Numerical Analysis - Part II

Anders C. Hansen

Lecture 15

Spectral Methods

Chebyshev polynomials

The Chebyshev polynomial of degree n is defined as

$$T_n(x) := \cos n \arccos x, \quad x \in [-1, 1],$$

or, in a more instructive form,

$$T_n(x) := \cos n\theta, \quad x = \cos\theta, \quad \theta \in [0, \pi]. \tag{1}$$

The three-term recurrence relation

1) The sequence (T_n) obeys the three-term recurrence relation

$$T_0(x) \equiv 1, \quad T_1(x) = x,$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1,$$

in particular, T_n is indeed an algebraic polynomial of degree n, with the leading coefficient 2^{n-1}. (The recurrence is due to the equality $\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta$ via substitution $x = \cos\theta$, expressions for T_0 and T_1 are straightforward.)

The recurrence yields

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2-1, \quad T_3(x) = 4x^3-3x, \ \ldots,$$

and T_n is called the nth *Chebyshev polynomial* (of the first kind).

Chebyshev polynomials - How they look

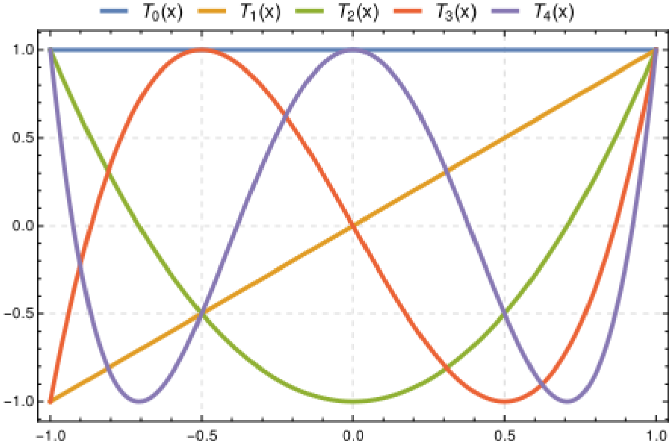

Figure : Chebyshev polynomials

Chebyshev polynomials are orthogonal

2) Also, (T_n) form a sequence of orthogonal polynomials with respect to the inner product $(f,g)_w := \int_{-1}^{1} f(x)g(x)w(x)dx$, with the weight function $w(x) := (1-x^2)^{-1/2}$. Namely, we have

$$
\begin{aligned}
(T_n, T_m)_w &= \int_{-1}^{1} T_m(x)\,T_n(x)\,\frac{dx}{\sqrt{1-x^2}} = \int_0^{\pi} \cos m\theta \cos n\theta \, d\theta \\
&= \begin{cases} \pi, & m = n = 0\,, \\ \frac{\pi}{2}, & m = n \geq 1\,, \\ 0, & m \neq n\,. \end{cases}
\end{aligned}
\tag{2}
$$

Chebyshev expansion

Since $(T_n)_{n=0}^{\infty}$ form an orthogonal sequence, a function f such that $\int_{-1}^{1} |f(x)|^2 w(x)\, dx < \infty$ can be expanded in the series

$$f(x) = \sum_{n=0}^{\infty} \check{f}_n T_n(x),$$

with the Chebyshev coefficients $\check{f}_n$. Making inner product of both sides with T_n and using orthogonality yields

$$\begin{aligned}(f, T_n)_w = \check{f}_n (T_n, T_n)_w \quad \Rightarrow \quad \check{f}_n &= \frac{(f, T_n)_w}{(T_n, T_n)_w} \\ &= \frac{c_n}{\pi} \int_{-1}^{1} f(x) T_n(x) \frac{dx}{\sqrt{1-x^2}}, \end{aligned} \tag{3}$$

where $c_0 = 1$ and $c_n = 2$ for $n \geq 1$.

Connection to the Fourier expansion

Letting $x = \cos\theta$ and $g(\theta) = f(\cos\theta)$, we obtain

$$\int_{-1}^{1} f(x)\,T_n(x)\,\frac{dx}{\sqrt{1-x^2}} = \int_0^{\pi} f(\cos\theta)\,T_n(\cos\theta)\,d\theta = \frac{1}{2}\int_{-\pi}^{\pi} g(\theta)\cos n\theta\,d\theta\,. \tag{4}$$

Given that $\cos n\theta = \frac{1}{2}(e^{in\theta} + e^{-in\theta})$, and using the Fourier expansion of the 2π-periodic function g,

$$g(\theta) = \sum_{n\in\mathbb{Z}} \widehat{g}_n \mathrm{e}^{in\theta}, \quad \text{where} \quad \widehat{g}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-int}\,dt, \quad n \in \mathbb{Z}\,,$$

we continue (4) as

$$\int_{-1}^{1} f(x)\,T_n(x)\frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}(\widehat{g}_{-n} + \widehat{g}_n)\,,$$

and from (3) we deduce that

$$\check{f}_n = \begin{cases} \widehat{g}_0, & n = 0\,, \\ \widehat{g}_{-n} + \widehat{g}_n, & n \geq 1\,. \end{cases}$$

Properties of the Chebyshev expansion

As we have seen, for a general integrable function f, the computation of its Chebyshev expansion is equivalent to the Fourier expansion of the function $g(\theta) = f(\cos\theta)$. Since the latter is periodic with period 2π, we can use a discrete Fourier transform (DFT) to compute the Chebyshev coefficients $\check{f}_n$. [Actually, based on this connection, one can perform a direct fast Chebyshev transform].

Also, if f can be analytically extended from $[-1, 1]$ (to the so-called Bernstein ellipse), then $\check{f}_n$ decays spectrally fast for $n \gg 1$ (with the rate depending on the size of the ellipse). Hence, the Chebyshev expansion inherits the rapid convergence of spectral methods without assuming that f is periodic.

The algebra of Chebyshev expansions

Let $\mathcal{B}$ be the set of analytic functions in $[-1, 1]$ that can be extended analytically into the complex plane. We identify each such function with its Chebyshev expansion. Like the set $\mathcal{A}$, the set $\mathcal{B}$ is a linear space and is closed under multiplication. In particular, we have

$$\begin{aligned} T_m(x)\,T_n(x) &= \cos(m\theta)\cos(n\theta) \\ &= \tfrac{1}{2}\left[\cos((m-n)\theta) + \cos((m+n)\theta)\right] \\ &= \tfrac{1}{2}\left[\,T_{|m-n|}(x) + T_{m+n}(x)\right] \end{aligned}$$

and hence,

$$\begin{aligned} f(x)g(x) &= \sum_{m=0}^{\infty} \breve{f}_m T_m(x) \cdot \sum_{n=0}^{\infty} \breve{g}_n T_n(x) \\ &= \frac{1}{2} \sum_{m,n=0}^{\infty} \breve{f}_m \breve{g}_n \left[T_{|m-n|}(x) + T_{m+n}(x)\right] \\ &= \frac{1}{2} \sum_{m,n=0}^{\infty} \breve{f}_m \Big(\breve{g}_{|m-n|} + \breve{g}_{m+n}\Big)\, T_n(x). \end{aligned}$$

Derivatives of Chebyshev polynomials

Lemma 1 (Derivatives of Chebyshev polynomials)

We can express derivatives T_n' in terms of (T_k) as follows,

$$T_{2n}'(x) = (2n) \cdot 2 \sum_{k=1}^{n} T_{2k-1}(x), \tag{5}$$

$$T_{2n+1}'(x) = (2n+1) \left[T_0(x) + 2 \sum_{k=1}^{n} T_{2k}(x) \right]. \tag{6}$$

Proof. From (1), we deduce

$$T_m(x) = \cos m\theta \quad \Rightarrow \quad T'_m(x) = \frac{m \sin m\theta}{\sin \theta} \qquad x = \cos\theta\,.$$

So, for $m = 2n$, (5) follows from the identity

$$\frac{\sin 2n\theta}{\sin\theta} = 2\sum_{k=1}^{n} \cos(2k-1)\theta,$$

which is verified as

$$\begin{aligned} 2\sin\theta \sum_{k=1}^{n} \cos(2k-1)\theta &= \sum_{k=1}^{n} 2\cos(2k-1)\theta \sin\theta \\ &= \sum_{k=1}^{n} \big[\sin 2k\theta - \sin(2k-1)\theta\big] = \sin 2n\theta. \end{aligned}$$

Proof. Cont. For $m = 2n + 1$, (6) turns into identity

$$\frac{\sin(2n+1)\theta}{\sin\theta} = 1 + 2\sum_{k=1}^{n} \cos 2k\theta,$$

and that follows from

$$\begin{aligned}\sin\theta\Big(1 + 2\sum_{k=1}^{n}\cos 2k\theta\Big) &= \sin\theta + \sum_{k=1}^{n}\big[\sin(2k+1)\theta - \sin(2k-1)\theta\big] \\ &= \sin(2n+1)\theta.\end{aligned}$$

□

Remark 2 (Application to PDEs)

With Lemma 1 all derivatives of u can be expressed in an explicit form as a Chebyshev expansion (cf. Exercise 19 on Example Sheets). For the computation of the Chebyshev coefficients the function f has to be sampled at the so-called Chebyshev points $\cos(2\pi k/N)$, $k = -N/2+1, \ldots, N/2$. This results into a grid, which is denser towards the edges. For elliptic problems this is not problematic, however for initial value PDEs such grids can cause numerical instabilities.

Chebyshev expansion for the derivatives

For an analytic function u, the coefficients $\breve{u}_n^{(k)}$ of the Chebyshev expansion for its derivatives are given by the following recursion,

$$\breve{u}_n^{(k)} = c_n \sum_{\substack{m=n+1 \\ n+m \text{ odd}}}^{\infty} m \, \breve{u}_m^{(k-1)}, \quad \forall \, k \geq 1,$$

where $c_0 = 1$ and $c_n = 2$ for $n \geq 1$. This can be derived from Lemma 1 (the case $m = 1$ is the topic of Ex. 19 on the Example Sheets).

The spectral method for evolutionary PDEs

We consider the problem

$$\begin{cases} \dfrac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t), & x \in [-1,1], \quad t \geq 0\,, \\ u(x,0) = g(x), & x \in [-1,1], \end{cases} \tag{7}$$

with appropriate boundary conditions on $\{-1,1\} \times \mathbb{R}_+$ and where $\mathcal{L}$ is a linear operator, e.g. a differential operator (acting on the x variable). We want to solve this problem by the method of lines (semi-discretization), using a spectral method for the approximation of u and its derivatives in the spatial variable x. Then, in a general spectral method, we seek solutions $u_N(x,t)$ with

$$u_N(x,t) = \sum_n c_n(t)\,\varphi_n(x), \tag{8}$$

where $c_n(t)$ are expansion coefficients and φ_n are basis functions chosen according to the specific structure of (7).

The spectral method for evolutionary PDEs

For example, we may take:

1) the *Fourier expansion* with $c_n(t) = \widehat{u}_n(t)$, $\varphi_n(x) = e^{i\pi nx}$ for periodic boundary conditions,

2) a polynomial expansion such as the *Chebyshev expansion* with $c_n(t) = \breve{u}_n(t)$, $\varphi_n(x) = T_n(x)$ for other boundary conditions.

The spectral approximation in space (8) results into a $N \times N$ system of ODEs for the expansion coefficients $\{c_n(t)\}$:

$$\mathbf{c}' = B\mathbf{c}\,, \tag{9}$$

where $B \in \mathbb{R}^{N\times N}$, and $\mathbf{c} = \{c_n(t)\} \in \mathbb{R}^N$. We can solve it with standard ODE solvers (Euler, Crank-Nikolson, etc.) which as we have seen are approximations to the matrix exponent in the exact solution $\mathbf{c}(t) = \mathrm{e}^{tB}\mathbf{c}(0)$.

The diffusion equation

Consider the diffusion equation for a function $u = u(x, t)$,

$$\begin{cases} u_t = u_{xx}, & (x, t) \in [-1, 1] \times \mathbb{R}_+ \,, \\ u(x, 0) = g(x), & x \in [-1, 1] \,. \end{cases} \tag{10}$$

with the periodic boundary conditions $u(-1, t) = u(1, t)$, $u_x(-1, t) = u_x(1, t)$, and standard normalisation $\int_{-1}^{1} u(x, t)\, dx = 0$, both imposed for all values $t \geq 0$.

For each t, we approximate $u(x, t)$ by its N-th order partial Fourier sum in x,

$$u(x, t) \approx u_N(x, t) \; = \sum_{n \in \Gamma_N} \widehat{u}_n(t)\, e^{i\pi n x} \,, \quad \Gamma_N := \{-N/2{+}1, ..., N/2\} \,.$$

The diffusion equation

Then, from (10), we see that each coefficient $\widehat{u}_n$ fulfills the ODE

$$\widehat{u}_n'(t) = -\pi^2 n^2 \widehat{u}_n(t)\,. \qquad n \in \Gamma_N \tag{11}$$

Its exact solution is $\widehat{u}_n(t) = \mathrm{e}^{-\pi^2 n^2 t}\,\widehat{g}_n$ for $n \neq 0$ and we set $\widehat{u}_0(t) = 0$ due to the normalisation condition, so that

$$u_N(x,t) = \sum_{n \in \Gamma_N} \widehat{g}_n\, \mathrm{e}^{-\pi^2 n^2 t}\, \mathrm{e}^{i\pi n x}\,,$$

which is the exact solution truncated to N terms.

Here, we were able to find the exact solution without solving ODE numerically due to the special structure of the Laplacian. However, for more general PDE we will need a numerical method, and thus the issue of stability arises, so we consider this issue on that simplified example.

Stability analysis

The system (11) has the form

$$\widehat{\mathbf{u}}' = B\widehat{\mathbf{u}}\,, \qquad B = \mathrm{diag}\,\{-\pi^2 n^2\}\,, \quad n \in \Gamma_N\,,$$

and we note that (a) all the eigenvalues of B are negative, and that (b) they consist of the eigenvalues $\lambda_n^{(2)}$ of the second order differentiation operator, with $\max|\lambda_n^{(2)}| = (\frac{N}{2})^2$.

If we approximate this system with the Euler method:

$$\widehat{\mathbf{u}}^{k+1} = (I + \tau B)\widehat{\mathbf{u}}^k, \qquad \tau := \Delta t,$$

then we see that, for stability condition $\|I + \tau B\| \leq 1$, we need to scale the time step $\tau = \Delta t \sim N^{-2}$.

Note that, for the Crank-Nikolson scheme, since the spectrum of B is negative, we get stability for any time step $\tau > 0$.
For a general linear operator $\mathcal{L}$ in (7) with constant coefficients, the matix B is again diagonal (hence normal), and provided that it spectrum is negative, for stability we must scale the time step $\tau \sim N^{-m}$, where m is the maximal order of differentiation.

The scaling $\tau \sim N^{-2}$ may seem similar to the scaling $k \sim h^2$ in difference methods which we viewed as a disadvantage, however in spectral methods we can take N, the order of partial Fourier or Chebyshev sums to achieve a good appoximation, rather small. (We may still need to choose τ small enough to get a desired accuracy.)

The diffusion equation with non-constant coefficient

We want to solve the diffusion equation with a non-constant coefficient $a(x) > 0$ for a function $u = u(x,t)$

$$\begin{cases} u_t = (a(x)u_x)_x, & (x,t) \in [-1,1] \times \mathbb{R}_+\,, \\ u(x,0) = g(x), & x \in [-1,1]\,, \end{cases} \tag{12}$$

with boundary and normalization conditions as before. Approximating u by its partial Fourier sum results in the following system of ODEs for the coefficients $\widehat{u}_n$

$$\widehat{u}_n'(t) = -\pi^2 \sum_{m\in\mathbb{Z}} mn\,\widehat{a}_{n-m}\,\widehat{u}_m(t), \qquad n \in \mathbb{Z}.$$

For the discretization in time we may apply the Euler method, this gives

$$\widehat{u}_n^{k+1} = \widehat{u}_n^k - \tau\,\pi^2 \sum_{m\in\Gamma_N} mn\,\widehat{a}_{n-m}\,\widehat{u}_m^k\,, \qquad \tau = \Delta t\,, n \in \Gamma_N$$

or in the vector form

$$\widehat{\mathbf{u}}^{k+1} = (I + \tau B)\widehat{\mathbf{u}}^k,$$

where $B = (b_{m,n}) = (-\pi^2 mn\,\widehat{a}_{n-m})$. For stability of Euler method, we again need $\|I + \tau B\| \leq 1$, but analysis here is less straightforward.

Remark 3 (Chebyshev methods for evolutionary problems)

In general, the boundary conditions for the considered PDEs have to be implemented in the Chebyshev expansion. If the boundary conditions are to be imposed exactly, either the basis functions have to be slightly modified, e.g., to $T_n(x) - 1$ instead of $T_n(x)$ for the boundary condition $u(1) = 0$, or we get additional conditions on the expansion coefficients $\breve{u}_n$ (cf. Exercise 20 from the Example Sheets).

While the exact imposition is in general not a problem for the numerical treatment of elliptic PDEs, as soon as the boundary conditions depend on time we may run into serious stability issues. One way around this is the use of penalty methods in which the boundary conditions is added to the scheme later as a penalty term.

Iterative methods for linear algebraic systems

Solving linear systems with iterative methods

The general *iterative* method for solving $Ax = b$ is a rule $\mathbf{x}^{k+1} = f_k(\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^k)$. We will consider the simplest ones: *linear, one-step, stationary* iterative schemes:

$$\mathbf{x}^{k+1} = H\mathbf{x}^k + \mathbf{v}, \qquad \mathbf{x}^0, \mathbf{v} \in \mathbb{R}^n. \tag{13}$$

Here one chooses H and v so that x^*, a solution of $A\mathbf{x} = \mathbf{b}$, satisfies $\mathbf{x}^* = H\mathbf{x}^* + \mathbf{v}$, i.e. it is the fixed point of the iteration (13) (if the scheme converges). Standard terminology:

the *iteration matrix* H, the *error* $\mathbf{e}^k := \mathbf{x}^* - \mathbf{x}^k$, the *residual* $\mathbf{r}^k := A\mathbf{e}^k$

For a given class of matrices A (e.g. positive definite matrices, or even a single particular matrix), we are interested in *convergent* methods, i.e. the methods such that $\mathbf{x}^k \to \mathbf{x}^* = A^{-1}\mathbf{b}$ for every starting value $\mathbf{x}^0$. Subtracting $\mathbf{x}^* = H\mathbf{x}^* + \mathbf{v}$ from (13) we obtain

$$\mathbf{e}^{k+1} = H\mathbf{e}^k = \cdots = H^{k+1}\mathbf{e}^0, \tag{14}$$

i.e., a method is convergent if $\mathbf{e}^k = H^k\mathbf{e}^0 \to 0$ for any $\mathbf{e}^0 \in \mathbb{R}^n$.

(Iterative refinement). This is the scheme

$$\mathbf{x}^{k+1} = \mathbf{x}^k - S(A\mathbf{x}^k - \mathbf{b})\,.$$

If $S = A^{-1}$, then $\mathbf{x}^{k+1} = A^{-1}\mathbf{b} = \mathbf{x}^*$, so it is suggestive to choose S as an approximation to A^{-1}. The iteration matrix for this scheme is $H_S = I - SA$.

Solving linear systems – Splitting

(Splitting). This is the scheme

$$(A - B)\mathbf{x}^{k+1} = -B\mathbf{x}^k + \mathbf{b}\,,$$

with the iteration matrix $H = -(A - B)^{-1}B$. Any splitting can be viewed as an iterative refinement (and vice versa) because

$$\begin{aligned}(A - B)\mathbf{x}^{k+1} = -B\mathbf{x}^k + \mathbf{b} \quad &\Leftrightarrow \quad (A - B)\mathbf{x}^{k+1} = (A - B)\mathbf{x}^k - (A\mathbf{x}^k - \mathbf{b}) \\ &\Leftrightarrow \quad \mathbf{x}^{k+1} = \mathbf{x}^k - (A - B)^{-1}(A\mathbf{x}^k - \mathbf{b}),\end{aligned}$$

so we should seek a splitting such that $S = (A - B)^{-1}$ approximates A^{-1}.

Theorem 4

Let $H \in \mathbb{R}^{n\times n}$*. Then* $\lim_{k\to\infty} H^k\mathbf{z} = 0$ *for any* $\mathbf{z} \in \mathbb{R}^n$ *if and only if* $\rho(H) < 1$*.*

Proof. 1) Let λ be an eigenvalue of (the real) H, real or complex, such that $|\lambda| = \rho(H) \geq 1$, and let $\mathbf{w}$ be a corresponding eigenvector, i.e., $H\mathbf{w} = \lambda\mathbf{w}$. Then $H^k\mathbf{w} = \lambda^k\mathbf{w}$, and

$$\|H^k\mathbf{w}\|_\infty = |\lambda|^k\|\mathbf{w}\|_\infty \geq \|\mathbf{w}\|_\infty =: \gamma > 0. \tag{15}$$

If $\mathbf{w}$ is real, we choose $\mathbf{z} = w$, hence $\|H^k\mathbf{z}\|_\infty \geq \gamma$, and this cannot tend to zero.
If $\mathbf{w}$ is complex, then $\mathbf{w} = \mathbf{u} + \mathrm{i}\mathbf{v}$ with some real vectors $\mathbf{u}, \mathbf{v}$. But then at least one of the sequences $(H^k\mathbf{u}), (H^k\mathbf{v})$ does not tend to zero. For if both do, then also $H^k\mathbf{w} = H^k\mathbf{u} + \mathrm{i}H^k\mathbf{v} \to 0$, and this contradicts (15).

Proof. Cont. 2) Now, let $\rho(H) < 1$, and assume for simplicity that H possesses n linearly independent eigenvectors $(\mathbf{w}_j)$ such that $H\mathbf{w}_j = \lambda_j \mathbf{w}_j$. Linear independence means that every $\mathbf{z} \in \mathbb{R}^n$ can be expressed as a linear combination of the eigenvectors, i.e., there exist $(c_j) \in \mathbb{C}$ such that $\mathbf{z} = \sum_{j=1}^n c_j \mathbf{w}_j$. Thus,

$$H^k \mathbf{z} = \sum_{j=1}^n c_j \lambda_j^k \mathbf{w}_j \,,$$

and since $|\lambda_j| \leq \rho(H) < 1$ we have $\lim_{k\to\infty} H^k \mathbf{z} = 0$, as required. $\square$