6. Matrix representation. Neumann and Periodic Boundary Conditions

A general, implicit one–step algorithm can be represented in the form

$$\sum_{n=1}^{N-1} A_{mn} U_n^{j+1} = \sum_{n=1}^{N-1} B_{mn} U_n^j + c_m^j \qquad \text{for} \quad 1 \leqslant m \leqslant N-1, \quad 1 \leqslant j \leqslant J\,. \tag{6.1}$$

We introduce the solution and truncation error vectors $\mathbf{Z}^j = (z_1^j,\ z_2^j, \ldots,\ z_{N-1}^j)$ and $\varepsilon \mathbf{d}^j = (R_1^j,\ R_2^j, \ldots,\ R_{N-1}^j)$ where ε is small, typically $\varepsilon = O(k^2 + kh^2)$. A and B are $(N-1) \times (N-1)$ matrices and then the error–vector obeys the equation

$$A\mathbf{Z}^{j+1} = B\mathbf{Z}^j + \varepsilon \mathbf{d}^j. \tag{6.2}$$

Assuming the matrix inverse A^{-1} exists (else the entire method collapses as we will be unable to find U_n^{j+1} from U_n^j) we can write

$$\mathbf{Z}^{j+1} = \left(A^{-1}B\right)\mathbf{Z}^j + \varepsilon A^{-1}\mathbf{d}^j = P\mathbf{Z}^j + \varepsilon \mathbf{f}^j \qquad \text{say.} \tag{6.3}$$

The initial error, $\mathbf{Z}^0 = 0$. Applying the above relation iteratively leads to

$$\mathbf{Z}^j = \varepsilon \left[(P)^{j-1}\mathbf{f}^0 + (P)^{j-2}\mathbf{f}^1 + \ldots + P\mathbf{f}^{j-2} + \mathbf{f}^{j-1}\right]\,. \tag{6.4}$$

The matrix P describes the *propagation* of errors from one time–step to the next. For the method (6.1) to be stable, we must have $\mathbf{Z}^J \to 0$ as $J \to \infty$ and $k \to 0$. We therefore want to consider the vector $(P)^j\mathbf{x}$ where $\mathbf{x}$ is any vector and j is large.

Suppose the matrix P has eigenvectors $\mathbf{e}_m$ and eigenvalues λ_m. Then

$$P\mathbf{e}_m = \lambda_m \mathbf{e}_m\,, \qquad (P)^2\mathbf{e}_m = \lambda_m^2 \mathbf{e}_m \qquad \text{and} \quad (P)^j\mathbf{e}_m = (\lambda_m)^j \mathbf{e}_m\,. \tag{6.5}$$

Thus, if one of the eigenvalues has modulus greater than one it is possible for $|(P)^j\mathbf{x}|$ to increase without limit. If this happens the method (6.1) will be unstable (although see the subtlety overleaf.) Conversely, suppose that $|\lambda_m| \leqslant 1$ for all m. Then it is easy to show that when the eigenvectors of P are *complete* so that we may expand $\mathbf{x} = \sum a_m \mathbf{e}_m$, then

$$\left|(P)^j\mathbf{x}\right| = \left|\sum_{m=1}^{N-1} a_m (P)^j \mathbf{e}_m\right| = \left|\sum a_m (\lambda_m)^j \mathbf{e}_m\right| \leqslant \sum |a_m||\lambda_m|^j \leqslant \sum |a_m|\,. \tag{6.6}$$

This is bounded and so the error vector $\mathbf{Z}^j$ in (6.6) remains small. Indeed, if all the eigenvalues are strictly less than one in modulus, then the propagated errors decrease as j increases. The solution error is then dominated by the local truncation error.

We have shown that the method (6.1) is stable provided $\max|\lambda_m| \leqslant 1$ and the eigenvectors are complete. In fact this result is true for incomplete sets of eigenvectors also, but is harder to prove. The maximum value of $|\lambda_m|$ is often called the **spectral radius** of P, and sometimes written as $\rho(P)$. When calculating the eigenvalues, it is usually best to consider the equation $|B - \lambda A| = 0$ rather than calculating P.

A subtlety: Bounded growth. Even if the spectral radius is a little greater than one, the method may still be stable, provided the errors grow in a bounded fashion. Consider a time interval $0 < t < T = Jk$. The errors over this period are bounded if for all m and some constant G,

$$|\lambda_m|^J \leqslant G \iff |\lambda_m| \leqslant G^{1/J} = e^{(\ln G)k/T} \approx 1 + k\frac{\ln G}{T} + O(k^2) \qquad \text{as} \quad k \to 0\ . \quad (6.7)$$

The stability condition allowing for bounded growth is $|\lambda_m| \leqslant 1 + O(k)$.

(Nearly) Tridiagonal systems

Last time we showed that an important case has A and B tridiagonal, which permits efficient solution. The Crank-Nicolson ($\theta = \frac{1}{2}$) method for the equation $u_t = u_{xx}$ had

$$A = \begin{pmatrix} 1+r & -\mu r & 0 & \ddots & d \\ -r/2 & 1+r & -r/2 & \ddots & 0 \\ 0 & -r/2 & 1+r & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots & -r/2 \\ d & 0 & \ddots & -\mu r & 1+r \end{pmatrix}, \qquad (6.8)$$

where $\mu = 1/2$ and $d = 0$. For more general problems, the diagonal entries vary with the row number, while keeping the tridiagonal structure.

Neumann Boundary Conditions: Slight differences occur if we change the boundary conditions on $x = 0, 1$. The condition $u = 0$ on a boundary is called a **Dirichlet** condition. An important alternative are the **Neumann** boundary conditions, $u_x = 0$ at $x = 0$, say. Physically, this corresponds to no flux across the boundary, so that if u corresponds to temperature, this would be an insulating boundary. In this case the value of u is unknown on the wall. One could require that $U_0^j = U_1^j$, for example, which keeps the same number of unknowns as in the Dirichlet case. This would change the (1,1) element of A to $1 + \frac{1}{2}r$ (and $\mu = \frac{1}{2}$). A better way of dealing with this condition is to introduce a fictitious point at $n = -1$, and then require $U_{-1} = U_1$. Then $(\delta^2 u)_0 = 2u_1 - 2u_0$, which we can use in the FD scheme evaluated on the boundary. If we have Neumann conditions at both ends, we would then have a matrix A given by (6.8) where $\mu = 1$ and $d = 0$. However, in that case we may have 2 more rows and columns in A as we have 2 more unknowns.

Periodic Boundary Conditions: Another important case to consider are **Periodic** boundary conditions, where we identify $x = 0$ and $x = 1$. The new boundary value $U_N = U_0$, and the fictitious points U_{-1} and U_{N+1} are identified as U_{N-1} and U_1 respectively. When we evaluate the FDM on the boundary, we get an extra equation, and the matrix A takes the form (6.3) with $\mu = 1/2$ and $d = -\frac{1}{2}r$. The matrix is no longer strictly tri-diagonal, and we have to modify our solving routine.