M.Sc. in Mathematical Modelling and Numerical Analysis

Paper B (Numerical Analysis)

Friday 23 April, 1999, 9.30 a.m. – 12.30 p.m.

Candidate's 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. (a) What is the Gaussian Elimination (GE) algorithm for the solution of $Ax = b$? Assuming that there is no failure in GE, show that it is mathematically equivalent to a factorisation of the coefficient matrix as $A = LU$ where L is lower triangular with 1's on the diagonal and U is upper triangular.

If a sequence of matrices A_k is constructed using LU factorisation as above by $A_1 = A$

$$\left.\begin{array}{ll} \text{Factor} & A_k = L_k U_k \\ \text{Form} & A_{k+1} = U_k L_k \end{array}\right\} \text{for} \quad k = 1, 2, \ldots \quad (*),$$

show that all the matrices A_k, $(k = 1, 2, \ldots)$ are similar. If, further, as $k \to \infty$, A_k tends to a triangular matrix, of what use is the algorithm defined by (*)?

(b) What is Jacobi iteration for a linear system of equations

$$\sum_{j=1}^{n} a_{i,j} x_j = b_i, \quad i = 1, \ldots, n?$$

Show that the xector $x \in \mathbb{R}^n$ with components

$$x_j = 2^{j/2} \sin \frac{rj\pi}{n+1}$$

is an eigenvector of the tridiagonal matrix

$$A = \begin{pmatrix} 3 & -1 & & \\ -2 & 3 & -1 & \\ & \ddots & \ddots & -1 \\ & & -2 & 3 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

for any $r \in \{1, 2, \ldots, n\}$. What is the corresponding eigenvalue? Will Jacobi iteration converge for a linear system with A as coefficient matrix?

Numerical Solution of Ordinary Differential Equations

2. The linear multistep method

$$y_{n+1} - (1+a)y_n + ay_{n-1} = h(\alpha f_{n+1} + \beta f_n + \gamma f_{n-1})$$

is used to find an approximate solution of the first-order ordinary differential equation $y' = f(x, y)$. For what range of values of a is the method zero-stable?

In the case $a = 0$, find values α, β, γ such that the truncation error is of highest possible order if

(i) we require $\alpha = 0$,

or if

(ii) we do not require $\alpha = 0$.

What is the significance of taking $\alpha = 0$ in the practical implementation of the method? How may the methods (i) and (ii) be combined to produce an effective method and what is the order of the truncation error?

For what class of differential equations are the above methods likely to prove ineffective and why? Suggest a method which would be effective.

Numerical Solution of Partial Differential Equations

3. (a) Explain in a general way, without writing any formulas or worrying about technical details, the distinction between hyperbolic and parabolic partial differential equations. Also explain the distinction between explicit and implicit finite difference discretisations. Finally, explain why it is that often, explicit finite difference discretisations are suitable for hyperbolic problems whereas implicit ones are needed for parabolic problems.

(b) Consider the PDE

$$u_t = u_x + Au_{xx}, \qquad x \in R,$$

with the usual notation $u_x = \partial u/\partial x$, etc. Here A is a fixed constant. Write down the obvious explicit discretisation of this equation on a regular grid with space step Δx and time step Δt, using three points at the current space step. (The grid is infinite, so in practice one would need boundary conditions, but ignore this fact.) Insert an appropriate Fourier mode with wave number k to obtain a formula for the amplification factor $\lambda = \lambda(k, \Delta x, \Delta t, A)$ by which a Fourier mode is multiplied from one time step to the next.

(c) Sketch the curve in the complex λ-plane described by $\lambda(k, \Delta x, \Delta t, A)$ as k ranges over all possible values, for fixed Δx, Δt, and A. What condition on this curve is necessary and sufficient for the discretisation to be stable? Derive algebraic stability conditions on Δt and Δx, and discuss their dependence on A in the light of part (a).

4. (a) Suppose we wish to solve Poisson's equation on the unit cube,

$$\nabla^2 u = f(x, y, z), \qquad 0 < x, y, z < 1$$

for the unknown function $u = u(x, y, z)$ with Dirichlet boundary conditions

$$u(x, y, z) = g(x, y, z) \qquad \text{on boundary.}$$

Describe carefully what matrix problem $Av = b$ you get when this problem is approximated by the standard centred second-order-accurate finite difference discretisation on the regular mesh with $\Delta x = \Delta y = \Delta z = h = 1/n$. What is b and what is v? Describe the structure of A carefully, explaining exactly of what blocks of what dimensions A is composed, what these blocks are composed of, and so on. Is A symmetric or positive definite?

(b) Suppose one solved this problem $Av = b$ by Gaussian elimination or Cholesky factorisation, taking no advantage of the sparsity of A. For A of dimension $n \times n$, approximately how many floating-point operations would be required for this solution, and how much storage? Now take $n = 100$ and assume one is working in standard IEEE double precision arithmetic on a machine that performs at the rate of 10 megaflops per second. How many megabytes of storage and how many minutes or hours of computing time would be required? Don't worry about getting the answers exactly right, but try to get them within a factor of two.

(c) In practice, very different methods would be used to solve this problem $Av = b$. Among the possibilities are banded Gaussian elimination or Cholesky factorisation, SOR iteration, nested dissection, fast Poisson solvers, multigrid, conjugate gradients, and conjugate gradients with a preconditioner. Pick one of these methods and describe it in some detail for this problem. How much would you expect it to improve the figures of part (b)?

Finite Element Methods

5. Consider the boundary value problem

$$-u'' + c(x)u = f(x) \quad \text{for } x \in (0,1), \qquad u(0) = 0, \quad u(1) + u'(1) = 0,$$

where c and f are continuous functions defined on the interval $[0,1]$ and $c(x) \geq 0$ for all x in $[0,1]$.

State the weak formulation of this boundary value problem.

Using continuous piecewise linear basis functions on a uniform mesh of size $h = 1/N$, $N \geq 2$, formulate the finite element approximation of the boundary value problem; show that this has a unique solution u_h.

Expand u_h in terms of the finite element basis functions ϕ_j, $j = 1, \ldots, N$, where $\phi_j(x) = (1 - |x - x_j|/h)_+$, by writing

$$u_h(x) = \sum_{j=1}^{N} U_j \phi_j(x)$$

to obtain a system of linear equations for the vector of unknowns $(U_1, \ldots, U_N)^T$. Comment on the sparsity structure of the matrix of this system.

Suppose that $c(x) \equiv 0$, $f(x) \equiv 1$ and $h = 1/3$. Solve the resulting system of linear equations and compare the corresponding numerical solution $u_h(x)$ with the analytical solution $u(x)$ at the mesh points x_1, x_2 and x_3.

6. Suppose that Ω is a convex polygonal domain in the plane and $\mathcal{T} = \{K\}$ is a triangulation of Ω. Consider a triangle $K \in \mathcal{T}$ whose vertices have position vectors $\mathbf{r}_i = (x_i, y_i)$, $i = 1, 2, 3$, labelled in an anti-clockwise direction. Given that $\mathbf{r} = (x, y)$ is a point in K, write $\mathbf{r} = \mathbf{r}(\xi, \eta) = (1 - \xi - \eta)\mathbf{r}_1 + \xi\mathbf{r}_2 + \eta\mathbf{r}_3$ with $0 \le \xi \le 1$, $0 \le \eta \le 1 - \xi$.

Show that any linear polynomial $v_h(x, y)$ defined on K may be expressed as

$$v_h(x, y) = v_h(\mathbf{r}(\xi, \eta)) = \sum_{i=1}^{3} V_i^K \hat{\psi}_i(\xi, \eta),$$

where $V_i^K = v_h(x_i, y_i)$, and $\{\hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3\}$ is a suitable nodal basis for the set of linear polynomials in ξ and η on the canonical triangle $\hat{K} = \{(\xi, \eta) \,:\, 0 \le \xi \le 1, \quad 0 \le \eta \le 1-\xi\}$. Show further that

$$\frac{\partial}{\partial \xi} v_h(\mathbf{r}(\xi, \eta)) = V_2^K - V_1^K, \qquad \frac{\partial}{\partial \eta} v_h(\mathbf{r}(\xi, \eta)) = V_3^K - V_1^K.$$

Deduce that

$$\int_K |\nabla v_h(x, y)|^2 \mathrm{d}x\mathrm{d}y = [V_1^K, V_2^K, V_3^K] A^K \begin{bmatrix} V_1^K \\ V_2^K \\ V_3^K \end{bmatrix},$$

where A^K is the symmetric 3×3 element stiffness matrix given as

$$A^K = \frac{1}{4A_{123}} \begin{bmatrix} a_{11}^K & a_{12}^K & a_{13}^K \\ & a_{22}^K & a_{23}^K \\ & & a_{33}^K \end{bmatrix}$$

and A_{123} is the area of K; you need only calculate explicitly one of the entries of the matrix, say a_{11}^K, in terms of the position vectors $\mathbf{r}_i$.

Show that if M denotes the number of triangles in T, there is a matrix L^K of size $M \times 3$ such that the global stiffness matrix A can be expressed as

$$A = \sum_{K \in \mathcal{T}} L^K A^K (L^K)^T,$$

where $(L^K)^T$ is the transpose of L^K.

Explain briefly the relevance of this result in the assembly of the global stiffness matrix for the piecewise linear finite element approximation of Poisson's equation $-\Delta u = f(x, y)$ in Ω subject to a homogeneous Dirichlet boundary condition on $\partial\Omega$.