Dr H. Fawzi

Mathematical Tripos Part II: Michaelmas Term 2022

Numerical Analysis – Lecture 6

Semidiscretization Let $u_m(t) = u(mh, t)$, $m = 1...M$, $t \geq 0$. Approximating $\partial^2/\partial x^2$ as before, we deduce from the PDE that the *semidiscretization*

$$\frac{du_m}{dt} = \frac{1}{h^2}(u_{m-1} - 2u_m + u_{m+1}), \qquad m = 1...M \tag{2.2}$$

carries an error of $\mathcal{O}(h^2)$. This is an ODE system, and we can solve it by any ODE solver. Thus, Euler's method yields (2.2), while backward Euler results in

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

This approach is commonly known as *the method of lines.* Much (although not all!) of the theory of finite-difference methods for PDEs of evolution can be presented as a two-stage task: first semidiscretize, getting rid of space variables, then use an ODE solver. Typically, each stage is conceptually easier than the process of discretizing in unison in both time and in space (so-called *full discretization*).

Example 2.4 (The Crank–Nicolson scheme for the diffusion equation) Discretizing the ODE (2.2) with the trapezoidal rule, we obtain

$$u_m^{n+1} - \tfrac{1}{2}\mu(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n + \tfrac{1}{2}\mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n), \qquad m = 1...M. \tag{2.3}$$

Thus, each step requires the solution of an $M \times M$ symmetric tridiagonal system. The error of the scheme is $\mathcal{O}(k^3 + kh^2)$, so basically the same as with Euler's method. However, as we will see, Crank–Nicolson enjoys superior stability features, as compared with the method (2.2).

Note further that (2.3) is an *implicit* method: advancing each time step requires to solve a linear algebraic system. However, the matrix of the system is symmetric tridiagonal and its solution by sparse Cholesky factorization can be done in $\mathcal{O}(M)$ operations.

Stability: Let's now analyze the stability of the Crank-Nicolson scheme. The recurrence equation (2.3) can be written as $B\boldsymbol{u}^{n+1} = C\boldsymbol{u}^n$, where the matrices B and C are Toeplitz symmetric tridiagonal (TST),

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

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

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

Consequently Crank–Nicolson is stable for all $\mu > 0$.

Convergence: We can now analyze the convergence of the Crank-Nicolson scheme. 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

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

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 2.5 (Crank–Nicolson for advection equation) Consider the advection equation:

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}.$$

If we discretize the right-hand side by $\frac{\partial u}{\partial x} = \frac{1}{2h}(u(x+h,t) - u(x-h,t)) + \mathcal{O}(h^2)$ we end up with the ODE

$$\frac{du_m}{dt} = \frac{1}{2h}(u_{m+1} - u_{m-1}).$$

Using the trapezoidal rule this yields

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

where $\mu = k/h$. In this case, $\boldsymbol{u}^{n+1} = B^{-1}C\boldsymbol{u}^n$, where the matrices B and C are given by

$$B = \begin{bmatrix} 1 & -\frac{1}{4}\mu & & \\ \frac{1}{4}\mu & 1 & \ddots & \\ & \ddots & \ddots & -\frac{1}{4}\mu \\ & & \frac{1}{4}\mu & 1 \end{bmatrix} = I - \frac{1}{4}\mu A_*, \qquad C = \begin{bmatrix} 1 & \frac{1}{4}\mu & & \\ -\frac{1}{4}\mu & 1 & \ddots & \\ & \ddots & \ddots & \frac{1}{4}\mu \\ & & -\frac{1}{4}\mu & 1 \end{bmatrix} = I + \frac{1}{4}\mu A_*$$

where

$$A_* = \begin{bmatrix} 0 & 1 & & \\ -1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & 0 \end{bmatrix}.$$

Note that A_* is skew-symmetric ($A_*^T = -A_*$) and so it is normal, and its eigenvalues are pure imaginary, i.e., $\lambda_k(A_*) = \mathrm{i}\gamma_k$ where $\gamma_k \in \mathbb{R}$ for $\ell = 1, \ldots, M$ (one can show that $\gamma_k = 2\cos(k\pi h)$). Reasoning like in the previous example, the eigenvalues of $A = B^{-1}C$ are thus given by

$$\lambda_k(A) = \frac{\lambda_k(C)}{\lambda_k(B)} = \frac{1 + \frac{\mu}{4}\,\mathrm{i}\,\gamma_k}{1 - \frac{\mu}{4}\,\mathrm{i}\,\gamma_k} \quad \Rightarrow \quad |\lambda_k(A)| = 1, \qquad k = 1...M.$$

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

Example 2.6 (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 $\mu \le 2$ (note $1 - \mu$ is the only eigenvalue of A), it is the matrix induced norm of A that matters. For this example, it is easier to work with $\|A\|_{\infty\to\infty}$ which we see is given by $|1-\mu| + \mu$ (by the formula in Lecture 5), and this is smaller than 1 precisely when $\mu \le 1$.