JMAT 7302
JACM 7302
JACM 7C63
JACM 7C65

M.Sc. in Mathematical Modelling and Scientific Computing

Paper B (Numerical Analysis)

Friday 19 April, 2002, 9.30 a.m. – 12.30 p.m.

Candidates may attempt as many questions as they wish. All questions will carry equal marks.

Do not turn this page until you are told that you may do so

Numerical Linear Algebra

1. What is a Householder matrix? Show that any column vector may have all but its first entry mapped to zero when it is premultiplied by an appropriate Householder matrix.

What is an upper Hessenberg matrix? Explain how any real square matrix A may be reduced to an upper Hessenberg matrix B by orthogonal similarity transformation. If A is symmetric, what is special about the form of B?

An upper Hessenberg matrix $B \in \mathbb{R}^{n\times n}$, $B = \{b_{i,j} : i, j = 1, \dots, n\}$, is said to be *reduced* if $b_{i,i-1} = 0$ for some i; otherwise it is called *unreduced*. Show that if B is reduced, then the eigenvalues of B are independent of the $(i-1)*(n-i+1)$, generally non-zero, entries of B.

Show that every right eigenvector, $x = (x_1, x_2, \dots, x_n)^T$, of an unreduced n-by-n upper Hessenberg matrix satisfies $x_n \neq 0$. Further, show that if an upper Hessenberg matrix has a multiple eigenvalue with at least two corresponding eigenvectors, then it must be reduced.

2. (a) Show how the iterates $\{x_k\}$ of a simple iteration

$$x_{k+1} = Tx_k + c\,, \qquad k = 0, 1, 2, \dots\,,$$

with starting guess x_0, for the solution of a linear system of equations $Ax = b$, can be linearly combined to give a polynomial iterative method.

What criteria determine a good choice of polynomials if T is diagonalisable, and why are the Chebyshev polynomials scaled and shifted onto an approriate interval a good choice? (You may quote without proof any results that you need.)

(b) Show that solving $Ax = b$ for $x \in \mathbb{R}^n$, given $b \in \mathbb{R}^n$ and a real symmetric and positive definite matrix $A \in \mathbb{R}^{n\times n}$, is equivalent to the problem of minimising the functional

$$\Phi(y) = \frac{1}{2}y^T Ay - y^T b$$

over $y \in \mathbb{R}^n$.

If $r = b - Ay$ is the residual for a particular $y \in \mathbb{R}^n$, show that $\Phi(y+\alpha r)$ is minimised when $\alpha = r^T r / r^T Ar$.

Quoting any results that you need, briefly describe the steps in the Conjugate Gradient method. (You do not need to give any formulae.) What is the role of *preconditioning* and what properties should an effective preconditioner have?

Numerical Solution of Differential Equations

3. Consider the initial value problem $y' = f(x, y)$, $y(x_0) = y_0$, on the closed interval $[x_0, x_N]$ of the real line, where f is a smooth function of its arguments and y_0 is a given real number, and the one-step method

$$y_{n+1} = y_n + \alpha h f(x_n, y_n) + \beta h f(x_n + \gamma h, y_n + \gamma h f(x_n, y_n))$$

for the numerical solution of this initial value problem on the mesh

$$\{x_n \,:\, x_n = x_0 + nh\,,\ n = 0, \ldots, N\}$$

of spacing h, $h > 0$, where α, β and γ are real parameters.

(i) Define the truncation error T_n of this method. Show that

$$T_n = (1 - \alpha - \beta)y'(x_n) + \frac{1}{2}(1 - 2\beta\gamma)hy''(x_n) + \mathcal{O}(h^2)\,, \quad \text{as } h \to 0\,.$$

Hence deduce that the method is consistent if, and only if, $\alpha + \beta = 1$.

Show further that there exists a choice of the parameters α, β and γ such that the method is second-order accurate.

(ii) By considering the method applied to the problem $y' = y$, $y(0) = 1$, show that there is no choice of α, β and γ such that the order of accuracy exceeds 2.

(iii) Suppose that a second–order method of the above form is applied to the initial value problem $y' = \lambda y$, $y(0) = 1$, where $\lambda < 0$ is a fixed constant. The solution $y(x) = \mathrm{e}^{\lambda x}$ to this initial value problem is monotonic decreasing. Find the set of all $h > 0$ such that the corresponding sequence of numerical approximations $(y_n)_{n=0,\ldots,N}$, with $y_0 = 1$, is monotonic decreasing. Discuss the practical implications of the resulting restriction on h when $\lambda \ll -1$.

4. Consider the system of linear equations

$$-a_jU_{j-1} + b_jU_j - c_jU_{j+1} = d_j\,, \qquad j = 1, \ldots, J-1\,,$$

with $J \geq 2$, and

$$U_0 = 0\,, \qquad U_J = 0\,,$$

where $a_j > 0$, $b_j > 0$, $c_j > 0$ and $b_j > a_j + c_j$ for all j.

(i) Show that

$$U_j = e_jU_{j+1} + f_j\,, \qquad j = J-1, J-2, \ldots, 1\,,$$

where

$$e_j = \frac{c_j}{b_j - a_je_{j-1}}\,, \qquad f_j = \frac{d_j + a_jf_{j-1}}{b_j - a_je_{j-1}}\,, \qquad j = 1, 2, \ldots, J-1\,,$$

with $e_0 = 0$ and $f_0 = 0$.

How would you use these formulae to solve the system of linear equations?

(ii) Show by induction that $0 < e_j < 1$ for $j = 1, 2, \ldots, J-1$.

Comment on the practical significance of this result in relation to the sensitivity of the algorithm to rounding errors.

(iii) Consider the heat equation $u_t = u_{xx} + u_{yy}$ on the unit square $\Omega = (0,1)^2$, and $t \in (0, T]$, subject to homogeneous Dirichlet boundary condition and the initial condition $u(x, y, 0) = u_0(x, y)$.

Set up an ADI scheme, based on the Crank–Nicolson method, for the numerical solution of this initial boundary value problem, on a uniform spatial mesh of spacings $\Delta x = 1/I$ and $\Delta y = 1/J$ in the x and y co-ordinate directions, respectively; $I, J \geq 2$.

Explain how the results of part (i) of this question can be exploited in this scheme.

Finite Element Methods

5. Consider the two-point boundary value problem

$$-u'' + c(x)u = f(x)\,, \quad x \in (a,b)\,, \qquad u(a) = 0\,, \quad u'(b) = 0$$

on the nonempty open interval (a,b) of the real line, where c is a bounded nonnegative function defined on (a,b) and f is a real-valued square-integrable function defined on (a,b).

(i) Show, using the Lax–Milgram Theorem, that the boundary value problem has a unique weak solution in a subspace $\mathcal{H}$ of $H^1(a,b)$ that you should carefully define.

(ii) Consider a nonuniform mesh $a = x_0 < x_1 < \ldots < x_{N-1} < x_N = b$ on the interval $[a,b]$. Using continuous piecewise linear basis functions defined on this mesh, formulate the corresponding finite element approximation of the boundary value problem, and show that this has a unique solution.

(iii) Let S_h be the N-dimensional linear subspace of $\mathcal{H}$ consisting of continuous piecewise linear functions defined on the mesh $a = x_0 < x_1 < \ldots < x_{N-1} < x_N = b$, with $h = \max_{1\le i\le N}(x_i - x_{i-1})$. Show that there exists a positive constant K such that

$$\|u - u_h\|_{H^1(a,b)} \le K \min_{v_h \in S_h} \|u - v_h\|_{H^1(a,b)}\,.$$

Hence deduce that

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

[Any bound on the error between u and its continuous piecewise linear finite element interpolant $I_h u \in S_h$ may be used without proof.]

6. Suppose that

$$\Omega = \{(x, y) \in \mathbb{R}^2 \,:\, x > 0\,,\; y > 0\,,\; x + y < 1\}\,,$$

and let $f \in L_2(\Omega)$. Consider Poisson's equation $-\nabla^2 u = f$ on Ω subject to homogeneous Dirichlet boundary condition on $\partial\Omega$.

(i) A triangulation of Ω is constructed by first considering a square mesh on Ω with horizontal and vertical meshlines, of spacing $h = 1/N$, $N \geq 2$, and then subdividing each mesh-square by the diagonal of negative slope. Define carefully the continuous piecewise linear finite element approximation u_h to u on this triangulation.

(ii) Show that the finite element method from (i) can be reformulated as a system of linear algebraic equations with a symmetric positive definite square matrix $\mathbf{A}$ that has $(N-1)(N-2)/2$ rows and columns. Show further that $\mathbf{A}$ is a sparse matrix whose diagonal entries are equal to 4.

(iii) Define the continuous piecewise linear finite element interpolant $I_h u$ of u on the mesh from (i). How much reduction in the error $\|u - u_h\|_{L_2(\Omega)}$ do you expect to observe when the mesh-size h is halved? Justify your answer.

[Bounds on the error between a function and its finite element interpolant may be used without proof.

You may assume that if $-\nabla^2 z = g$ in Ω with $g \in L_2(\Omega)$ and $z|_{\partial\Omega} = 0$, then $z \in H^2(\Omega)$ and $\|z\|_{H^2(\Omega)} \leq Const. \|g\|_{L_2(\Omega)}$.]