JMAT 7302
JACM 7C65
JACM 7C63

---

JMAT 7302

**Degree Master of Science in Mathematical Modelling and Scientific Computing**

**Numerical Analysis**

**Friday, 25th April 2003, 9:30 a.m. – 12:30 p.m.**

*Candidates may attempt as many questions as they wish.*

---

JACM 7C65

**Degree Master of Science in Applied & Computational Mathematics**

**Numerical Solution of Differential Equations, Numerical Linear Algebra & Finite Element Methods**

**Friday, 25th April 2003, 9:30 a.m. – 12:30 p.m.**

*Candidates may attempt as many questions as they wish.*

---

JACM 7C63

**Degree Master of Science in Applied & Computational Mathematics**

**Numerical Solution of Differential Equations & Numerical Linear Algebra**

**Friday, 25th April 2003, 9:30 a.m. – 11:30 a.m.**

*Candidates may attempt questions 1,2,3 only.*

---

Please start the answer to each question on a new page.

All questions will carry equal marks.

**Do not turn over until told that you may do so.**

## Numerical Solution of Differential Equations

### Question 1

(a) State the general form of a linear $k$-step method for the numerical solution of the initial value problem $y' = f(x, y)$, $y(x_0) = y_0$ on the mesh $\{x_j : x_j = x_0 + jh\}$ of uniform spacing $h > 0$.

(b) Define the *truncation error* of a linear $k$-step method. What is meant by saying that a linear $k$-step method is *consistent*?

(c) What is meant by saying that a linear $k$-step method is *zero-stable*? Formulate an equivalent characterisation of zero-stability in terms of the roots of a certain polynomial of degree $k$.

(d) Consider the two-parameter family of linear two-step methods defined by

$$ay_{n+2} - (1 + a)y_{n+1} + y_n = bhf_{n+2}$$

where $f_j = f(x_j, y_j)$, and $a$ and $b$ are real numbers, $ab \neq 0$. Determine the set of all values of the parameters $a$ and $b$ such that the method is zero-stable.

(e) Show that there exists a unique choice of the parameters $a$ and $b$ such that the method is second-order accurate. Is the method convergent for these values of $a$ and $b$? Justify your answer by using Dahlquist's Theorem which you should carefully state.

### Question 2

Consider the initial boundary value problem

$$\begin{aligned} \rho(x)\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, &\qquad 0 < x < 1, \quad 0 < t \leq T, \\ u(x, 0) = u_0(x), &\qquad 0 < x < 1, \\ u(0, t) = A, \quad u(1, t) = B, &\qquad t \in [0, T], \end{aligned}$$

where $A$ and $B$ are given real numbers, $T$ is a fixed positive real number, $\rho$ is a continuous function of $x$, $\rho(x) \geq c_0 > 0$, and $u_0$ is a real-valued function defined and continuous on the closed real interval $[0, 1]$.

(a) Formulate the Crank–Nicolson scheme for the numerical solution of this initial value problem on a mesh with uniform spacings $\Delta x = 1/J$ and $\Delta t = T/M$ in the $x$ and $t$ co-ordinate directions, respectively, where $J$ and $M$ are positive integers.

(b) Define the truncation error of the scheme and show that it is of size $\mathcal{O}((\Delta x)^2 + (\Delta t)^2)$ as $\Delta x, \Delta t \to 0$. [You may assume that $u$ has as many bounded and continuous partial derivatives with respect to $x$ and $t$ as are required by your proof.]

(c) Now, suppose that $\Delta t \leq c_0(\Delta x)^2$. Show that

$$\max_{0 \leq j \leq J} |U_j^m| \leq \max_{0 \leq j \leq J} |U_j^0|$$

for all $m$, $1 \leq m \leq M$.

# Numerical Linear Algebra

## Question 3

(i) Describe a method for constructing the factorisation of $A \in \mathbb{R}^{m\times n}$ as $QR$ where $Q \in \mathbb{R}^{m\times m}$ is orthogonal and $R \in \mathbb{R}^{m\times n}$ is upper triangular. Explain how this factorisation may be employed in solving the least squares problem:

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2, \quad m > n.$$

If $m = n$ what other factorisation is more commonly employed to solve the linear system $Ax = b$? Demonstrate this other factorisation, carefully indicating any additional steps that are required, for the matrix

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

(ii) if $A = M - N \in \mathbb{R}^{n\times n}$ with $A$ symmetric and positive definite and $M$ symmetric and invertible, prove that the vector sequence generated by

$$Mx^{(k)} = Nx^{(k-1)} + b, \quad k = 1, 2, \ldots$$

converges to the solution of $Ax = b$ for any initial guess $x^{(0)}$ if and only if the eigenvalues of $I - M^{-1}A$ lie inside the unit disc.

(You may wish to note that a symmetric matrix is diagonalisable and that a symmetric and positive definite matrix has a Cholesky factorisation).

## Question 4

Let $\Pi_k$ denote the set of real polynomials of degree at most $k$. Suppose that $A \in \mathbb{R}^{n\times n}$ is symmetric and positive definite.

Say why any vector in the Krylov subspace

$$\mathcal{K}_k(A,b) = \operatorname{span}\{b, Ab, A^2b, \ldots, A^{k-1}b\}$$

can be written as $p(A)b$ where $p(A)$ is a polynomial in $A$. Prove that any iterative method for the solution of $Ax = b$ with $x_0 = 0$ which has an update formula for the iterates $x_j$ of the form

$$x_k = x_{k-1} + \alpha_{k-1} r_{k-1}$$

where $r_j = b - Ax_j$, necessarily has $x_k$ in $\mathcal{K}_k(A,b)$. Deduce that $r_k = p(A)r_0$, $p \in \Pi_k, p(0) = 1$. Suppose in fact that the residuals satisfy the optimality property

$$\|r_k\|_\star = \min_{p\in\Pi_k, p(0)=1} \|p(A)r_0\|_\star$$

for some norm $\|\cdot\|_\star$. Prove that $x_\ell = x$ if $r_0$ lies in a subspace spanned by $\ell$ eigenvectors of $A$ which correspond to $\ell$ distinct eigenvalues. Is it a more or less favourable situation if the $\ell$ eigenvalues are not all distinct?

For the matrix

$$A = \begin{pmatrix} 2I_n + H(\omega) & I_n \\ I_n & 2I_n \end{pmatrix}$$

where $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix and $H(\omega)$ is a Householder matrix, show that $x_\ell = x$ where $\ell$ can be no greater than $4$ whatever the vector $b$.

# Finite Element Methods

## Question 5

(a) Given that $(a, b)$ is a bounded nonempty open interval of the real line, define the Sobolev space $\mathrm{H}^1(a, b)$ and the Sobolev norm $\|\cdot\|_{\mathrm{H}^1(a,b)}$.

(b) What is meant by saying that $u$ is a weak solution in $\mathrm{H}^1(a, b)$ of the boundary value problem

$$-u'' + (x^2 + 1)u = f(x), \quad x \in (a, b); \qquad u'(a) = 0, \quad u(b) + u'(b) = 1,$$

where $f \in \mathrm{L}_2(a, b)$?

Show that the bilinear form associated with the weak formulation of this problem is coercive on $\mathrm{H}^1(a, b)$.

(c) Consider the continuous piecewise linear basis functions $\varphi_i$, $i = 0, 1, \dots, N$, defined by $\varphi_i(x) = (1 - |x - x_i|/h)_+$, $x \in [a, b]$, on the uniform mesh of size $h = (b - a)/N$, $N \geq 2$, with mesh-points $x_i = a + ih$, $i = 0, 1, \dots, N$. Using the basis functions $\varphi_i$, $i = 0, 1, \dots, N$, define the finite element approximation of the boundary value problem and show that it has a unique solution $u_h$.

(d) Expand $u_h$ in terms of the basis functions $\varphi_i$, $i = 0, 1, \dots, N$, by writing

$$u_h(x) = \sum_{i=0}^{N} U_i \varphi_i(x)$$

where $\mathbf{U} = (U_0, U_1, \dots, U_N)^T \in \mathbb{R}^{N+1}$, to obtain a system of linear algebraic equations for the vector of unknowns $\mathbf{U}$. Show that the matrix $\mathcal{A}$ of this linear system is symmetric (*i.e.*, $\mathcal{A}^T = \mathcal{A}$) and positive definite (*i.e.*, $\mathbf{V}^T \mathcal{A} \mathbf{V} > 0$ for all $\mathbf{V} \in \mathbb{R}^{N+1}$, $\mathbf{V} \neq \mathbf{0}$).

Show also that $\|u - u_h\|_{\mathrm{H}^1(a,b)} = \mathcal{O}(h)$ as $h \to 0$.

[Any bound on the error between $u$ and its finite element interpolant $\mathcal{I}_h u$ may be used without proof, but must be stated carefully.]

## Question 6

Let $u(x,t)$ denote the solution to the initial boundary value problem

$$\begin{aligned} \frac{\partial u}{\partial t} + q(x)u &= \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t \leq T, \\ \frac{\partial u}{\partial x}(0,t) = 0, \quad u(1,t) &= 0, \qquad 0 \leq t \leq T, \\ u(x,0) &= u_0(x), \qquad 0 < x < 1, \end{aligned}$$

where $T > 0$, $u_0 \in \mathrm{L}_2(0,1)$, and $q$ is a continuous function defined on the closed interval $[0,1]$ of the real line, such that $q(x) \geq c_0 > 0$ for all $x \in [0,1]$.

(a) Construct a finite element method for the numerical solution of this problem, based on the backward Euler scheme with time step $\Delta t = T/M$, $M \geq 2$, and a piecewise linear approximation in $x$ on a uniform subdivision of spacing $h = 1/N$, $N \geq 2$, of the interval $[0,1]$, denoting by $u_h^m$ the finite element approximation to $u(\cdot, t^m)$ where $t^m = m\Delta t$, $0 \leq m \leq M$.

(b) Show that, for $0 \leq m \leq M-1$,

$$(1 + 2c_0\Delta t)\|u_h^{m+1}\|^2_{\mathrm{L}_2(0,1)} \leq \|u_h^m\|^2_{\mathrm{L}_2(0,1)},$$

where $\|\cdot\|_{\mathrm{L}_2(0,1)}$ is the $\mathrm{L}_2$-norm on the interval $(0,1)$.

Hence deduce that the method is unconditionally stable in the $\mathrm{L}_2$-norm in the sense that, for any $\Delta t$, independent of the choice of $h$,

$$\|u_h^m\|^2_{\mathrm{L}_2(0,1)} \leq (1 + 2c_0\Delta t)^{-m}\,\|u_h^0\|^2_{\mathrm{L}_2(0,1)}, \qquad 1 \leq m \leq M.$$

(c) Show that, for each $m$, $0 \leq m \leq M-1$, $u_h^{m+1}$ can be obtained from $u_h^m$ by solving a system of linear algebraic equations with a symmetric tridiagonal matrix $\mathcal{A}$ whose entries you should define in terms of the standard piecewise linear basis functions $\varphi_i$, $i = 0, 1, \ldots, N-1$. Assuming that $q(x) \equiv 1$, compute the diagonal entries of the matrix $\mathcal{A}$.