GRAU D'ENGINYERIA INFORMÀTICA

Introducció a la Computació Científica

Semestre Tardor 2020 - Prova parcial de 9 de novembre

CADA APARTAT VAL 20 PUNTS I ES QUALIFICA SOBRE 100. RESOL ELS 2 EXERCICIS EN FULLS SEPARATS.

Exercici 1 [Errors]

Considerem la data d'avui escrita així: $A = 20201109 = 1001101000011111010010101_{2)}$.

(a) Troba el nombre màquina $\bar{A}$ que és la representació IEEE amb precisió simple d'A i l'error exacte $E = A - \bar{A}$ comès en aquesta representació.

(b) Mitjançant la fórmula aproximada de propagació de l'error, troba una fita aproximada de l'error absolut en calcular A^2 usant el valor representat $\bar{A}$: $\bar{A}^2$.

(c) Es vol calcular la diferència exacta D entre els quadrats d'A i $\bar{A}$:

$$D = A^2 - \bar{A}^2 \ .$$

Intenta calcular-la directament i explica la problemàtica amb què et trobes si no pots fer servir moltes xifres. Fes servir una fórmula millor des del punt de vista numèric per calcular-la, justificant breument perquè és millor.

Exercici 2 [Sistemes lineals]

El polinomi interpolador $p_3(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ de grau màxim 3 a la funció $f(x)$ en els nodes x_k $(k = 0, 1, 2)$ es pot trobar resolent el sistema d'equacions lineals $p_3(x_k) = f(x_k)$ $(k = 0, 1, 2, 3)$ que compleixen els coeficients a_0, a_1, a_2, a_3:

$$\begin{array}{ccccccccc} a_0 & + & x_0a_1 & + & x_0^2a_2 & + & x_0^3a_3 & = & f(x_0) \\ a_0 & + & x_1a_1 & + & x_1^2a_2 & + & x_1^3a_3 & = & f(x_1) \\ a_0 & + & x_2a_1 & + & x_2^2a_2 & + & x_2^3a_3 & = & f(x_2) \\ a_0 & + & x_3a_1 & + & x_3^2a_2 & + & x_3^3a_3 & = & f(x_3) \end{array}$$

Considera la funció $f(x) = x^4$. Es vol trobar el polinomi interpolador a $f(x)$ en les nodes $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 3$, on $f(x_0) = 0, f(x_1) = 1, f(x_2) = 16, f(x_3) = 81$.

El sistema d'equacions lineals que compleixen els coeficients a_0, a_1, a_2, a_3 de $p_3(x)$ és:

$$\mathbf{Aa} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 16 \\ 81 \end{pmatrix} = \mathbf{b} \ .$$

(a) Troba la solució del sistema d'equacions $\mathbf{Aa} = \mathbf{b}$ pel mètode de Gauss i escriu el polinomi interpolador $p_3(x)$.

(b) Calcula el determinant de la matriu $\det \mathbf{A}$ i el de la seva matriu inversa $\det \mathbf{A}^{-1}$ i la darrera columna $\mathbf{x}^{(4)}$ de la matriu inversa.

(c) Escriu les iteracions del mètode de Jacobi aplicat al sistema $\mathbf{Aa} = \mathbf{b}$ i indica com sabries si és convergent o no.

Idees per a la solució. **[Errors]**

(a) Fem la representació en punt flotant en base 2:

$$A = 1001101000011111010010101_{2)} = 1.001101000011111010010101_{2)} 2^{24} = \sigma(1+f)2^{q-1} \ .$$

Representació IEEE simple d'A:

- Signe σ: $s = 0$.
- Mantissa fraccionària amb 23 bits arrodonits: $\mathrm{fl}_{23}(f) = 0.00110100001111101001011_{2)}$.
- Exponent: $e = q - 1 + 127 = 24 + 127 = 10010111_{2)}$

La representació usant 4 bytes seria

$$\boxed{0|10010111|00110100001111101001011} \ .$$

El nombre màquina de representació d'A (24 bits significatius arrodonits) és així:

$$\bar{A} = 1001101000011111010010110_{2)} = 20201110 \ .$$

Error absolut exacte: $E = A - \bar{A} = -1$.

Fita de l'error absolut: $\epsilon_a(A) = 1$.

(b) Una fita aproximada de l'error absolut comès en aproximar A^2 per $\bar{A}^2$ ve donada per la fórmula de propagació de l'error:

$$\epsilon_a(A^2) \approx 2\bar{A}\epsilon_a(A) = 40402220 \ .$$

(c) Si no es pot treballar amb molts dígits significatius, l'ús directe de l'expressió té un problema de cancel·lació que no permet conèixer gaire xifres de la diferència. Treballant, per exemple, amb 12 dígits significatius s'obté:

$$D = A^2 - \bar{A}^2 = 0.408084845232 \cdot 10^{15} - 0.408084804830 \cdot 10^{15} = 0.000000040402 \cdot 10^{15} = 40402000$$

amb només 5 dígits significatius. Amb una fórmula millor, que evita la cancel·lacions, s'obté un resultat més precís:

$$A^2 - \bar{A}^2 = (A - \bar{A})(A + \bar{A}) = E(A + \bar{A}) = A + \bar{A} = 40402219 \ .$$

Idees per a la solució. **[Sistemes lineals]**

(a) La matriu ampliada del sistema és:

$$\left(\begin{array}{cccr|r} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 3 & 9 & 27 & 81 \end{array}\right) .$$

Aplicant el mètode de Gauss, es troben les matrius ampliades:

$$\left(\begin{array}{cccr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 2 & 4 & 8 & 16 \\ 0 & 3 & 9 & 27 & 81 \end{array}\right) , \quad \left(\begin{array}{cccr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 6 & 14 \\ 0 & 0 & 6 & 24 & 81 \end{array}\right) , \quad \left(\begin{array}{cccc|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 6 & 14 \\ 0 & 0 & 0 & 6 & 36 \end{array}\right) .$$

La solució del sistema lineal triangular superior resultant es troba per substitució enrera:

$$a_3 = 6 \ , \quad a_2 = -11 \ , \quad a_1 = 6 \ , \quad a_0 = 0 \ .$$

El polinomi interpolador és llavors:

$$p_3(x) = 6x^3 - 11x^2 + 6x \ .$$

(b) El determinat d'**A** coincideix amb el de la matriu triangular superior, que és el producte dels seus elements diagonals:

$$\det \mathbf{A} = 1 \cdot 1 \cdot 2 \cdot 6 = 12 \ .$$

El determinant de la inversa d'**A** és el recíproc del determinant d'**A**:

$$\det \mathbf{A}^{-1} = \frac{1}{12} \ .$$

Per trobar la quarta columna $\mathbf{x}^{(4)}$ de la inversa cal resoldre el sistema lineal $\mathbf{A}\mathbf{x}^{(4)} = \mathbf{e}^{(4)}$, on $\mathbf{e}^{(4)}$ és la quarta columna de la matriu identitat.

S'escriuen els sistemes lineals corresponents en la matriu estesa

$$\left(\begin{array}{cccr|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 8 & 0 \\ 1 & 3 & 9 & 27 & 1 \end{array}\right) .$$

Aplicant el mètode de Gauss, es troben les matrius ampliades:

$$\left(\begin{array}{cccr|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 2 & 4 & 8 & 0 \\ 0 & 3 & 9 & 27 & 1 \end{array}\right) , \quad \left(\begin{array}{cccr|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & 6 & 0 \\ 0 & 0 & 6 & 24 & 1 \end{array}\right) , \quad \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & 6 & 0 \\ 0 & 0 & 0 & 6 & 1 \end{array}\right) .$$

El terme independent no es veu afectat per la triangularització.

La solució del sistema lineal triangular superior resultant es troba per substitució enrera:

$$x_3^{(4)} = \frac{1}{6} \ , \quad x_2^{(4)} = -\frac{1}{2} \ , \quad x_1^{(4)} = \frac{1}{3} \ , \quad x_0^{(4)} = 0 \ .$$

La quarta columna d'$\mathbf{A}^{-1}$ és:

$$\mathbf{x}^{(4)} = \begin{pmatrix} 0 \\ \frac{1}{3} \\ -\frac{1}{2} \\ \frac{1}{6} \end{pmatrix} \ .$$

(c) Les iteracions del mètode de Jacobi s'escriuen així

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}^{(k+1)} = \begin{pmatrix} 0 \\ 1 \\ 4 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ \frac{1}{4} & \frac{1}{2} & 0 & 2 \\ \frac{1}{27} & \frac{1}{9} & \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}^{(k)} \ .$$

Tot i que no es demana fer-ho explícitament sinó explicar com es faria, s'indica la resolució completa a continuació.

La matriu d'iteració és així:

$$\mathbf{B}_J = -\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ \frac{1}{4} & \frac{1}{2} & 0 & 2 \\ \frac{1}{27} & \frac{1}{9} & \frac{1}{3} & 0 \end{pmatrix} \ .$$

Com que les normes $\|\mathbf{B}_J\|_\infty$, $\|\mathbf{B}_J\|_1$ són més grans que 1, caldria veure si el radi espectral de $\mathbf{B}_J$ és més petit que 1 o no, per decidir si és convergent o no.

L'equació carecterística que cumpleixen els valors propis és:

$$\det(\mathbf{B}_J - \lambda\mathbf{I}) = \begin{vmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 1 & 1 \\ \frac{1}{4} & \frac{1}{2} & \lambda & 2 \\ \frac{1}{27} & \frac{1}{9} & \frac{1}{3} & \lambda \end{vmatrix} = \lambda(\lambda^3 - \frac{23}{18}\lambda + \frac{7}{18}) = 0 \ .$$

En el cas que tots els valors propis tinguessin mòdul més petit que 1, seria convergent. Si hi hagués algun valor propi amb mòdul més gran que 1, seria divergent.

Un valor propi és zero i els altres compleixen:

$$p_3(\lambda) = 18\lambda^3 - 23\lambda + 7 = 0 \ .$$

Notem que $p_3(\lambda)$ és molt negatiu per a valors molt negatius de λ i que $p_3(-1) = 12 > 0$. Té doncs una solució més petita que -1, de mòdul més gran que 1.

El radi espectral de $\mathbf{B}_J$ és més gran que 1 i, per tant, el mètode seria divergent.