Mathematical Tripos Part II: Michaelmas Term 2020

Numerical Analysis – Lecture 6

Definition 2.4 (Normal matrices) We say that a matrix A is *normal* if $A = QD\bar{Q}^T$, where D is a (complex) diagonal matrix and Q is a unitary matrix (such that $Q\bar{Q}^T = I$, where the bar in $\bar{Q}$ means complex conjugation). In other words, a matrix is normal if it has a complete set of orthonormal eigenvectors.

Examples of the real normal matrices, besides the familiar symmetric matrices ($A = A^T$), inlclude also the matrices which are skew-symmetric ($A = -A^T$), and more generally the matrices with skew-symmetric off-diagonal part.

Proposition 2.5 *If A is normal, then $\|A\| = \rho(A)$.*

Proof. Let $\boldsymbol{u}$ be any vector (complex-valued as well). We can expand it in the basis of the orthonormal eigenvectors $\boldsymbol{u} = \sum_{i=1}^n a_i \boldsymbol{q}_i$. Then $A\boldsymbol{u} = \sum_{i=1}^n \lambda_i a_i \boldsymbol{q}_i$, and since $\boldsymbol{q}_i$ are orthonormal, we obtain

$$\|A\|_2 := \sup_{\boldsymbol{u}} \frac{\|A\boldsymbol{u}\|_2}{\|\boldsymbol{u}\|_2} = \sup_{a_i} \frac{\{\sum_{i=1}^n |\lambda_i a_i|^2\}^{1/2}}{\{\sum_{i=1}^n |a_i|^2\}^{1/2}} = |\lambda_{\max}| \,.$$

Remark 2.6 More generally, one can prove that, for any matrix A, we have $\|A\|_2 = [\rho(A\bar{A}^T)]^{1/2}$, and the previous result for normal matrices can be deduced from that formula.

Example 2.7 (Crank–Nicolson method for diffusion equation) Let

$$u_m^{n+1} - \tfrac{1}{2}\mu(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n + \tfrac{1}{2}\mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n), \qquad m = 1...M.$$

Then $B\boldsymbol{u}^{n+1} = C\boldsymbol{u}^n$, where the matrices B and C are Toeplitz symmetric tridiagonal (TST),

$$\boldsymbol{u}^{n+1} = B^{-1}C\boldsymbol{u}^n, \qquad \begin{array}{l} B = I - \tfrac{1}{2}\mu A_* \,, \\ C = I + \tfrac{1}{2}\mu A_* \,, \end{array} \qquad A_* = \begin{bmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix}_{M\times M} .$$

All $M \times M$ TST matrices share the same eigenvectors, hence so does $B^{-1}C$. Moreover, these eigenvectors are orthogonal. Therefore, also $A = B^{-1}C$ is normal and its eigenvalues are

$$\lambda_k(A) = \frac{\lambda_k(C)}{\lambda_k(B)} = \frac{1 - 2\mu \sin^2 \frac{1}{2}\pi k h}{1 + 2\mu \sin^2 \frac{1}{2}\pi k h} \quad \Rightarrow \quad |\lambda_k(A)| \le 1, \qquad k = 1...M.$$

Consequently Crank–Nicolson is stable for all $\mu > 0$. [Note: Similarly to the situation with stiff ODEs, this *does not* mean that $k = \Delta t$ may be arbitrarily large, but that the only valid consideration in the choice of $k = \Delta t$ vs $h = \Delta x$ is accuracy.]

Example 2.8 (Convergence of the Crank-Nicolson method for diffusion equation) It is not difficult to verify that the local error of the Crank-Nicolson scheme is $\eta_m^n = \mathcal{O}(k^3 + kh^2)$, where $\mathcal{O}(k^3)$ is inherited from the trapezoidal rule (compared to $\mathcal{O}(k^2)$ for the Euler method). We also have

$$\|\boldsymbol{\eta}^n\| = \{h \textstyle\sum_{m=1}^M |\eta_m^n|^2\}^{1/2} = \mathcal{O}(k^3 + kh^2) \,.$$

Hence, for the error vectors $\boldsymbol{e}^n$ we have

$$B\boldsymbol{e}^{n+1} = C\boldsymbol{e}^n + \boldsymbol{\eta}^n \quad \Rightarrow \quad \|\boldsymbol{e}^{n+1}\| \le \|B^{-1}C\| \cdot \|\boldsymbol{e}^n\| + \|B^{-1}\| \cdot \|\boldsymbol{\eta}^n\| \,.$$

We have just proved that $\|B^{-1}C\| \le 1$, and we also have $\|B^{-1}\| \le 1$, because all the eigenvalues of B are greater than 1 (by Gershgorin's theorem). Therefore, $\|\boldsymbol{e}^{n+1}\| \le \|\boldsymbol{e}^n\| + \|\boldsymbol{\eta}^n\|$, and

$$\|\boldsymbol{e}^n\| \le \|\boldsymbol{e}^0\| + n\|\boldsymbol{\eta}\| = n\|\boldsymbol{\eta}\| \le \tfrac{cT}{k}(k^3 + kh^2) = cT(k^2 + h^2).$$

Thus, taking $k = \alpha h$ will result in $\mathcal{O}(h^2)$ error of approximation.

Example 2.9 (Crank–Nicolson for advection equation) Let

$$u_m^{n+1} - u_m^n = \tfrac{1}{4}\mu(u_{m+1}^{n+1} - u_{m-1}^{n+1}) + \tfrac{1}{4}\mu(u_{m+1}^n - u_{m-1}^n), \qquad m = 1...M\,.$$

(This is the trapezoidal rule applied to the semidiscretization of advection equation $\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}$). In this case, $\boldsymbol{u}^{n+1} = B^{-1}C\boldsymbol{u}^n$, where the matrices B and C are Toeplitz antisymmetric tridiagonal,

$$B = \begin{bmatrix} 1 & -\frac{1}{4}\mu & & \\ \frac{1}{4}\mu & 1 & \ddots & \\ & \ddots & \ddots & -\frac{1}{4}\mu \\ & & \frac{1}{4}\mu & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & \frac{1}{4}\mu & & \\ -\frac{1}{4}\mu & 1 & \ddots & \\ & \ddots & \ddots & \frac{1}{4}\mu \\ & & -\frac{1}{4}\mu & 1 \end{bmatrix}.$$

Similarly to Exercise 4, the eigenvalues and eigenvectors of the matrix

$$S = \begin{bmatrix} \alpha & \beta & & \\ -\beta & \alpha & \ddots & \\ & \ddots & \ddots & \beta \\ & & -\beta & \alpha \end{bmatrix},$$

are given by $\lambda_k = \alpha + 2\,\mathrm{i}\,\beta\cos kx$, and $\boldsymbol{w}_k = (\mathrm{i}^m \sin kmx)_{m=1}^M$, where $x = \pi h = \frac{\pi}{M+1}$. So, all such S are normal and share the same eigenvectors, hence so does $A = B^{-1}C$, hence A is normal and

$$\lambda_k(A) = \frac{\lambda_k(C)}{\lambda_k(B)} = \frac{1 + \frac{1}{2}\,\mathrm{i}\,\mu\cos kx}{1 - \frac{1}{2}\,\mathrm{i}\,\mu\cos kx} \quad \Rightarrow \quad |\lambda_k(A)| = 1, \qquad k = 1...M.$$

So, Crank–Nicolson is again stable for all $\mu > 0$.

Example 2.10 (Euler for advection equation) Finally, consider the Euler method for advection equation

$$u_m^{n+1} - u_m^n = \mu(u_{m+1}^n - u_m^n), \qquad m = 1...M\,.$$

We have $\boldsymbol{u}^{n+1} = A\boldsymbol{u}^n$, where

$$A = \begin{bmatrix} 1-\mu & \mu & & \\ & 1-\mu & \ddots & \\ & & \ddots & \mu \\ & & & 1-\mu \end{bmatrix},$$

but A is *not* normal, and although its eigenvalues are bounded by 1 for $\mu \le 2$, it is the spectral radius of AA^T that matters, and we have $\rho(AA^T) \approx (|1-\mu| + |\mu|)^2$, so that the method is stable only if $\mu \le 1$.