Prof. A. C. Hansen

Mathematical Tripos Part II: Michaelmas Term 2024

Numerical Analysis – Lecture 10

Technique 2.31 (Splitting for the 2D diffusion equation) Recall that for the 2D diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x^2}$$

using the five-point discretisation scheme for the Laplacian yields the following ODE

$$\frac{d\boldsymbol{u}}{dt} = \frac{1}{h^2}(A_x + A_y)\boldsymbol{u}$$

where the matrices A_x and A_y are expressed as $A_x = G \otimes I$ and $A_y = I \otimes G$, where $\otimes$ is the Kronecker product, and G is the $M \times M$ tridiagonal matrix

$$G = \begin{bmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix} \in \mathbb{R}^{M \times M}.$$

It is straightforward to verify that A_x and A_y commute; namely $A_xA_y = A_yA_x = G\otimes G$ (check out the basic rules of multiplication with the kronecker product `https://en.wikipedia.org/wiki/Kronecker_product`). This should not come as a suprise since the operators $\partial^2/\partial x^2$ and $\partial^2/\partial y^2$, which A_x/h^2 and A_y/h^2 approximate, are known to commute. So we can write

$$\mathrm{e}^{k(A_x+A_y)/h^2} = \mathrm{e}^{kA_x/h^2}\mathrm{e}^{kA_y/h^2}.$$

This means that the solution of the semi-discretized diffusion equation in 2D, with zero boundary conditions, satisfies

$$\boldsymbol{u}^{n+1} = \mathrm{e}^{kA_x/h^2}\mathrm{e}^{kA_y/h^2}\boldsymbol{u}^n. \tag{2.17}$$

The split Crank-Nicolson scheme: In the split Crank-Nicolson scheme, we approximate each exponential map in (2.17) by the rational function

$$r(z) = (1 + z/2)(1 - z/2)^{-1},$$

which leads to

$$\boldsymbol{u}^{n+1} = (I + \tfrac{\mu}{2}A_x)(I - \tfrac{\mu}{2}A_x)^{-1}(I + \tfrac{\mu}{2}A_y)(I - \tfrac{\mu}{2}A_y)^{-1}\boldsymbol{u}^n. \tag{2.18}$$

Note that computing $\boldsymbol{u}^{n+1/2} = (I + \frac{\mu}{2}A_y)(I - \frac{\mu}{2}A_y)^{-1}\boldsymbol{u}^n$ can be done efficiently in $\mathcal{O}(M^2)$ time as A_y is block-diagonal, and the matrices G are tridiagonal (each tridiagonal solve requires $\mathcal{O}(M)$ time, and we have M of these). Computing $\boldsymbol{u}^{n+1} = (I+\frac{\mu}{2}A_x)(I-\frac{\mu}{2}A_x)^{-1}\boldsymbol{u}^{n+1/2}$ can also be done in $\mathcal{O}(M^2)$ time, since A_x is also block-diagonal provided we appropriately permute the rows and columns so that the grid ordering is by rows instead of columns. This means that the update step (2.18) of Split-Crank-Nicolson can be performed in time $\mathcal{O}(M^2)$ and only requires tridiagonal matrix solves (no FFT needed).

Stability: One can easily verify stability of the split Crank-Nicolson scheme. Indeed, we can write

$$\|r(\mu A_x)r(\mu A_y)\|_2 \le \|r(\mu A_x)\|_2\|r(\mu A_y)\|_2 \le 1$$

since, as seen in previous lectures, $\|r(\mu A_x)\|_2 = \|(I + \frac{\mu}{2}A_x)(I - \frac{\mu}{2}A_x)^{-1}\|_2 \le 1$ since A_x is symmetric and its eigenvalues are ≤ 0. (Same for $\|r(\mu A_y)\|_2$.)

Exercise: Check the consistency of the scheme

$$\boldsymbol{u}^{n+1} = r(\mu A_x)r(\mu A_y)\boldsymbol{u}^n.$$

In particular, show that split Crank-Nicolson has the 'same' local error as the classical Crank-Nicolson scheme. That is the local error is $\mathcal{O}(k^3 + kh^2)$.

Example 2.32 Consider the general diffusion equation

$$\frac{\partial u}{\partial t} = \nabla^\top \left(a(x,y)\nabla u\right) + f(x,y) = \frac{\partial}{\partial x}\left(a(x,y)\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(a(x,y)\frac{\partial u}{\partial y}\right) + f(x,y), \tag{2.19}$$

where $a(x,y) > \alpha > 0$ and $f(x,y)$ are given, together with initial conditions on $[0,1]^2$ and Dirichlet boundary conditions along $\partial[0,1]^2 \times [0,\infty)$. Replace each space derivative by *central differences* at midpoints,

$$\frac{\mathrm{d}g(\xi)}{\mathrm{d}\xi} \approx \frac{g(\xi+\frac{1}{2}h) - g(\xi-\frac{1}{2}h)}{h},$$

resulting in the ODE system

$$\begin{aligned} u'_{\ell,m} \;\; = \;\; & \tfrac{1}{h^2}\Big[a_{\ell-\frac12,m}u_{\ell-1,m} + a_{\ell+\frac12,m}u_{\ell+1,m} + a_{\ell,m-\frac12}u_{\ell,m-1} + a_{\ell,m+\frac12}u_{\ell,m+1} \\ & \quad - \big(a_{\ell-\frac12,m} + a_{\ell+\frac12,m} + a_{\ell,m-\frac12} + a_{\ell,m+\frac12}\big)u_{\ell,m}\Big] + f_{\ell,m}. \end{aligned} \tag{2.20}$$

Assuming zero boundary conditions, we have a system $\boldsymbol{u}' = A\boldsymbol{u}$, and the matrix A can be split as $A = \frac{1}{h^2}(A_x + A_y)$. Here, A_x and A_y are again constructed from the contribution of discretizations in the x- and y-directions respectively, namely A_x includes all the $a_{\ell\pm\frac12,m}$ terms, and A_y consists of the remaining $a_{\ell,m\pm\frac12}$ components. The resulting operators A_x and A_y do not necessarily commute, and so the splitting scheme

$$\boldsymbol{u}^{n+1} = \mathrm{e}^{kA_x/h^2}\mathrm{e}^{kA_y/h^2}\boldsymbol{u}^n$$

will carry an error of $\mathcal{O}(k^2)$.

Strang splitting : One can obtain better splitting approximations of $\mathrm{e}^{t(B+C)}$. For example it is not hard to prove that $\mathrm{e}^{\frac12 tB}\mathrm{e}^{tC}\mathrm{e}^{\frac12 tB}$ gives a $\mathcal{O}(t^3)$ approximation of $\mathrm{e}^{t(B+C)}$, i.e.,

$$\mathrm{e}^{t(B+C)} = \mathrm{e}^{\frac12 tB}\mathrm{e}^{tC}\mathrm{e}^{\frac12 tB} + \mathcal{O}(t^3). \tag{2.21}$$

Technique 2.33 (Splitting methods) Recall that, for $z_1, z_2 \in \mathbb{C}$, we have $\mathrm{e}^{z_1+z_2} = \mathrm{e}^{z_1}\mathrm{e}^{z_2}$ and had this been true for the matrices, i.e. that $\mathrm{e}^{tA} = \mathrm{e}^{t(B+C)} = \mathrm{e}^{tB}\mathrm{e}^{tC}$, we could have approximated each component of the exponent of $A = A_x + A_y$ with the trapezoidal rule, say, to produce

$$\boldsymbol{u}^{n+1} = \big(I - \tfrac12\mu A_x\big)^{-1}\big(I + \tfrac12\mu A_x\big)\big(I - \tfrac12\mu A_y\big)^{-1}\big(I + \tfrac12\mu A_y\big)\boldsymbol{u}^n, \qquad \mu = k/h^2, \tag{2.22}$$

and since both $I - \frac12\mu A_x$ and $I - \frac12\mu A_y$ are tridiagonal, this system can be solved very cheaply.

Unfortunately, the assumption that $\mathrm{e}^{t(B+C)} = \mathrm{e}^{tB}\mathrm{e}^{tC}$ is, in general, false. Not all hope is lost, though, and we will demonstrate that, suitably implemented, splitting is a powerful technique to reduce drastically the expense of numerical solution.

Method 2.34 (Splitting of inhomogeneous systems) Our exposition so far has been limited to the case of zero boundary conditions. In general, the linear ODE system is of the form

$$\boldsymbol{u}' = A\boldsymbol{u} + \boldsymbol{b}, \qquad \boldsymbol{u}(0) = \boldsymbol{u}^0, \tag{2.23}$$

where $\boldsymbol{b}$ originates in boundary conditions (and, possibly, in a forcing term $f(x,y)$ in the original PDE (2.19)). Note that our analysis should accommodate $\boldsymbol{b} = \boldsymbol{b}(t)$, since boundary conditions might vary in time! The *exact* solution of (2.23) is provided by the *variation of constants* formula

$$\boldsymbol{u}(t) = \mathrm{e}^{tA}\boldsymbol{u}(0) + \int_0^t \mathrm{e}^{(t-s)A}\boldsymbol{b}(s)\,\mathrm{d}s, \qquad t \ge 0,$$

therefore

$$\boldsymbol{u}(t_{n+1}) = \mathrm{e}^{kA}\boldsymbol{u}(t_n) + \int_{t_n}^{t_{n+1}} \mathrm{e}^{(t_{n+1}-s)A}\boldsymbol{b}(s)\,\mathrm{d}s.$$

The integral on the right-hand side can be evaluated using quadrature.

For example, the trapezoidal rule $\int_0^k g(\tau)\,\mathrm{d}\tau = \frac{1}{2}k[g(0)+g(k)] + \mathcal{O}(k^3)$ gives

$$\boldsymbol{u}(t_{n+1}) \;\approx\; \mathrm{e}^{kA}\boldsymbol{u}(t_n) + \tfrac{1}{2}k[\mathrm{e}^{kA}\boldsymbol{b}(t_n) + \boldsymbol{b}(t_{n+1})],$$

with a local error of $\mathcal{O}(k^3)$. We can now replace exponentials with their splittings. For example, Strang's splitting (2.21), together with the rational function approximation $r(z) = (1+z/2)/(1-z/2)$ of the exponential map, results in

$$\boldsymbol{u}^{n+1} \;=\; r\big(\tfrac{1}{2}kB\big)\, r\big(kC\big)\, r\big(\tfrac{1}{2}kB\big)\big[\boldsymbol{u}^n + \tfrac{1}{2}k\boldsymbol{b}^n\big] + \tfrac{1}{2}k\boldsymbol{b}^{n+1}.$$

As before, everything reduces to (inexpensive) solution of tridiagonal systems.