JMAT 7302
JACM 7C63

JMAT 7302

Degree Master of Science in Mathematical Modelling and Scientific Computing

Numerical Analysis

Friday, 21st April 2006, 9:30 a.m. – 12:30 p.m.

Candidates may attempt as many questions as they wish.

TRINITY TERM 2006

JACM 7C63

Degree Master of Science in Applied & Computational Mathematics

Numerical Solution of Differential Equations & Numerical Linear Algebra

Friday, 21st April 2006, 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) Suppose that $\theta \in [0,1]$ and let N be an integer, $N \geq 2$. Consider the one-step method

$$y_{n+1} = y_n + h\,f(x_n + \theta h, y_n + \theta h f(x_n, y_n)), \qquad n = 0, 1, \ldots, N-1, \qquad (\star)$$

for the numerical solution of the initial-value problem $y' = f(x, y(x))$, $y(0) = y_0$, over the uniform mesh $\{x_n : x_n = nh, n = 0, 1, \ldots, N\}$ of spacing $h = X_M/N > 0$ contained in the closed interval $[0, X_M]$.

Define the truncation error T_n of the method. Show that there exists $\theta = \theta_0 \in [0,1]$ such that $T_n = \mathcal{O}(h^2)$ as $h \to 0$, $n \to \infty$, with $nh = x_n$ fixed.

By considering $f(x,y) \equiv \lambda y$ where λ is a nonzero real number, show that there is no $r > 2$ such that $T_n = \mathcal{O}(h^r)$ as $h \to 0$, $n \to \infty$, with $nh = x_n$ fixed. Hence deduce that for $\theta = \theta_0$ the method $(\star)$ is second-order accurate.

(8+5 marks)

(b) Consider the initial-value problem

$$y' = \lambda y, \qquad y(0) = y_0, \qquad (\star\star)$$

where $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) < 0$. Determine $\lim_{x \to +\infty} y(x)$.

Apply the method $(\star)$ to the initial-value problem $(\star\star)$ over the mesh $\{x_n : x_n = nh,\ n = 0, 1, \ldots\}$, $h > 0$, with $y_n \in \mathbb{C}$ denoting the numerical approximation to $y(x_n) \in \mathbb{C}$ at $x = x_n$. Show that there exists a complex number $z = z(\lambda h, \theta)$ such that $y_{n+1} = z(\lambda h, \theta) y_n$ for $n = 0, 1, \ldots$.

Characterise in terms of $|z(\lambda h, \theta)|$ the region of absolute stability

$$\mathcal{H} = \{\lambda h \in C : \lim_{n \to +\infty} y_n = 0 \quad \text{for any } y_0 \in \mathbb{C}\}.$$

Sketch $\mathcal{H}$ in the complex plane for: (a) $\theta = 0$, (b) $\theta = \theta_0$.

(2+2+3+5 marks)

Question 2

Suppose that a is a real number, $T > 0$, and u_0 is a real-valued, four times continuously differentiable function of $x \in (-\infty, \infty)$ which is equal to zero outside a bounded subinterval of $(-\infty, \infty)$. Consider the initial value problem

$$u_t + au_x = 0, \quad -\infty < x < \infty, \quad 0 < t \leq T; \qquad u(x,0) = u_0(x), \quad -\infty < x < \infty.$$

(a) Eliminate the partial derivatives u_t, u_{tt} from a truncated Taylor series expansion of $u(x, t + \Delta t)$ about the point (x, t) using the differential equation above to show that

$$\frac{u(x,t+\Delta t) - u(x,t)}{\Delta t} + au_x(x,t) = \frac{1}{2}a^2 \Delta t\, u_{xx}(x,t) + \mathcal{O}((\Delta t)^2).$$

Explain how this identity can be used to construct the Lax–Wendroff finite difference scheme

$$\frac{U_j^{m+1} - U_j^m}{\Delta t} + a\frac{U_{j+1}^m - U_{j-1}^m}{2\Delta x} = \frac{1}{2}a^2\Delta t\,\frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta x)^2}, \quad \begin{cases} 0 \leq m \leq M-1, \\ j \in \mathbb{Z}, \end{cases}$$

$$U_j^0 = u_0(x_j), \qquad j \in \mathbb{Z},$$

where $\mathbb{Z}$ denotes the set of all integers, $\Delta x > 0$, $\Delta t = T/M$, and M is a positive integer.

(7 marks)

(b) Define the truncation error T_j^m of the scheme. Show that $\max_{0 \leq m \leq M-1} \max_{j\in\mathbb{Z}} |T_j^m| \leq K$, where $K = C_1(\Delta x)^2 + C_2(\Delta t)^2$, and C_1, C_2 are constants which you should define in terms of a and upper bounds on absolute values of certain partial derivatives of the solution u.

(6 marks)

(c) By writing

$$U_j^m = \frac{1}{2\pi}\int_{-\pi/\Delta x}^{\pi/\Delta x} \hat{U}^m(k)\mathrm{e}^{\mathrm{i}kj\,\Delta x}\mathrm{d}k,$$

where $\hat{U}^m$ denotes the semidiscrete Fourier transform of the mesh-function $j \in \mathbb{Z} \mapsto U_j^m$ defined by the Lax–Wendroff scheme, show that

$$\hat{U}^{m+1}(k) = \lambda(k)\hat{U}^m(k),$$

where $\lambda(k)$ is a function that you should define on the interval $[-\pi/\Delta x, \pi/\Delta x]$ in terms of CFL number $\mu = a\Delta t/\Delta x$, the wave number k and the spatial mesh-size Δx.

Show that if $|\mu| \leq 1$ then the Lax–Wendroff scheme is stable in the ℓ_2 norm.

(6+6 marks)

Numerical Linear Algebra

Question 3

Gram–Schmidt factorization (for all students)

Throughout this problem, A is an $m \times n$ real matrix with $m \geq n$ and rank n, and $A = QR$ is a reduced (or "skinny" or "thin" or "economy sized") QR factorization of A.

(a) State the dimensions and orthogonality/sparsity properties of Q and R. **(2 marks)**

(b) One way to compute Q and R is by Gram–Schmidt orthogonalization, in which a sequence of orthonormal vectors $q_1, q_2, \ldots$ is constructed that span the same sequences of nested spaces as the column vectors $a_1, a_2, \ldots$ of A. Give formulae showing how the following quantities can be computed in sequence by a Gram–Schmidt process:

$$r_{11},\ q_1,\ r_{12},\ r_{22},\ q_2,\ r_{13},\ r_{23},\ r_{33},\ q_3.$$

Then write down the general algorithm (e.g. in "pseudocode" or Matlab) to compute all of Q and R. (You do not have to worry about whether your algorithm corresponds to "classical" or "modified" Gram–Schmidt.) **(9 marks)**

(c) Show that in any QR factorization of A, whether computed by Gram–Schmidt factorization or not, $r_{jj} \neq 0$ for each j, $1 \leq j \leq n$. Show using (b) that there exists a QR factorization of A with $r_{jj} > 0$ for each j, and that it is unique.

(7 marks)

(d) Show that the singular values of A are the same as those of R. For the special case $m = n = 2$, sketch the behaviour of a "typical" matrix A as a mapping from $\mathbf{R}^2$ to $\mathbf{R}^2$, showing where the axes $(1,0)^T$ and $(0,1)^T$ might be mapped by A and where the singular values σ_1 and σ_2 appear in the picture. If $A = QR$, what will the corresponding sketch for R look like? **(7 marks)**

Question 4

The cosine of a matrix. (for students in Mathematical Modelling and Scientific Computing)

Let A be an $n \times n$ real matrix. In the study of the second-order differential equation $u'' = Au$, where u is a time-dependent n-vector, it is sometimes useful to consider the *cosine* of A, $\cos(A)$.

(a) One way to define $\cos(A)$ is by the usual Taylor series for $\cos(x)$, except with x replaced by A. Write down this series definition for $\cos(A)$. Prove by using the fact that $\|A\| < \infty$ (the choice of norm doesn't matter) that this series converges for any A.

(6 marks)

(b) Suppose that $\|A\| \leq 1$. Exactly which powers of A will be needed to determine $\cos(A)$ to an accuracy of better than 10^{-10}, using the series definition?

(4 marks)

(c) Show from the result of (b) that for A with $\|A\| \leq 1$, $\cos(A)$ can be computed to 10-digit accuracy by means of six matrix multiplications together with various less expensive operations. To leading order in n as $n \to \infty$, what floating point operation count does this correspond to?

(6 marks)

(d) Define an *eigenvalue decomposition* of a square matrix A, and state what special form the eigenvalue decomposition can take if $A = A^T$.

(4 marks)

(e) Assuming $A = A^T$, explain how the eigenvalue decomposition can be used to compute $\cos(A)$. If approximately $9n^3$ flops are required to compute the decomposition, does the operation count for this method come out much better, much worse, or about the same as that of parts (a)–(c)? (5 marks)

Finite Element Methods

Question 5

(a) Suppose that $f \in \mathrm{L}^2(1,2)$. State what is meant by saying that u is a weak solution in $\mathrm{H}^1(1,2)$ of the boundary value problem

$$-(xu')' + u' + u = f(x), \quad x \in (1,2); \qquad u(1) - u'(1) = 0, \quad u(2) + 2u'(2) = 0.$$

(3 marks)

(b) Show that the bilinear form associated with the weak formulation of this problem is coercive on $\mathrm{H}^1(1,2)$.

(3 marks)

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

(3 marks)

(d) Expand u_h in terms of the basis functions φ_j, $j = 0, 1, \ldots, N$, by writing

$$u_h(x) = \sum_{j=0}^{N} U_j \varphi_j(x)$$

where $\mathbf{U} = (U_0, U_1, \ldots, U_N)^\top \in \mathbb{R}^{N+1}$, 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 positive definite (i.e. $\mathbf{V}^\top \mathcal{A} \mathbf{V} > 0$ for all $\mathbf{V} \in \mathbb{R}^{N+1}$, $\mathbf{V} \neq \mathbf{0}$). Is the matrix $\mathcal{A}$ symmetric (i.e. $\mathcal{A}^\top = \mathcal{A}$)? Justify your answer.

(3+3+3 marks)

(e) Show that $\|u - u_h\|_{\mathrm{H}^1(1,2)} = \mathcal{O}(h)$ as $h \to 0$.

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

(7 marks)

Question 6

Suppose that Ω is a bounded polygonal domain in $\mathbb{R}^2$ with boundary Γ which has been subdivided into two disjoint nonempty parts Γ_D and Γ_N. It is assumed that each of Γ_D and Γ_N is the union of certain edges of the polygon Γ. Let $f \in \mathrm{L}^2(\Omega)$ and define $\mathrm{H}^1_\mathrm{D}(\Omega) = \{v \in \mathrm{H}^1(\Omega) : v|_{\Gamma_\mathrm{D}} = 0\}$. Consider the quadratic functional $J : v \in \mathrm{H}^1_\mathrm{D}(\Omega) \mapsto J(v) \in \mathbb{R}$ defined by

$$J(v) = \frac{1}{2}\int_\Omega \left(|\nabla v|^2 + v^2\right) \mathrm{d}\Omega - \int_\Omega f v \, \mathrm{d}\Omega.$$

(a) Show that if $u \in \mathrm{H}^1_\mathrm{D}(\Omega)$ is such that $J(u) \leq J(v)$ for all $v \in \mathrm{H}^1_\mathrm{D}(\Omega)$, then there exist a bilinear functional $a(\cdot,\cdot)$ defined on $\mathrm{H}^1_\mathrm{D}(\Omega) \times \mathrm{H}^1_\mathrm{D}(\Omega)$ and a linear functional $\ell(\cdot)$ defined on $\mathrm{H}^1_\mathrm{D}(\Omega)$ such that

$$a(u,v) = \ell(v) \qquad \forall v \in \mathrm{H}^1_\mathrm{D}(\Omega). \tag{P}$$

(7 marks)

(b) Show that (P) is the weak formulation of the elliptic boundary-value problem

$$-\nabla^2 u + u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \Gamma_\mathrm{D}, \qquad \frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma_\mathrm{N},$$

where $\frac{\partial u}{\partial n} = \nabla u \cdot \mathbf{n}$ and $\mathbf{n}$ denotes the unit outward normal vector to Γ. (6 marks)

(c) Suppose that Ω is the unit square $(0,1) \times (0,1)$ with boundary Γ which has been subdivided into $\Gamma_\mathrm{N} = \{(x,1) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$ and $\Gamma_\mathrm{D} = \Gamma \backslash \Gamma_\mathrm{N}$, and let $\mathcal{T}_h$ be a triangulation of Ω constructed from a uniform square mesh of spacing $h = 1/N$ by subdividing each mesh square by the diagonal of negative slope. Formulate the piecewise linear finite element approximation (P_h) of problem (P) on the triangulation $\mathcal{T}_h$. Show that P_h has a unique solution u_h and that $J(u) \leq J(u_h) \leq J(v_h)$ for any continuous piecewise linear function v_h defined on the triangulation $\mathcal{T}_h$.

(3+3+3+3 marks)