quantia gasta para produzir e promover o produto 2 e os α_i e β_i são constantes definidas. P , x_1 e x_2 estão em unidades de 10^5 euros. Calcule o lucro máximo para as seguintes condições:

$$\alpha_1 = 3, \alpha_2 = 4, \alpha_3 = 1, \beta_1 = 1.2, \beta_2 = 1.5, \text{ e } \beta_3 = 1.$$

- Resolva o problema usando o método quasi-Newton sem fornecer as primeiras derivadas da função objetivo. Considere a aproximação inicial $(1, 1)$.
- Resolva o problema usando o método quasi-Newton com procura unidimensional, fornecendo as primeiras derivadas da função objetivo. Considere a aproximação inicial da alínea anterior.
- Resolva novamente o problema mas selecione agora o método de Newton com regiões de confiança.

3. Suponha que pretendia representar um número positivo A na forma de um produto de quatro fatores positivos x_1, x_2, x_3, x_4 . Para $A = 2401$, determine esses fatores de tal forma que a sua soma seja a menor possível.

Formule o problema como um problema de otimização sem restrições em função das três variáveis x_1, x_2 e x_3 .

A partir da aproximação inicial $(x_1, x_2, x_3)^{(1)} = (6, 7, 5)$, use o método quasi-Newton (com fórmula DFP), para calcular esses fatores. Na paragem do processo iterativo use TolX=TolFun=0.0001.

4. Resolva o problema *Epistatic Michalewicz*

$$\min_x f(x) \equiv - \sum_{i=1}^n \sin(y_i) \left(\sin \left(\frac{iy_i^2}{\pi} \right) \right)^{2m}$$

$$y_i = \begin{cases} x_i \cos(\theta) - x_{i+1} \sin(\theta), & i = 1, 3, 5, \dots, < n \\ x_i \sin(\theta) + x_{i+1} \cos(\theta), & i = 2, 4, 6, \dots, < n \\ x_i & i = n \end{cases}$$

pelo método quasi-Newton (sem fornecer derivadas) para $n = 5$ e para $n = 10$. Considere

$$\theta = \frac{\pi}{6}, m = 10 \text{ e o valor inicial } x^{(1)} = \begin{cases} 2, & i = 1, 3, 5, \dots, \leq n \\ 1, & i = 2, 4, 6, \dots, \leq n \end{cases}.$$

5. Considere o problema *Griewank*

$$\min_x f(x) \equiv 1 + \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos \left(\frac{x_i}{\sqrt{i}} \right).$$

Resolva-o pelo método quasi-Newton com fórmula DFP para $n = 10$ e $n = 25$. Considere o valor inicial $x^{(1)} = (1, 1, \dots, 1)^T$.

Capítulo 14

Método de Nelder-Mead

O método do simplex de Nelder-Mead é um dos métodos mais conhecidos na classe dos métodos de otimização sem derivadas (“derivative-free optimization”). Estes são os métodos adequados quando se está em presença de uma função objetivo $f(x)$ não suave, não linear, descontínua e não convexa.

Definição 14.0.1 *A função $f(x)$ diz-se suave se é continuamente diferenciável até à segunda ordem no seu domínio, isto é, as primeiras e as segundas derivadas existem em todos os pontos do domínio. Uma função não suave pode ser contínua e não ser diferenciável em alguns pontos do domínio.*

Estas características também podem estar presentes nas restrições do problema, mas este capítulo diz respeito apenas a problemas sem restrições.

Exemplo 14.1 *Considere-se a seguinte função não diferenciável*

$$f(x_1, x_2) = \begin{cases} 5\sqrt{9x_1^2 + 16x_2^2} & \text{para } x_1 \geq |x_2| \\ 9x_1 + 16|x_2| & \text{para } 0 < x_1 < |x_2| \\ 9x_1 + 16x_2 - x_1^9 & \text{para } x_1 \leq 0, x_2 \geq 0 \\ 9x_1 - 16x_2 - x_1^9 & \text{para } x_1 \leq 0, x_2 < 0 \end{cases}$$

A função f não é diferenciável em pontos que verificam $x_1 \leq 0$ e $x_2 = 0$. Significa que f não é suave.

Se se usar um método baseado em derivadas para calcular um ponto estacionário a partir de um ponto inicial que verifique $x_1 > |x_2| > (9/16)^2|x_1|$, o processo converge para $(0,0)$.

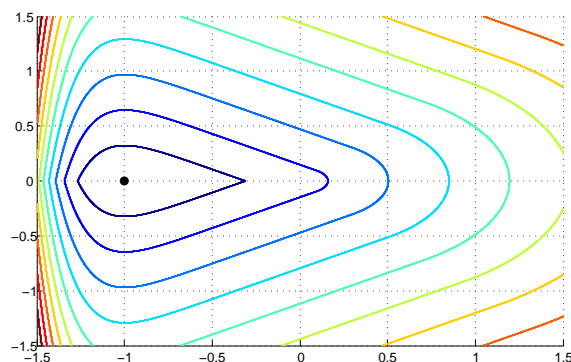


Figura 14.1: Função não diferenciável.

Neste ponto o gradiente pode não ser nulo. Na realidade, a solução para o problema $\min f(x)$ é $x^* = (-1, 0)^T$.

Se a função é não suave não devem usar-se métodos que recorram a derivadas para calcular a solução, como pode verificar-se no Exemplo 14.1. Mesmo que a derivada exista num certo ponto y e se $y \approx \bar{x}$, em que \bar{x} é um ponto onde a derivada não existe, não deve usar-se para prever o comportamento de f em \bar{x} . A informação sobre a derivada também não deve usar-se, neste caso, para terminar o processo iterativo pois não é necessariamente nula perto de \bar{x} .

Casos como o descrito no Exemplo 14.1 têm originado um aumento da popularidade dos métodos sem derivadas.

O método de Nelder-Mead é um método baseado num simplex. Um simplex em \mathbb{R}^n pode ser entendido como um poliedro com $n + 1$ vértices distintos. Quando os lados do simplex têm todos o mesmo comprimento, este diz-se regular. Assim, em \mathbb{R}^2 um simplex regular é um triângulo equilátero e em \mathbb{R}^3 um tetraedro. Se os vértices de um simplex em \mathbb{R}^n forem denotados por X_i , $i = 1, \dots, n + 1$ e substituirmos um dos vértices por W , obtém-se um novo simplex. Pode assim gerar-se uma sequência de simplex alterando um vértice de cada vez.

Se for estabelecido um conjunto de regras para alterar um dado vértice do simplex e essas regras se basearem apenas no valor da função objetivo f , é possível gerar uma sequência de simplex, cada um deles gerado com base no anterior, avaliando a função objetivo. Se estas regras forem tais que esta sequência contém o minimizante de f , x^* , então deverá existir um método de estimar x^* , sendo que a precisão da estimativa depende do tamanho do simplex final que contém x^* .

O método de Nelder-Mead é iterativo e define em cada iteração um simplex, isto é, um

poliedro. Em \mathbb{R}^n , sejam $x_1, x_2, \dots, x_n, x_{n+1}$ os $n + 1$ pontos do simplex de dimensão n . Então

$$S_k = \langle X_1, X_2, \dots, X_n, X_{n+1} \rangle$$

representa o simplex da iteração k em que os vértices já estão ordenados por ordem crescente dos valores da função objetivo, isto é, $f(X_1) \leq f(X_2) \leq \dots \leq f(X_n) \leq f(X_{n+1})$. Os vértices mais importantes do simplex são

- X_1 – o melhor vértice;
- X_n – o segundo pior vértice;
- X_{n+1} – o pior vértice.

Por exemplo, em \mathbb{R}^2 , o simplex formado pelos $n + 1 = 3$ pontos define um triângulo (Figura 14.2).

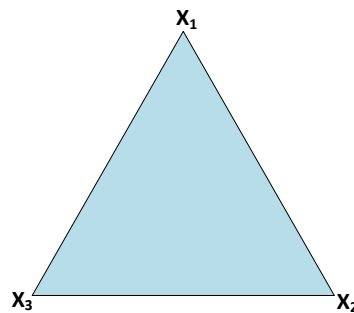


Figura 14.2: Simplex em \mathbb{R}^2 .

Em cada iteração deste método definem-se pontos auxiliares, que são candidatos a vértices do novo simplex. Serão rejeitados ou aceites de acordo com a comparação do seu valor da função objetivo com os valores da função objetivo nos vértices mais importantes do simplex referidos acima - $f(X_1)$, $f(X_n)$ e $f(X_{n+1})$.

Há um conjunto de operações básicas que permitem construir os pontos auxiliares:

- refletir;
- expandir;
- contrair para o interior;
- contrair para o exterior;

- encolher.

Seja

$$S_k = \langle X_1, \dots, X_{n+1} \rangle$$

o simplex de uma iteração k já ordenado.

A explicação que se segue é ilustrada para \mathbb{R}^2 , mas em dimensões superiores os princípios a seguir são os mesmos, aumentando apenas o número de vértices no simplex.

Em cada iteração começa-se por calcular o centróide do simplex (Figura 14.3), que é o ponto médio do hiperplano definido por X_1, \dots, X_n , e é dado por

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i.$$

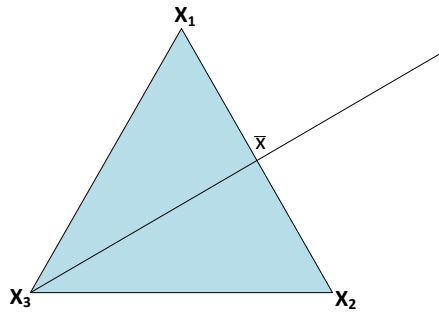


Figura 14.3: Centróide em \mathbb{R}^2 .

De seguida, calcula-se o vértice refletido (Figura 14.4).

$$x_r = (1 + \alpha)\bar{x} - \alpha X_{n+1}.$$

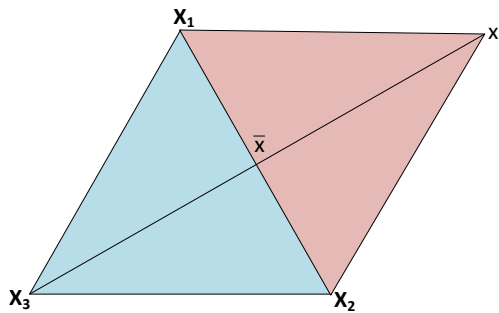


Figura 14.4: Vértice refletido em \mathbb{R}^2 com $\alpha = 1$.

Se x_r for bom, isto é, $f(X_1) \leq f(x_r) < f(X_n)$, aceita-se x_r e o simplex da iteração seguinte é

$$S_{k+1} = \langle X_1, \dots, X_n, x_r \rangle.$$

Se x_r for muito bom, isto é, $f(x_r) < f(X_1)$, faz-se uma expansão do simplex (Figura 14.5),

$$x_e = \gamma x_r + (1 - \gamma)\bar{x}.$$

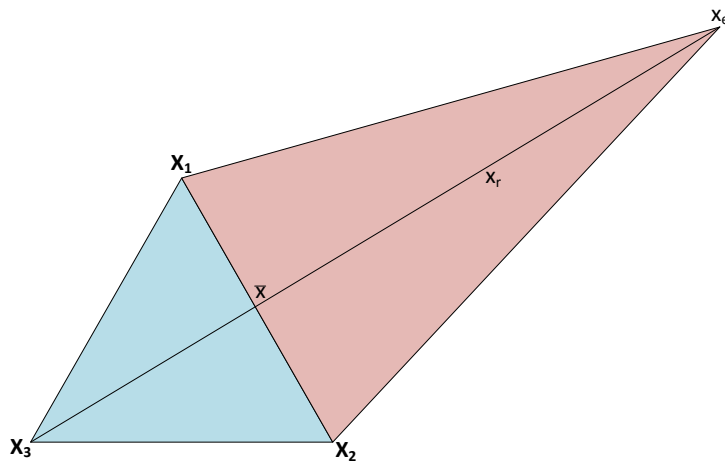


Figura 14.5: Vértice expandido em \mathbb{R}^2 com $\gamma = 2$.

Se x_e for muito bom, isto é, $f(x_e) < f(X_1)$, aceita-se x_e e o simplex da iteração seguinte é

$$S_{k+1} = \langle X_1, \dots, X_n, x_e \rangle,$$

caso contrário aceita-se x_r e o simplex da iteração seguinte é

$$S_{k+1} = \langle X_1, \dots, X_n, x_r \rangle.$$

Se x_r for fraco, isto é, $f(X_n) \leq f(x_r) < f(X_{n+1})$, faz-se uma contração para o exterior (Figura 14.6),

$$\hat{x}_c = \beta x_r + (1 - \beta)\bar{x}.$$

Se \hat{x}_c for bom, isto é, $f(\hat{x}_c) < f(X_n)$, aceita-se \hat{x}_c e o simplex da próxima iteração é

$$S_{k+1} = \langle X_1, \dots, X_n, \hat{x}_c \rangle,$$

caso contrário encolhe-se o simplex (Figura 14.8).

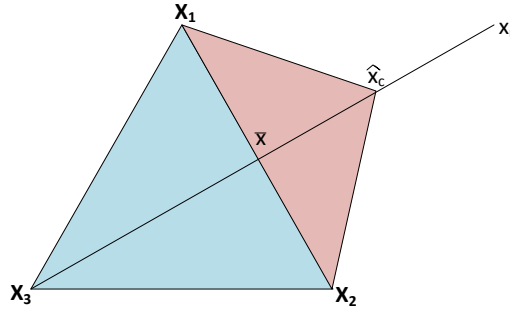


Figura 14.6: Vértice contraído para o exterior em \mathbb{R}^2 com $\beta = 1/2$.

Se x_r for muito fraco, isto é, $f(x_r) \geq f(X_{n+1})$, faz-se uma contração para o interior (Figura 14.7),

$$x_c = \beta X_{n+1} + (1 - \beta)\bar{x}.$$

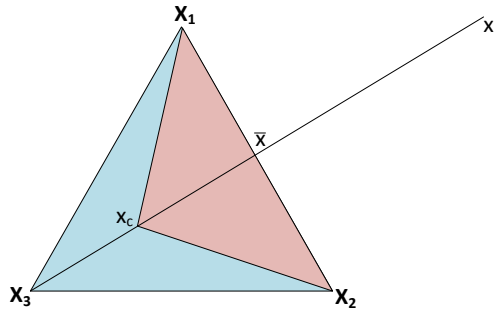


Figura 14.7: Vértice contraído para o interior.

Se x_c for bom, isto é, $f(x_c) < f(X_n)$, aceita-se x_c e o simplex da iteração seguinte é

$$S_{k+1} = \langle X_1, \dots, X_n, x_c \rangle,$$

caso contrário encolhe-se o simplex (Figura 14.8).

Encolher o simplex consiste em substituir cada um dos vértices X_i , $i = 2, \dots, n + 1$ pelo segmento médio do segmento que une X_i a X_1 , isto é

$$x_i = \frac{X_i + X_1}{2},$$

e o simplex da iteração seguinte é

$$S_{k+1} = \langle X_1, x_2, \dots, x_n, x_{n+1} \rangle.$$

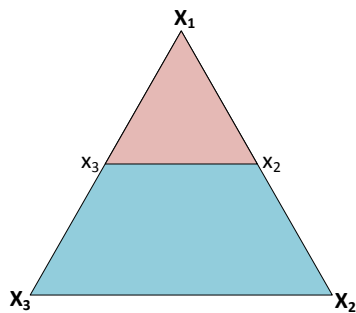


Figura 14.8: Encolher o simplex em \mathbb{R}^2 .

Em suma,

$$x_r \left\{ \begin{array}{ll} \begin{array}{l} \text{muito bom} \\ f(x_r) < f(X_1) \end{array} & \Rightarrow x_e \left\{ \begin{array}{ll} \text{muito bom} & \Rightarrow \text{aceita-se } x_e \\ f(x_e) < f(X_1) & \\ \text{caso contrário} & \Rightarrow \text{aceita-se } x_r \end{array} \right. \\ \\ \begin{array}{l} \text{bom} \\ f(X_1) \leq f(x_r) < f(X_n) \end{array} & \Rightarrow \text{aceita-se } x_r \\ \\ \begin{array}{l} \text{fraco} \\ f(X_n) \leq f(x_r) < f(X_{n+1}) \end{array} & \Rightarrow \hat{x}_c \left\{ \begin{array}{ll} \text{bom} & \Rightarrow \text{aceita-se } \hat{x}_c \\ f(\hat{x}_c) < f(X_n) & \\ \text{caso contrário} & \Rightarrow \text{encolhe-se o simplex} \end{array} \right. \\ \\ \begin{array}{l} \text{muito fraco} \\ f(x_r) \geq f(X_{n+1}) \end{array} & \Rightarrow x_c \left\{ \begin{array}{ll} \text{bom} & \Rightarrow \text{aceita-se } x_c \\ f(x_c) < f(X_n) & \\ \text{caso contrário} & \Rightarrow \text{encolhe-se o simplex} \end{array} \right. \end{array} \right.$$

Critério de paragem

O critério de paragem do método de Nelder-Mead consiste em verificar se o tamanho relativo do simplex é inferior ou igual a uma quantidade positiva pequena. Significa que o processo iterativo para se

$$\frac{1}{\Delta} \max_{2 \leq i \leq n+1} \|X_i - X_1\|_2 \leq \varepsilon,$$

com $\Delta = \max(1, \|X_1\|_2)$. Para verificar o critério de paragem é necessário que o simplex se encontre ordenado.

Se o critério de paragem for verificado, o vértice do simplex com menor valor da função objectivo, X_1 , é considerado como a melhor aproximação calculada à solução. Caso o critério de paragem não se verifique, o processo iterativo continua. Outras implementações diferentes do método de Nelder-Mead podem surgir com critérios de paragem diferentes.

O método de Nelder-Mead encontra-se descrito no Algoritmo 14.6.

Algoritmo 14.6 Método de Nelder-Mead

ler: X_1, \dots, X_{n+1} e ε , $\alpha = 1$, $\beta = 0.5$, $\gamma = 2$

ordenar o simplex de acordo com o valor da função ($f(X_1) \leq \dots \leq f(X_{n+1})$) $S_1 \leftarrow \langle X_1, \dots, X_{n+1} \rangle$ $k \leftarrow 1$ **repetir**

$$\bar{x} \leftarrow \frac{1}{n} \sum_{i=1}^n X_i$$

$$x_r \leftarrow (1 + \alpha)\bar{x} - \alpha X_{n+1}$$

se $f(x_r) < f(X_n)$ **então****se** $f(x_r) \geq f(X_1)$ **então**

$$S_{k+1} \leftarrow \langle X_1, \dots, X_n, x_r \rangle$$

senão

$$x_e \leftarrow \gamma x_r + (1 - \gamma)\bar{x}$$

se $f(x_e) < f(X_1)$ **então**

$$S_{k+1} \leftarrow \langle X_1, \dots, X_n, x_e \rangle$$

senão

$$S_{k+1} \leftarrow \langle X_1, \dots, X_n, x_r \rangle$$

fim se**fim se****senão****se** $f(x_r) \geq f(X_{n+1})$ **então**

$$x_c \leftarrow \beta X_{n+1} + (1 - \beta)\bar{x}$$

se $f(x_c) < f(X_n)$ **então**

$$S_{k+1} \leftarrow \langle X_1, \dots, X_n, x_c \rangle$$

senão

$$x_i \leftarrow \frac{X_i + X_1}{2}, \quad i = 2, \dots, n + 1$$

$$S_{k+1} \leftarrow \langle X_1, x_2, \dots, x_{n+1} \rangle$$

fim se**senão**

$$\hat{x}_c \leftarrow \beta x_r + (1 - \beta)\bar{x}$$

se $f(\hat{x}_c) < f(X_n)$ **então**

$$S_{k+1} \leftarrow \langle X_1, \dots, X_n, \hat{x}_c \rangle$$

senão

$$x_i \leftarrow \frac{X_i + X_1}{2}, \quad i = 2, \dots, n + 1$$

$$S_{k+1} \leftarrow \langle X_1, x_2, \dots, x_{n+1} \rangle$$

fim se**fim se****fim se**ordenar o simplex de acordo com o valor da função ($f(X_1) \leq \dots \leq f(X_{n+1})$)

$$S_{k+1} \leftarrow \langle X_1, \dots, X_{n+1} \rangle$$

$$k \leftarrow k + 1$$

$$\Delta = \max(1, \|X_1\|_2)$$

até $\frac{1}{\Delta} \times \max_{2 \leq i \leq n+1} \|X_i - X_1\|_2 \leq \varepsilon$

$$x_{\min} \leftarrow X_1$$

Exemplo 14.1 Considere um sistema de duas molas em que é aplicada uma força de deformação P com duas componentes P_1 e P_2 . Pretende-se determinar os deslocamentos x_1 e x_2 das molas que minimizam a energia potencial total EP , definida pela seguinte expressão:

$$EP(x_1, x_2) = \frac{1}{2}K_1 \left(\sqrt{x_1^2 + (l_1 - x_2)^2} - l_1 \right)^2 + \frac{1}{2}K_2 \left(\sqrt{x_1^2 + (l_2 + x_2)^2} - l_2 \right)^2 - P_1x_1 - P_2x_2.$$

Sabendo que as características do sistema são: $l_1 = 10$, $l_2 = 10$, $K_1 = 8$, $K_2 = 1$, $P_1 = 5$ e $P_2 = 5$, resolva o problema através do método de Nelder-Mead com $\varepsilon = 0.5$ (ou duas iterações).

Considere os seguintes pontos iniciais: $(5, 2)$, $(3.25, 2.5)$ e $(0, 0)$.

Resolução:

$$f(x_1, x_2) = 4 \left(\sqrt{x_1^2 + (10 - x_2)^2} - 10 \right) + 0.5 \left(\sqrt{x_1^2 + (10 + x_2)^2} - 10 \right) - 5x_1 - 5x_2$$

Iniciar o algoritmo de Nelder-Mead ($\varepsilon = 0.5$):

$$\left\langle \underbrace{\begin{pmatrix} 5 \\ 2 \end{pmatrix}}_{-29.2185}, \underbrace{\begin{pmatrix} 3.25 \\ 2.5 \end{pmatrix}}_{-11.1610}, \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_0 \right\rangle$$

• 1ª iteração

$$\bar{x} = \begin{pmatrix} 4.125 \\ 2.25 \end{pmatrix}$$

Calcular x_r

$$x_r = 2 \times \begin{pmatrix} 4.125 \\ 2.25 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 8.25 \\ 4.5 \end{pmatrix}}_{-41.3920}$$

$f(x_r) < f(X_2)$ (verdadeiro) $f(x_r) \geq f(X_1)$ (falso), logo expandir o simplex

$$x_e = 2 \times \begin{pmatrix} 8.25 \\ 4.5 \end{pmatrix} - \begin{pmatrix} 4.125 \\ 2.25 \end{pmatrix} = \underbrace{\begin{pmatrix} 12.375 \\ 6.75 \end{pmatrix}}_{-5.7885}$$

$f(x_e) < f(X_1)$ (falso), aceitar x_r

$$\left\langle \underbrace{\begin{pmatrix} 8.25 \\ 4.5 \end{pmatrix}}_{-41.3920}, \underbrace{\begin{pmatrix} 5 \\ 2 \end{pmatrix}}_{-29.2185}, \underbrace{\begin{pmatrix} 3.25 \\ 2.5 \end{pmatrix}}_{-11.1610} \right\rangle$$

- Critério de Paragem

$$\|X_1\|_2 = 9.3975 \quad \|X_2 - X_1\|_2 = 4.1003 \quad \|X_3 - X_1\|_2 = 5.3852$$

$$\Delta = \max(1, \|X_1\|_2) = 9.3975$$

$$\frac{1}{\Delta} \times \max(\|X_2 - X_1\|_2, \|X_3 - X_1\|_2) = \frac{5.3852}{9.3975} = 0.5730 \leq \varepsilon \quad (falso)$$

- 2ª iteração

$$\bar{x} = \begin{pmatrix} 6.625 \\ 3.25 \end{pmatrix}$$

Calcular x_r

$$x_r = 2 \times \begin{pmatrix} 6.625 \\ 3.25 \end{pmatrix} - \begin{pmatrix} 3.25 \\ 2.5 \end{pmatrix} = \underbrace{\begin{pmatrix} 10 \\ 4 \end{pmatrix}}_{-32.9988}$$

$f(x_r) < f(X_2)$ (verdadeiro), logo aceitar x_r

$$\left\langle \underbrace{\begin{pmatrix} 8.25 \\ 4.5 \end{pmatrix}}_{-41.3920}, \underbrace{\begin{pmatrix} 10 \\ 4 \end{pmatrix}}_{-32.9988}, \underbrace{\begin{pmatrix} 5 \\ 2 \end{pmatrix}}_{-29.2185} \right\rangle$$

- Critério de Paragem

$$\|X_1\|_2 = 9.3975 \quad \|X_2 - X_1\|_2 = 1.8200 \quad \|X_3 - X_1\|_2 = 4.1003$$

$$\Delta = \max(1, \|X_1\|_2) = 9.3975$$

$$\frac{1}{\Delta} \times \max(\|X_2 - X_1\|_2, \|X_3 - X_1\|_2) = \frac{4.1003}{9.3975} = 0.4363 \leq \varepsilon \quad (falso)$$

$$x_{\min} \approx \begin{pmatrix} 8.25 \\ 4.5 \end{pmatrix} \text{ e } f_{\min} \approx -41.3920$$

14.1 Exercícios

1. Considere um sistema de duas molas em que é aplicada uma força de deformação P com duas componentes P_1 e P_2 . Pretende-se determinar os deslocamentos x_1 e x_2 das molas que minimizam a energia potencial total EP , definida pela seguinte expressão:

$$EP(x_1, x_2) = \frac{1}{2}K_1 \left(\sqrt{x_1^2 + (l_1 - x_2)^2} - l_1 \right)^2 + \frac{1}{2}K_2 \left(\sqrt{x_1^2 + (l_2 + x_2)^2} - l_2 \right)^2 - P_1x_1 - P_2x_2.$$

Sabendo que as características do sistema são: $l_1 = 10$, $l_2 = 10$, $K_1 = 8$, $K_2 = 1$, $P_1 = 5$ e $P_2 = 5$, resolva o problema através do método de Nelder-Mead com $\varepsilon = 0.5$ (ou duas iterações). Considere os seguintes pontos iniciais: $(5, 2)$, $(3.25, 2.5)$ e $(0, 0)$.

2. Calcule o mínimo da função $f(x)$ definida por

$$f(x_1, x_2) = \max((x_1 - 1)^2, x_1^2 + 4(x_2 - 1)^2)$$

implementando o método de Nelder-Mead, tomando para conjunto inicial os vetores

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle.$$

e $\varepsilon = 0.5$.

3. Calcule o mínimo da função $f(x)$ definida por

$$f(x_1, x_2) = \max(|x_1|, |x_2 - 1|)$$

implementando o método de Nelder-Mead, tomando para conjunto inicial os vetores

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0 \end{pmatrix} \right\rangle.$$

e $\varepsilon = 0.5$.

4. Calcule o máximo da seguinte função não diferenciável

$$f(x_1, x_2) = -|x_1x_2| - x_2^2$$

usando o método de Nelder-Mead. Inicie o processo iterativo com o seguinte simplex:

$$\left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\rangle.$$

Para a paragem do processo iterativo use $\varepsilon = 0.5$ ou $n_{\max} = 4$.

14.2 Soluções

1. 2 iterações; $x_{\min} \approx \begin{pmatrix} 8.25 \\ 4.5 \end{pmatrix}$; $f_{\min} \approx -41.3920$
2. 4 iterações; $x_{\min} \approx \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}$; $f_{\min} \approx 0.25$
3. 4 iterações; $x_{\min} \approx \begin{pmatrix} -0.1875 \\ 0.875 \end{pmatrix}$; $f_{\min} \approx 0.1875$
4. 4 iterações; $x_{\min} \approx \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; $f_{\min} \approx 0$

14.3 MATLAB

14.3.1 fminsearch

A função `fminunc` encontra o mínimo de uma função multidimensional usando o método de Nelder-Mead. A sua sintaxe é

```
[xmin,fmin,exitflag,output] = fminsearch('func',x0,options)
```

em que os parâmetros de saída definem o minimizante em `xmin`, o valor mínimo da função `fmin`, a `exitflag` define a forma como parou o algoritmo (a desejável é 1 - significa que o processo convergiu para um ponto estacionário) e na estrutura `output` encontra-se informação sobre o processo iterativo (número de iterações, número de cálculos da função, algoritmo usado...).

Os parâmetros de entrada a fornecer são a função que se pretende minimizar (caso o problema seja de maximização tem de se trocar o sinal da função *a priori*), que deve ser escrita numa `m-file` do tipo função. `x0` é o valor inicial e na estrutura `options` podem alterar-se alguns dos valores que o MATLAB tem por defeito. Para se saber que opções podem ser alteradas e quais os valores que se encontram por defeito no MATLAB, basta escrever na janela de comandos `fminsearch('defaults')`. Para se alterar a estrutura `options` deve usar-se a sintaxe

```
options = optimset('param1',value1,'param2',value2,...)
```

Os valores que são *strings* devem indicar-se entre plicas (').

Exemplo 14.2 *Resolva o problema*

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2)$$

com $f(x_1, x_2) = \max\{|x_1|, |x_2 - 1|\}$. Como aproximação inicial considere o ponto $(1, 1)$.

Começa-se por colocar a função a minimizar numa `m-file`:

```
function f=NM1(x)
f=max(abs(x(1)),abs(x(2)-1));
```

Os comandos para resolver o problema são

```
>> x1=[1 1];
>> [x,f,e,o]=fminsearch('NM1',x1)
```

14.3.2 Exercícios para resolver em MATLAB

1. Resolva o problema

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2)$$

com $f(x_1, x_2) = \max\{|x_1|, |x_2 - 1|\}$. Como aproximação inicial considere o ponto $(1, 1)$.

2. Considere o seguinte problema não diferenciável

$$\min_{x \in \mathbb{R}^2} f(x) \equiv \max\{x_1^2 + x_2^4, (2 - x_1)^2 + (2 - x_2)^2, 2e^{-x_1+x_2}\}.$$

A partir da aproximação inicial $x = (1, -0.1)^T$, calcule a solução, usando o método mais adequado. Repita o processo com a seguinte aproximação inicial $x = (2, 2)^T$.

Resolva novamente o problema a partir de $x = (-10, -10)^T$.

Com qual das aproximações iniciais o processo exigiu menos cálculos da função objetivo?

3. Considere o seguinte problema não diferenciável

$$\min_{x \in \mathbb{R}^n} f(x) \equiv n \left(\max_{1 \leq i \leq n} x_i \right) - \sum_{i=1}^n x_i.$$

Para $n = 2$ e a partir da aproximação inicial $x_i = i - (\frac{n}{2} + 0.5)$, $i = 1, \dots, n$, calcule a solução.

Repita a resolução considerando agora $n = 5$ e TolX = 10^{-20} . Resolva ainda acrescentando a opção MaxFunEvals=10000. Acrescente ainda a opção MaxIter=10000. Comente os resultados.

4. Considere o seguinte problema não diferenciável

$$\min_{x \in \mathbb{R}^n} f(x) \equiv \prod_{i=1}^n x_i - \left(\min_{1 \leq i \leq n} x_i \right).$$

Para $n = 2$ e a partir da aproximação inicial $x_i = i - (\frac{n}{2} + 0.5)$, $i = 1, \dots, n$, calcule a solução.

Repita a resolução considerando agora $n = 5$ e MaxFunEvals = 5000.

5. Considere o seguinte problema não diferenciável

$$\min_{x \in \mathbb{R}^2} f(x) \equiv \max\{x_1^2 + x_2^2, x_1^2 + x_2^2 + \omega(-4x_1 - x_2 + 4), x_1^2 + x_2^2 + \omega(-x_1 - 2x_2 + 6)\}.$$

A partir da aproximação inicial $x = (-1, 5)^T$, calcule a solução, usando o método mais adequado e considerando $\omega = 500$. A partir da mesma aproximação inicial, volte a resolver o problema, mas agora fazendo $\omega = 1000$.

Repita mais uma vez considerando $\omega = 1500$.

Para que valor de ω , o processo iterativo é mais eficiente?

6. Considere o seguinte problema não diferenciável

$$\min_{x \in \mathbb{R}^4} f(x) \equiv \max_{1 \leq i \leq 21} |u_i(x)|$$

em que

$$u_i(x) = x_4 - (x_1 t_i^2 + x_2 t_i + x_3)^2 - \sqrt{t_i}$$

para $1 \leq i \leq 21$.

A partir da aproximação inicial $x_i = 1$, $i = 1, \dots, 4$, calcule a solução, usando o método mais adequado e os seguintes valores $t_i = 0.25 + 0.75(i - 1)/20$, $i = 1, \dots, 21$.

Repita o processo mas agora considere os seguintes parâmetros $t_i = 0.2i$, $i = 1, \dots, 21$.