Numerical Analysis – Lecture 20[1]

Technique 5.9 *(Eigenvalue analysis of stability)* We say that a matrix A is *normal* if $A = QD\bar{Q}^\top$, where D is a (complex) diagonal matrix and Q is a unitary matrix. Examples: symmetric and skew-symmetric matrices.
An interesting alternative definition of a (complex) normal matrix A is that it commutes with its adjoint $\bar{A}^\top$. Can you prove that the first definition implies the second?

Proposition 5.10 *A normal $\Rightarrow \|A\|_2 = \rho(A)$.*

Proof Recall that, for a general complex matrix B, $\|B\| = \max_{\|\boldsymbol{v}\|=1} \|B\boldsymbol{v}\|$. In particular, let $\boldsymbol{w}$ be an eigenvector of B, $B\boldsymbol{w} = \lambda\boldsymbol{w}$, $\|\boldsymbol{w}\| = 1$. Thus, $\|B\boldsymbol{w}\| = |\lambda|$ and we deduce that $\|B\| \geq \rho(B)$ for every matrix B and norm $\|\cdot\|$. Next let A be normal and recall that $\|Q\boldsymbol{v}\|_2 = \|\boldsymbol{v}\|_2\ \forall\boldsymbol{v}$ (unitary matrices are *isometries* w.r.t. the Euclidean norm). Therefore $\|A\boldsymbol{v}\|_2 = \|QD\bar{Q}^\top\boldsymbol{v}\|_2 = \|D\bar{Q}^\top\boldsymbol{v}\|_2$. Let $\boldsymbol{u} = \bar{Q}^\top\boldsymbol{v}$ (this is the same as rendering $\boldsymbol{v}$ in the basis spanned by the rows of $\bar{Q}$). Hence $\|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2$ and $\|A\boldsymbol{v}\|_2 = \|D\boldsymbol{u}\|_2$. D is diagonal, therefore

$$\begin{aligned} &\operatorname{diag} D = \sigma(A) \quad\Rightarrow\quad \|D\|_2 = \rho(A) \quad\Rightarrow\quad \|A\boldsymbol{v}\|_2 \leq \|D\|_2\|\boldsymbol{u}\|_2 = \rho(A)\|\boldsymbol{v}\|_2 \\ \Rightarrow\quad &\|A\|_2 \leq \rho(A) \quad\Rightarrow\quad \|A\|_2 = \rho(A) \end{aligned}$$

and the proof follows. □

Suppose that a numerical method (with zero boundary conditions) can be written in the form $\boldsymbol{u}^{n+1}_{\Delta x} = A_{\Delta x}\boldsymbol{u}^n_{\Delta x}$, $n \in \mathbb{Z}^+$, where $A_{\Delta x}$ is *normal* for all small $\Delta x > 0$. Induction $\Rightarrow \boldsymbol{u}^n_{\Delta x} = A^n_{\Delta x}\boldsymbol{u}^0_{\Delta x}$. *[Note: A normal $\Rightarrow A^n$ normal for all $n \in \mathbb{Z}^+$, since A and A^n share the same eigenvectors.]* Let

$$\|\boldsymbol{v}_{\Delta x}\|_{2,\Delta x} = \left[(\Delta x)\sum_k |v_k|^2\right]^{1/2}.$$

Remarks:

1. In general, the dimension of $\boldsymbol{v}_{\Delta x}$ depends on Δx;
2. The reason for the factor of $(\Delta x)^{1/2}$ in the definition is to ensure that, because of the convergence of Riemann sums, $\|\boldsymbol{v}_{\Delta x}\|_{2,\Delta x} \overset{\Delta x\to 0}{\longrightarrow} = \left[\int |v(x)|^2 \mathrm{d}x\right]^{1/2} = \|v\|_2$, provided that v is an integrable function such that $v_{k,\Delta x} = v(k\Delta x)$.

We thus have

$$\|\boldsymbol{u}^n_{\Delta x}\|_{2,\Delta x} = \|A^n_{\Delta x}\boldsymbol{u}^0_{\Delta x}\|_{2,\Delta x} \leq [\rho(A_{\Delta x})]^n \|\boldsymbol{u}^0_{\Delta x}\|_{2,\Delta x}.$$

Since $\|\boldsymbol{u}^0_{\Delta x}\|_{2,\Delta x}$ can be uniformly bounded for a square-integrable initial value, it follows that *the method is stable if $\rho(A_{\Delta x}) \leq 1$ as $\Delta x \to 0$.*

Example 5.11 *(Crank–Nicolson)* Let

$$u^{n+1}_m - \tfrac{1}{2}\mu(u^{n+1}_{m-1} - 2u^{n+1}_m + u^{n+1}_{m+1}) = u^n_m + \tfrac{1}{2}\mu(u^n_{m-1} - 2u^n_m + u^n_{m+1}), \quad m = 1,2,\ldots,M,$$

Therefore $\boldsymbol{u}^{n+1} = B^{-1}C\boldsymbol{u}^n$, where the matrices B and C are TST,

$$B = \begin{bmatrix} 1+\mu & -\frac{1}{2}\mu & & \\ -\frac{1}{2}\mu & 1+\mu & & \\ & \ddots & \ddots & -\frac{1}{2}\mu \\ & & -\frac{1}{2}\mu & 1+\mu \end{bmatrix}, \qquad C = \begin{bmatrix} 1-\mu & \frac{1}{2}\mu & & \\ \frac{1}{2}\mu & 1-\mu & & \\ & \ddots & \ddots & \frac{1}{2}\mu \\ & & \frac{1}{2}\mu & 1-\mu \end{bmatrix}.$$

[1] Please email all corrections and suggestions to these notes to `A.Iserles@damtp.cam.ac.uk`. All handouts are available on the WWW at the URL `http://www.damtp.cam.ac.uk/user/na/PartII/Handouts.html`.

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

$$\frac{1-\mu+\mu\cos\frac{2\pi k}{M+1}}{1+\mu-\mu\cos\frac{2\pi k}{M+1}} = \frac{1-2\mu\sin^2\frac{\pi k}{M+1}}{1+2\mu\sin^2\frac{\pi k}{M+1}} \le 1, \qquad k = 1, 2, \ldots, M-1.$$

Consequently Crank–Nicolson is stable $\forall\, \mu > 0$.
[Note: Similarly to the situation with stiff ODEs, this does not mean that Δt may be arbitrarily large, but that the only valid consideration in the choice of Δt vs Δx is accuracy.]

Technique 5.12 *(Fourier analysis of stability)* Let us now assume a recurrence of the form

$$\sum_{k=-r}^{s} \alpha_k u_{m+k}^{n+1} = \sum_{k=-r}^{s} \beta_k u_{m+k}^{n}, \qquad n \in \mathbb{Z}^+, \tag{5.3}$$

where m ranges over $\mathbb{Z}$ (within our framework of discretizing PDEs of evolution, this corresponds to $-\infty < x < \infty$ in the undelying PDE and so there are no explicit boundary conditions but the initial condition must be square-integrable in $(-\infty, \infty)$: this is known as a *Cauchy problem.*). The coefficients α_k and β_k are independent of m, n, but typically depend upon μ. We investigate the stability of (5.3) by *Fourier analysis. [Note that it doesn't matter what is the underlying PDE: numerical stability is a feature of algebraic recurrences, not of PDEs!]*

Let $\boldsymbol{v} \in \ell_2[\mathbb{Z}]$. Its *Fourier transform* is the function

$$\hat{v}(\theta) = \sum_{m\in\mathbb{Z}} \mathrm{e}^{-\mathrm{i}m\theta} v_m, \qquad -\pi \le \theta \le \pi.$$

We equip sequences and functions with the norms

$$\|\boldsymbol{v}\| = \left\{ \sum_{m\in\mathbb{Z}} |v_m|^2 \right\}^{\frac{1}{2}} \qquad \text{and} \qquad |||\hat{v}||| = (2\pi)^{-\frac{1}{2}} \left\{ \int_{-\pi}^{\pi} |\hat{v}(\theta)|^2 \,\mathrm{d}\theta \right\}^{\frac{1}{2}}$$

respectively.

Lemma 5.13 $\|\boldsymbol{v}\| = |||\hat{v}||| \;\forall\, \boldsymbol{v} \in \ell_2[\mathbb{Z}]$.

Proof By definition,

$$\begin{aligned} |||\hat{v}|||^2 &= \tfrac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{m\in\mathbb{Z}} \mathrm{e}^{-\mathrm{i}m\theta} v_m \right|^2 \mathrm{d}\theta = \tfrac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{m\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} v_m \bar{v}_k \mathrm{e}^{-\mathrm{i}(m-k)\theta} \,\mathrm{d}\theta \\ &= \tfrac{1}{2\pi} \sum_{m\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} v_m \bar{v}_k \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}(m-k)\theta} \,\mathrm{d}\theta. \end{aligned}$$

But

$$\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}l\theta} \,\mathrm{d}\theta = \begin{cases} 2\pi, & l = 0, \\ 0, & l \in \mathbb{Z} \setminus \{0\}, \end{cases}$$

and we deduce that $|||\hat{v}||| = \|\boldsymbol{v}\|$. $\square$

The implication of the lemma is that the Fourier transform is an *isometry* of the Euclidean norm. This is an important reason underlying its many applications in mathematics and beyond.