UNIVERSIDADE DO MINHO

Licenciatura em Ciências da Computação

Análise Numérica

Duração: 2 horas 30 minutos	15 de janeiro de 2018
EXAME FINAL (COM CONSULTA)	

Deves escrever na tua folha de respostas todos os comandos executados no Matlab.

1. Considera o seguinte desenvolvimento em série de potências de x, válido para $|x| < 1$,

$$log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots + (-1)^{n+1}\frac{x^n}{n} + \cdots$$

onde $log(x)$ denota o logaritmo natural (de base e).

a) Para $x = 0.9$, qual o último termo que devemos adicionar para garantir que o erro de truncatura é inferior a 10^{-4}? Justifica a resposta.

b) Usa a *function* horner desenvolvida nas aulas para calcular a soma que aproxima o valor de $log(1.9)$ com erro inferior a 10^{-4}.

2. Se f é uma função derivável num ponto $x = a$, então

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

a) Calcula (no Matlab)

$$\frac{f(1+h) - f(1)}{h}$$

para $f(x) = log(x)$ e $h = 10^{-10}, h = 10^{-12}, h = 10^{-14}$ e $h = 10^{-16}$.

b) Por que é que as aproximações obtidas para a derivada $f'(1)$ são piores para valores de h mais próximos de zero?

3. No Matlab, executa sucessivamente as seguintes instruções

```
x1=fzero('(exp(x)+cos(x))/sin(x)',1)

x2=fzero('(exp(x)+cos(x))/sin(x)',2)

x3=fzero('(exp(x)+cos(x))/sin(x)',-2)
```

Indica qual ou quais dos valores x1, x2 e x3 aproximam uma raiz da equação

$$\frac{e^x + cos(x)}{sin(x)} = 0.$$

Justifica a tua resposta.

4. Considera os nós $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$ e os correspondentes valores nodais $y_i = sin(\frac{\pi}{2}x_i),\ i = 0, 1, 2, 3, 4$

   a) Para cada $i = 0, 1, 2, 3$, determina os valores de a_i e b_i tais que o polinómio $a_i + b_i(x - x_i)$ interpola (x_i, y_i) e (x_{i+1}, y_{i+1}).

   b) Sabendo que, neste caso,

$$\max_{x\in[x_i,x_{i+1}]} |(x - x_i)(x - x_{i+1})| = \frac{1}{4},$$

   determina um majorante para o erro

$$|a_i + b_i(x - x_i) - sin(\frac{\pi}{2}x)|$$

   onde $x \in [x_i, x_{i+1}]$. Apresenta os cálculos efetuados.

5. **a)** Usa a regra de Simpson composta e os valores dados na tabela

x	0	1	2	3	4
f(x)	0	1	0	-1	0

   para obter uma aproximação do valor de $\int_0^4 f(x)dx$. Apresenta os cálculos efetuados.

   b) Sabendo que a derivada de quarta ordem da função tabelada não excede, em módulo, no intervalo $[0, 4]$, o valor $(\frac{\pi}{2})^4$, determina um majorante para o erro cometido na aproximação calculada na alínea anterior.

6. Sejam L e U matrizes triangulares, inferior e superior, respetivamente, com entradas na diagonal principal todas diferentes de zero, e seja $A = L * U$.

   a) Qual é o processo mais eficiente para resolver o sistema $Ax = b$, dado um certo vetor b de termos independentes? Justifica a resposta.

   b) Explica, de forma sucinta, em que consiste o método de eliminação de Gauss com pivotação parcial e qual é a sua importância.

questão	1a	1b	2a	2b	3	4a	4b	5a	5b	6a	6b	Total
cotação	1,5	2	1,5	2	2	2	2	1,5	2	1,5	2	20

RESOLUÇÃO

1. a) Numa série alternada, o erro de truncatura (em valor absoluto) é inferior ao valor absoluto do primeiro termos que se despreza. Devemos adicionar até n=50 porque $0.9^{51}/51 = 9.09e-5$ e $0.9^{50}/50 = 1.03e-4$.

   b) Executando

   ```
   k=1:50; a=(-1).^(k+1)./k; horner([fliplr(a),0],0.9)
   ```

   obtemos o resultado 0.6418 que aproxima log(1.9) com erro inferior a 10^{-4}.

2. a) Executando

   ```
   h=[1e-10 1e-12 1e-14 1e-16]; (log(1+h)-log(1))./h
   ```

   obtemos

   ```
   1.0000    1.0001    0.9992         0
   ```

   b) Em teoria, isto é, em aritmética exata, as aproximações são melhores para valores de h menores. Na prática, isto não acontece por causa dos erros produzidos no cálculo de $log(1+h)$. Note-se que não se pode falar de cancelamento subtrativo no cálculo do numerador $log(1+h) - log(1)$ porque $log(1) = 0$ e não h portanto perda de algarismos significativos nesta subtração. O problema da perda de algarismos significativos no numerador é devido ao elevado erro de condição relativo de

   $$f(x) = log(1+x)$$

   para valores de x próximos de zero. Este número de condição é dado por

   $$xf'(x)/f(x) = \frac{\frac{x}{1+x}}{log(1+x)}$$

   que tende para infinito quando x tende para zero.

3. Obtem-se

$$\begin{aligned} x1 &= 2.9066e-16 \\ x2 &= 3.1416 \\ x3 &= -1.7461 \end{aligned}$$

x3 é a única raiz da equação. A função *fzero* assume como sendo raiz um ponto onde f muda de sinal o que acontece também para $x = 0$ e $x = \pi$ (note-se que x1 e x2 são aproximações de 0 e π, respetivamente). Mas este critério só pode aplicar-se se f for contínua o que não é verdade porque nestes pontos anula-se o denominador.

4. a) De

$$a_i + b_i(x_i - x_i) = y_i$$

resulta $a_i = y_i$ e de

$$a_i + b_i(x_{i+1} - x_i) = y_{i+1}$$

resulta

$$b_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i},$$

isto é, b_i é a diferença dividida de primeira ordem relativa aos nós x_i e x_{i+1}. Estes valores, que podem ser obtidos a partir da tabela produzida pela function TabDifDiv, são

$$b_0 = 1, b_1 = -1, b_2 = -1, b_3 = 1.$$

b) A expressão do erro do polinómio interpolador neste caso dá

$$f(x) - (a_i + b_i(x - x_i)) = (x - x_i)(x - x_{i+1})\frac{f''(\eta)}{2}$$

onde $\eta \in [x_i, x_{i+1}]$. Sendo $f(x) = sin(\frac{\pi}{2}x)$, é $f''(x) = -\frac{\pi^2}{4}sin(\frac{\pi}{2}x)$ e

$$\max_{x\in[x_i,x_{i+1}]} |f''(x)| = \pi^2/4$$

para $i = 0, 1, 2, 3, 4$. Assim, tem-se

$$|a_i + b_i(x - x_i) - sin(\frac{\pi}{2}x)| \leq \frac{1}{4}\frac{\pi^2/4}{2} = \pi^2/32 \approx 0.31$$

5. a) A regra composta de Simpson neste caso dá

$$\int_0^4 f(x)dx = \frac{h}{3}\left[f(0) + 4f(1) + 2f(2) + 4f(3) + f(4)\right] - \frac{h^2}{180}(4-0)f^{(iv)}(\eta)$$

onde $h = 1$ e $\eta \in [0, 4]$. No Matlab,

```
(0+4*1+2*0+4*(-1)+0)/3
```

produz o resultado zero.

b) Da expressão do erro de truncatura dada na alínea anterior conclui-se que o erro (em valor absoluto) é majorado por

$$\frac{1}{180}(4 - 0)(\pi/2)^4 \approx 0.1353$$

6. a) Conhecida a fatorização $A = LU$, o sistema $Ax = b$ escreve-se $L(Ux) = b$ ou seja $Ly = b$ com $y = Ux$. Assim, a solução x pode ser encontrada resolvendo os sistemas triangulares $Ly = b$ e $Ux = y$ (por esta ordem). Como a resolução de um sistema triangular (por substituição) requer apenas n^2 operações aritméticas, este processo é mais eficiente do que usar o método de eliminação de Gauss que requer $O(n^3)$operações aritméticas.

b) O método de eliminação de Gauss transforma o sistema $Ax = b$ (com n equações e n incógnitas) num sistema equivalente $Ux = b'$ onde U é triangular superior. Este processo envolve $n-1$ passos de redução, no k-ésimo passo de redução ($k = 1, \cdots, n-1$) são eliminadas as entradas na k-ésima coluna abaixo da diagonal principal. Porém, erros de arredondamento neste processo podem comprometer a estabilidade numérica do processo quando o pivot a_{kk} for, em valor absoluto, muito inferior á entrada a_{ik} ($i = k+1, \cdots, n$) a eliminar, originando multiplicadores muito grandes (em valor absoluto). Para evitar estes erros, é normalmente usada a técnica de pivotação parcial que consiste em determinar, no inicio do k-ésimo passo de redução, a posição p tal que

$$|a_{pk}| = \max_{i=k,\cdots n} |a_{ik}|$$

e proceder à troca das linhas p e k da matriz ampliada. Desta forma, os multiplicadores usados terão grandeza inferior á unidade e o método de eliminação de Gauss é numericamente estável (salvo raros casos em que só a chamada pivotação total pode resolver o problema dos erros).