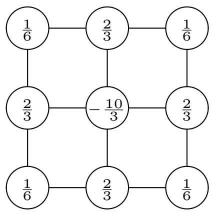
Dr H. Fawzi

Mathematical Tripos Part II: Michaelmas Term 2022

Numerical Analysis – Lecture 2

Approximation 1.6 The *nine-point method*:



$$u_{i,j} \quad = \quad h^2 f_{i,j}$$

As such, it again produces error of $\mathcal{O}(h^4)$. However, this can be remedied by a clever trick of adding the term $\frac{1}{12}h^4\nabla^2 f$ to the right-hand side, with the 5-point approximation to $h^2\nabla^2 f$, which increases the order to $\mathcal{O}(h^6)$ (see Exercise 1).

Problem 1.8 Finite-difference discretization of $\nabla^2 u = f$ replaces the PDE by a large system of linear equations. In the sequel we pay special attention to the *five-point formula,* which results in the approximation

$$h^2\nabla^2 u(x,y) \approx u(x-h,y) + u(x+h,y) + u(x,y-h) + u(x,y+h) - 4u(x,y)\,. \tag{1.5}$$

For the sake of simplicity, we restrict our attention to the important case of Ω being a *unit square*, where $h = \frac{1}{m+1}$ for some positive integer m. Thus, we estimate the m^2 unknown function values $u(ih,jh)_{i,j=1}^m$ (where $(ih,jh) \in \Omega$) by letting the right-hand side of (1.5) equal $h^2 f(ih,jh)$ at each value of i and j. This yields an $n \times n$ system of linear equations with $n = m^2$ unknowns $u_{i,j}$:

$$u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j} = h^2 f(ih,jh)\,. \tag{1.6}$$

(Note that when i or j is equal to 1 or m, then the values $u_{0,j}$, $u_{i,0}$ or $u_{i,m+1}$, $u_{m+1,j}$ are known boundary values and they should be moved to the right-hand side, thus leaving fewer unknowns on the left.) Having ordered grid points, we can write (1.6) as a linear system, say

$$A\boldsymbol{u} = \boldsymbol{b}\,.$$

Our present concern is to prove that, as $h \to 0$, the numerical solution (1.6) tends to the exact solution of the Poisson equation $\nabla^2 u = f$ (with appropriate Dirichlet boundary conditions).

Example 1.9 (Natural ordering) The way the matrix A of this system looks depends of course on the way how the grid points (ih,jh) are being assembled in the one-dimensional array. In the *natural ordering*, when the grid points are arranged by columns, A is the following block tridiagonal matrix:

$$A = \begin{bmatrix} B & I & & & \\ I & B & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & B & I \\ & & & I & B \end{bmatrix}, \qquad B = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ & & & 1 & -4 \end{bmatrix}.$$

Matlab demo (natural ordering animation): `www.damtp.cam.ac.uk/user/hf323/M21-II-NA/demos/partii.php`.

Before heading on let us prove the following simple but useful theorem whose importance will become apparent in the course of the lecture.

Theorem 1.10 (Gershgorin theorem) *All eigenvalues of an $n \times n$ matrix A are contained in the union of the Gershgorin discs in the complex plane:*

$$\sigma(A) \subset \cup_{i=1}^n \Gamma_i\,, \qquad \Gamma_i := \{z \in \mathbb{C} : |z - a_{ii}| \le r_i\}, \qquad r_i := \textstyle\sum_{j\ne i} |a_{ij}|\,.$$

Proof. For any matrix A, if $A\boldsymbol{x} = \lambda\boldsymbol{x}$ and $|x_i| = \max|x_j|$, then the ith equation of the relation $A\boldsymbol{x} = \lambda\boldsymbol{x}$ gives

$$|\lambda - a_{ii}| \cdot |x_i| = \Big|\sum_{j\neq i} a_{ij}x_j\Big| \le \sum_{j\neq i}|a_{ij}||x_j| \le |x_i|\sum_{j\neq i}|a_{ij}| =: |x_i|\, r_i\,,$$

and after dividing by $|x_i|$ we obtain $|\lambda - a_{ii}| \le r_i$. So, for any eigenvalue λ of A, the inequality $|\lambda - a_{ii}| \le r_i$ is valid for at least one value of i, hence the theorem. □

Lemma 1.11 *For any ordering of the grid points, the matrix A of the system (1.6) is symmetric and negative definite.*

Proof. Equation (1.6) implies that if $a_{ij} \neq 0$ for $i \neq j$, then the i-th and j-th points of the grid (ph, qh), are nearest neighbours. Hence $a_{ij} \neq 0$ implies $a_{ij} = a_{ji} = 1$, which proves the symmetry of A. Therefore A has real eigenvalues and eigenvectors.

It remains to prove that all the eigenvalues are negative. The arguments are parallel to the proof of Gershgorin theorem. Let $A\boldsymbol{x} = \lambda\boldsymbol{x}$, and let i be an integer such that $|x_i| = \max|x_j|$. With such an i we address the following identity (which is a reordering of the equation $(A\boldsymbol{x})_i = \lambda x_i$):

$$\underbrace{\big|(\lambda - a_{ii})\,x_i\big|}_{|\lambda+4|\,|x_i|} = \underbrace{\big|\,\textstyle\sum_{j\neq i}^{n} a_{ij}x_j\,\big|}_{\le 4\,|x_i|}\,. \tag{1.7}$$

Here $a_{ii} = -4$ and $a_{ij} \in \{0,1\}$ for $j \neq i$, with at most four nonzero elements on the right-hand side. It is seen that the case $\lambda > 0$ is impossible. Assuming $\lambda = 0$, we obtain $|x_j| = |x_i|$ whenever $a_{ij} = 1$, so we can alter the value of i in (1.7) to any of such j and repeat the same arguments. Thus, the modulus of every component of $\boldsymbol{x}$ would be $|x_i|$, but then the equations (1.7) that occur at the boundary of the grid and have fewer than four off-diagonal terms (see (1.6)) could not be true. Hence, $\lambda = 0$ is impossible too, hence $\lambda < 0$ which proves that A is negative definite. □

Proposition 1.12 *The eigenvalues of the matrix A are*

$$\lambda_{k,\ell} = -4\Big(\sin^2\frac{k\pi h}{2} + \sin^2\frac{\ell\pi h}{2}\Big), \qquad h = \frac{1}{m+1}\,, \qquad k,\ell = 1...m.$$

Proof. Let us show that, for every pair (k,ℓ), the vectors

$$v = (v_{i,j}), \quad v_{i,j} = \sin ix \sin jy, \quad \text{where} \quad x = k\pi h, \quad y = \ell\pi h,$$

are the eigenvectors of A. Indeed, for $i, j = 1...m$, we have

$$\begin{aligned}(Av)_{i,j} &= \sin(jy)\,\big[\sin(ix-x) - 2\sin(ix) + \sin(ix+x)\big] \\ &\quad + \sin(ix)\,\big[\sin(jy-y) - 2\sin(jy) + \sin(jy+y)\big] \\ &= \sin(jy)\sin(ix)\big[2\cos x - 2\big] + \sin(ix)\sin(jy)\big[2\cos y - 2\big] \;=\; \lambda v_{i,j}\,.\end{aligned}$$

Note that the terms $u_{i\pm1,j}, u_{i,j\pm1}$ do not appear in (1.6) for $i, j = 1$ or $i, j = m$, respectively, therefore (for such i, j) we should have dropped the corresponding components from above equation, but they are equal to zero because $\sin(i-1)x = 0$ for $i = 1$, while $\sin(i+1)x = 0$ for $i = m$, since $x = \frac{k\pi}{m+1}$. Thus, the eigenvalues are

$$\lambda_{k,\ell} = \big[2\cos x - 2\big] + \big[2\cos y - 2\big] = -4\Big(\sin^2\frac{x}{2} + \sin^2\frac{y}{2}\Big) = -4\Big(\sin^2\frac{k\pi h}{2} + \sin^2\frac{\ell\pi h}{2}\Big). \quad \square$$

Remark 1.13 As a matter of independent mathematical interest, note that for $1 \le k, \ell \ll m$ we have $\sin x \approx x$, hence the eigenvalues for the discretized Laplacian ∇_h^2 are

$$\frac{\lambda_{k,\ell}}{h^2} \approx -\frac{4}{h^2}\left[\frac{k^2\pi^2h^2}{4} + \frac{\ell^2\pi^2h^2}{4}\right] = -(k^2+\ell^2)\pi^2\,.$$

Now, recall (e.g. from the solution of the Poisson equation in a square by separation of variables in Maths Methods) that the *exact* eigenvalues of ∇^2 (in the unit square) are $-(k^2+\ell^2)\pi^2$, $k, \ell \in \mathbb{N}$, with the corresponding eigenfunctions $V_{k,\ell}(x,y) = \sin k\pi x \sin \ell\pi y$. So, the eigenvectors of the discretized ∇_h^2 are the values of $V_{k,\ell}(x,y)$ on the grid-points, and the eigenvalues of ∇_h^2 approximate those for continuous case.