2SM46 June 1997

TIME ALLOWED: TWO HOURS AND A HALF

Instructions to candidates

Full marks may be obtained for FIVE complete answers.

1. a) Describe the standard computer representation of floating point numbers in binary form. Explain why numbers larger than a certain amount can not be represented. What happens if a particular calculation would produce a number which is too large to be represented? Explain why numbers smaller than a certain amount can not be represented accurately. What happens if a calculation produces such a number?

b) Write down the hexadecimal representation of the smallest representable single precision number greater than 1 and the largest representable single precision number smaller than 1. [You may assume that in the single precision representation the mantissa has 24 bits.]

c) Find the representation in hexadecimal form of the decimal number $x =9.7$.

Write down the hexadecimal representation of the number stored in single precision in the computer memory x^* which represents the decimal number 9.7. Evaluate the relative error of this stored number. Write down the next larger representable number in hexadecimal form. [You may assume that in the single precision representation the mantissa has 24 bits.]

d) Describe the method of Simple Iteration for the solution of the equation $x = f(x)$. How can one tell if the method is going to converge?

Show that the equation $x = 4e^{-x}$ has a solution between $x = 1.2$ and $x = 1.21$. Will the method of Simple Iteration converge for this equation?

Show that the equations $x = x/2 + 2e^{-x}$, $x = 2x/3 + 4e^{-x}/3$ and $x = \ln(4/x)$ all have the same solution as the equation $x = 4e^{-x}$. Which of these equations can be used to obtain a solution by Simple Iteration and which converges the fastest?

[20 marks]

2. Verify that the matrix A can be decomposed into the product $A = L.U$, where

$$A = \begin{pmatrix} 3 & 6 & -3 & 3 \\ 6 & 14 & -4 & 4 \\ 3 & 12 & 4 & 0 \\ -3 & -8 & 3 & 4 \end{pmatrix}, L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ -1 & -1 & 2 & 1 \end{pmatrix}, U = \begin{pmatrix} 3 & 6 & -3 & 3 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Find two matrices D and M such that D is a diagonal matrix and M has its leading diagonal elements all equal to 1 such that $U = D.M$.

Find the inverses of the matrices L and U and show that the inverse of A is given by:

$$A^{-1} = \begin{pmatrix} -344/3 & 74 & -21 & 12 \\ 38 & -49/2 & 7 & -4 \\ -28 & 18 & -5 & 3 \\ 11 & -7 & 2 & -1 \end{pmatrix}.$$

Explain the significance of the condition number of a matrix. Using whichever norm you like, find the condition number for the matrix A.

[20 marks]

3. The Crout decomposition of the matrix $A = N.M$ is

$$A = \begin{pmatrix} 5 & 2 & -1 & 7 \\ 10 & 12 & 10 & 10 \\ 15/2 & 7 & 27/2 & 23/2 \\ 15 & 0 & -9 & 29 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 10 & 8 & 0 & 0 \\ 15/2 & 4 & 9 & 0 \\ 15 & -6 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2/5 & -1/5 & 7/5 \\ 0 & 1 & 3/2 & -1/2 \\ 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

a) Use exact [rational] arithmetic to find the general decomposition $A = L.D.M$, where L is a lower triangular matrix with 1's on its leading diagonal and D is a diagonal matrix. Also find the Doolittle decomposition $A = L.U$.

b) Solve the equation $A.x = b$, where B is the column vector $(52, 96, 80, 184)^T$.

c) If this calculation were to be performed using floating point arithmetic would it be desirable to use partial pivoting? If so find the permutation matrix P. Is the permuted matrix $P.A$ strictly diagonally dominant?

[20 marks]

4. Describe the Jacobi method and the Gauss-Siedel method for finding a solution to the set of linear equations $A.\mathbf{x} = \mathbf{b}$. Under what conditions do these methods converge?

For the case when A and $\mathbf{b}$ are given by

$$A = \begin{pmatrix} 99.10 & -3.70 & 5.55 \\ 7.63 & 75.65 & -3.50 \\ -4.89 & 6.32 & 55.10 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 202.3 \\ 231.7 \\ 210.5 \end{pmatrix},$$

starting with $\mathbf{x}^0 = (2.0, 3.0, 4.0)^T$, find the first three iterates using the Jacobi method rounding to at least 5 decimal digits.

[20 marks]

5. State Gerschgorin's circle theorem on the eigenvalues of a matrix.

Use Gershgorin's theorem to show that the matrix A below has three distinct eigenvalues and find the intervals within which they lie.

$$A = \begin{pmatrix} 105.7 & 12.3 & -2.7 \\ 12.3 & 55.5 & -2.5 \\ -2.7 & -2.5 & 17.7 \end{pmatrix}$$

Describe the power method for obtaining the largest eigenvalue of a matrix and explain how it works. Choose a suitable starting vector and use three steps of the power method (with scaling) rounding to at least 5 decimal digits to obtain an approximate value for the largest eigenvalue of the matrix A.

Describe the method of inverse iteration for finding eigenvalues of a matrix. Explain how you would use it to evaluate the intermediate eigenvalue of the matrix A.

[20 marks]

6. A linear least squares system $A.\mathbf{x} = \mathbf{b}$ is given with

$$A = \begin{pmatrix} 3.7 & 2.5 \\ 4.1 & 3.3 \\ 4.9 & 3.6 \\ 5.4 & 4.2 \\ 6.2 & 4.7 \\ 7.1 & 5.2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 52.00 \\ 54.00 \\ 70.00 \\ 78.00 \\ 82.00 \\ 100.00 \end{pmatrix}$$

a) Factorize A in the form $Q.R$, rounding to at least 5 decimal digits, where Q is a 6 by 2 matrix whose columns are normalised orthogonal vectors and R is an upper triangular matrix.

b) Find the least squares solution $\mathbf{x}$.

[20 marks]