Numerical Analysis - Part II

Anders C. Hansen

Lecture 13

Spectral Methods

General idea of spectral methods

The basic idea of spectral methods is simple. Consider a PDE of the form

$$\mathcal{L}u = f \tag{1}$$

where $\mathcal{L}$ is a differential operator (e.g., $\mathcal{L} = \frac{\partial^2}{\partial x^2}$, or $\mathcal{L} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, etc.) and f is a right-hand side function. We consider a finite-dimensional subspace of functions V spanned by a basis $\psi_1, \dots, \psi_N$. A typical choice for V is a space of (trigonometric) polynomials of finite degree. We seek an approximate solution to the PDE by a linear combination of the ψ_n, i.e., $u_N(x) = \sum_{n=1}^N c_n \psi_n(x)$. Plugging $u_N(x)$ in the PDE we get the following linear equation in the unknowns (c_n):

$$\sum_{n=1}^{N} c_n \mathcal{L}\psi_n = f. \tag{2}$$

General idea of spectral methods

In general the equation will not have a solution, as there is no reason to expect that the original PDE has a solution in the subspace V. However, we can seek to satisfy equation (2) approximately. Assume that the $(\psi_n)_{1\leq n\leq N}$ are an orthonormal family of functions, with respect to some inner product $\langle\cdot,\cdot\rangle$. Then instead of looking for (c_n) that satisfy (2), we will require only that the projection of $\mathcal{L}u_N - f$ on the subspace V is zero. This is the same as requiring that

$$\sum_{n=1}^{N} c_n \langle \mathcal{L}\psi_n, \psi_m\rangle = \langle f, \psi_m\rangle \qquad \forall m = 1, \ldots, N. \tag{3}$$

If we call A the matrix $A_{m,n} = \langle \mathcal{L}\psi_n, \psi_m\rangle$, we end up with a $N \times N$ linear system $Ac = \tilde{f}$, where $\tilde{f}_m = \langle f, \psi_m\rangle$.

In this chapter we will focus on two of the most common choices of basis functions (ψ_n); namely the Fourier basis, and the basis of Chebyshev polynomials.

We focus on one-dimensional problems on the domain $[-1,1]$. The basis of functions we consider here is

$$\psi_n(x) = e^{i\pi n x}, \qquad n \in \mathbb{Z}.$$

These functions are orthonormal with respect to the normalized L^2 inner product on $[-1,1]$, i.e.,

$$\langle \psi_n, \psi_m \rangle = \frac{1}{2}\int_{-1}^{1} \psi_n(x)\overline{\psi_m(x)} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{else.} \end{cases}$$

Definition 1 (Convergence at spectral speed)

An N-term approximation ϕ_N of a function f converges to f at *spectral* speed if $\|\phi_N - f\|$ decays faster than $\mathcal{O}(N^{-p})$ for any $p = 1, 2, \ldots$.

Remark 2

It is possible to prove that there exist constants $c_1, w > 0$ such that $\|\phi_N - f\| \leq c_1 e^{-wN}$ for all $N \in \mathbb{N}$ uniformly in $[-1, 1]$. Thus, convergence is at least at an exponential rate.

The algebra of Fourier expansions

Let $\mathcal{A}$ be the set of all functions $f : [-1, 1] \to C$, which are analytic in $[-1, 1]$, periodic with period 2, and that can be extended analytically into the complex plane. Then $\mathcal{A}$ is a linear space, i.e., $f, g \in \mathcal{A}$ and $\alpha \in \mathbb{C}$ then $f + g \in \mathcal{A}$ and $af \in \mathcal{A}$. In particular, with f and g expressed in its Fourier series, i.e.,

$$f(x) = \sum_{n=-\infty}^{\infty} \widehat{f}_n e^{i\pi nx}, \quad g(x) = \sum_{n=-\infty}^{\infty} \widehat{g}_n e^{i\pi nx}$$

we have

$$f(x) + g(x) = \sum_{n=-\infty}^{\infty} (\widehat{f}_n + \widehat{g}_n) e^{i\pi nx}, \quad \alpha f(x) = \sum_{n=-\infty}^{\infty} \alpha \widehat{f}_n e^{i\pi nx}. \quad (4)$$

The algebra of Fourier expansions

Moreover,

$$f(x) \cdot g(x) = \sum_{n=-\infty}^{\infty} \left(\sum_{m=-\infty}^{\infty} \widehat{f}_{n-m} \widehat{g}_m \right) e^{i\pi nx} = \sum_{n=-\infty}^{\infty} \left(\widehat{f} * \widehat{g} \right)_n e^{i\pi nx}, \tag{5}$$

where $*$ denotes the convolution operator, hence $\widehat{(f \cdot g)}_n = (\widehat{f} * \widehat{g})_n$.
Moreover, if $f \in \mathcal{A}$ then $f' \in \mathcal{A}$ and

$$f'(x) = i\pi \sum_{n=-\infty}^{\infty} n \cdot \widehat{f}_n e^{i\pi nx}. \tag{6}$$

Since $\{\widehat{f_n}\}$ decays faster than $\mathcal{O}(n^{-p})$ for any $p \in \mathbb{N}$, this provides that all derivatives of f have rapidly convergent Fourier expansions.

Consider the two-point boundary value problem: $y = y(x)$, $-1 \leq x \leq 1$, solves

$$y'' + a(x)y' + b(x)y = f(x), \quad y(-1) = y(1), \tag{7}$$

where $a, b, f \in \mathcal{A}$ and we seek a *periodic solution* $y \in \mathcal{A}$ for (7). Substituting y, a, b and f by their Fourier series and using (4)-(6) we obtain an infinite dimensional system of linear equations for the Fourier coefficients $\widehat{y}_n$:

$$-\pi^2 n^2 \widehat{y}_n + i\pi \sum_{m=-\infty}^{\infty} m\widehat{a}_{n-m}\widehat{y}_m + \sum_{m=-\infty}^{\infty} \widehat{b}_{n-m}\widehat{y}_m = \widehat{f}_n, \quad n \in \mathbb{Z}. \tag{8}$$

Application to differential equations

Since $a, b, f \in \mathcal{A}$, their Fourier coefficients decrease rapidly, like $\mathcal{O}(n^{-p})$ for every $p \in \mathbb{N}$. Hence, we can truncate (8) into the N-dimensional system

$$-\pi^2 n^2 \widehat{y}_n + i\pi \sum_{m=-N/2+1}^{N/2} m\widehat{a}_{n-m}\widehat{y}_m + \sum_{m=-N/2+1}^{N/2} \widehat{b}_{n-m}\widehat{y}_m = \widehat{f}_n, \quad (9)$$

where $n = -N/2+1, \ldots, N/2$.

Remark 3

The matrix of (9) is in general dense, but our theory predicts that fairly small values of N, hence very small matrices, are sufficient for high accuracy. For instance: choosing $a(x) = f(x) = \cos \pi x$, $b(x) = \sin 2\pi x$ (which incidentally even leads to a sparse matrix) we get

$N = 16$	error of size 10^{-10}
$N = 22$	error of size 10^{-15} (which is already hitting ϵ_{Mach}).

Computation of Fourier coefficients (DFT)

We have to compute

$$\widehat{f}_n = \frac{1}{2}\int_{-1}^{1} f(t)e^{-i\pi nt}\,dt, \quad n \in \mathbb{Z}. \tag{10}$$

For this, suppose we wish to compute the integral on $[-1, 1]$ of a function $h \in \mathcal{A}$ by means of the Riemann sums on the uniform partition

$$\int_{-1}^{1} h(t)\,dt \approx \frac{2}{N} \sum_{k=-N/2+1}^{N/2} h\left(\frac{2k}{N}\right). \tag{11}$$

This is known as a *rectangle rule*. We want to know how good this approximation is. As in the definition of the DFT, let $\omega_N = e^{2\pi i/N}$. Then we have

$$\begin{aligned} \frac{2}{N} \sum_{k=-N/2+1}^{N/2} h\left(\frac{2k}{N}\right) &= \frac{2}{N} \sum_{k=-N/2+1}^{N/2} \sum_{n=-\infty}^{\infty} \widehat{h}_n e^{2\pi ink/N} \\ &= \frac{2}{N} \sum_{n=-\infty}^{\infty} \widehat{h}_n \sum_{k=-N/2+1}^{N/2} \omega_N^{nk}. \end{aligned} \tag{12}$$

Computation of Fourier coefficients (DFT)

Since $\omega_N^N = 1$ we have

$$\sum_{k=-N/2+1}^{N/2} \omega_N^{nk} = \omega_N^{-n(N/2-1)} \sum_{k=0}^{N-1} \omega_N^{nk} = \begin{cases} N, & n \equiv 0 \pmod{N}, \\ 0, & n \not\equiv 0 \pmod{N}, \end{cases}$$

and we deduce that

$$\frac{2}{N} \sum_{k=-N/2+1}^{N/2} h\left(\frac{2k}{N}\right) = 2 \sum_{r=-\infty}^{\infty} \widehat{h}_{Nr} \, .$$

Computation of Fourier coefficients (DFT)

Hence, the error committed by the Riemann approximation is

$$
\begin{aligned}
e_N(h) \;&:=\; \frac{2}{N} \sum_{k=-N/2+1}^{N/2} h\left(\frac{2k}{N}\right) - \int_{-1}^{1} h(t)\,dt \;=\; 2\sum_{r=-\infty}^{\infty} \widehat{h}_{Nr} - 2\widehat{h}_0 \\
&=\; 2\sum_{r=1}^{\infty} \left(\widehat{h}_{Nr} + \widehat{h}_{-Nr}\right).
\end{aligned}
$$

Since $h \in \mathcal{A}$, its Fourier coefficients decay at spectral rate, namely $\widehat{h}_{Nr} = \mathcal{O}((Nr)^{-p})$, and hence the error of the Riemann sums approximation (11) decays spectrally as a function of N,

$$
e_N(h) = \mathcal{O}(N^{-p}) \quad \forall p \in \mathbb{N}\,.
$$

Computation of Fourier coefficients (DFT)

Going back to the computation of the Fourier coefficients (10), we see that we may compute the integral of $h(x) = \frac{1}{2} f(x) e^{-i\pi nx}$ by means of the Riemann sums, and this gives a spectral method for calculating the Fourier coefficients of f:

$$\widehat{f_n} \approx \frac{1}{N} \sum_{k=-N/2+1}^{N/2} f\left(\frac{2k}{N}\right) \omega_N^{-nk}, \qquad n = -N/2+1, \ldots, N/2\,. \tag{13}$$

The Poisson equation

We consider the *Poisson equation*

$$\nabla^2 u = f, \quad -1 \leq x, y \leq 1, \tag{14}$$

where f is analytic and obeys the periodic boundary conditions

$$f(-1, y) = f(1, y), \quad -1 \leq y \leq 1, \qquad f(x, -1) = f(x, 1), \quad -1 \leq x \leq 1.$$

Moreover, we add to (14) the following *periodic boundary conditions*

$$\begin{aligned} u(-1, y) = u(1, y), \quad u_x(-1, y) = u_x(1, y), \quad -1 \leq y \leq 1 \\ u(x, -1) = u(x, 1), \quad u_y(x, -1) = u_y(x, 1), \quad -1 \leq x \leq 1. \end{aligned} \tag{15}$$

With these boundary conditions alone, a solution of (14) is only defined up to an additive constant.

The Poisson equation

Hence, we add a *normalisation condition* to fix the constant:

$$\int_{-1}^{1}\int_{-1}^{1} u(x,y)\;dx\;dy = 0. \tag{16}$$

We have the spectrally convergent Fourier expansion

$$f(x,y) = \sum_{k,l=-\infty}^{\infty} \widehat{f}_{k,\ell} e^{i\pi(kx+\ell y)}$$

and seek the Fourier expansion of u

$$u(x,y) = \sum_{k,\ell=-\infty}^{\infty} \widehat{u}_{k,\ell} e^{i\pi(kx+\ell y)}.$$

Since

$$0 = \int_{-1}^{1}\int_{-1}^{1} u(x,y)\,dx\,dy = \sum_{k,\ell=-\infty}^{\infty} \widehat{u}_{k,\ell} \int_{-1}^{1}\int_{-1}^{1} e^{i\pi(kx+\ell y)}\,dx\,dy = \widehat{u}_{0,0},$$

and

$$\nabla^2 u(x,y) = -\pi^2 \sum_{k,\ell=-\infty}^{\infty} (k^2+\ell^2)\widehat{u}_{k,\ell} e^{i\pi(kx+\ell y)},$$

together with (14), we have

$$\begin{cases} \widehat{u}_{k,\ell} = -\dfrac{1}{(k^2+\ell^2)\pi^2}\widehat{f}_{k,\ell}, & k,\ell \in \mathbb{Z},\ (k,\ell) \neq (0,0) \\ \widehat{u}_{0,0} = 0. \end{cases}$$

Remark 4

Applying a spectral method to the Poisson equation is not representative for its application to other PDEs. The reason is the special structure of the Poisson equation. In fact, $\phi_{k,\ell} = e^{i\pi(kx+\ell y)}$ are the eigenfunctions of the Laplace operator with

$$\nabla^2 \phi_{k,\ell} = -\pi^2 (k^2 + \ell^2) \phi_{k,\ell},$$

and they obey periodic boundary conditions.

General second-order linear elliptic PDE

We consider the more general second-order linear elliptic PDE

$$\nabla^\top(a\nabla u) = f, \quad -1 \leq x, y \leq 1,$$

with $a(x, y) > 0$, and a and f periodic. We again impose the periodic boundary conditions (15) and the normalisation condition (16). We rewrite

$$\nabla^\top(a\nabla u) = \frac{\partial}{\partial x}(au_x) + \frac{\partial}{\partial y}(au_y) = f\,.$$

Recall that for the Fourier expansions

$$g(x, y) = \sum_{k,\ell\in\mathbb{Z}} \widehat{g}_{k,\ell}\phi_{k,\ell}(x, y), \qquad h(x, y) = \sum_{m,n\in\mathbb{Z}} \widehat{h}_{m,n}\phi_{m,n}(x, y),$$

(here the $\phi_{k,\ell}$s are the complex exponentials) we have that

$$\widehat{(g \cdot h)}_{k,\ell} = \sum_{m,n\in\mathbb{Z}} \widehat{g}_{k-m,\ell-n}\widehat{h}_{m,n}, \qquad \widehat{(g_x)}_{k,\ell} = i\pi k\,\widehat{g}_{k,\ell}\,, \quad \widehat{(g_y)}_{k,\ell} = i\pi\ell\,\widehat{g}_{k,\ell}\,,$$

$$\widehat{(h_x)}_{m,n} = i\pi m\,\widehat{h}_{m,n}\,, \quad \widehat{(h_y)}_{m,n} = i\pi n\,\widehat{h}_{m,n}\,.$$

This gives

$$-\pi^2 \sum_{k,\ell\in\mathbb{Z}} \sum_{m,n\in\mathbb{Z}} (km+\ell n)\, \widehat{a}_{k-m,\ell-n} \widehat{u}_{m,n} \phi_{k,\ell}(x,y) = \sum_{k,\ell\in\mathbb{Z}} \widehat{f}_{k,\ell} \phi_{k,\ell}(x,y)\,.$$

In the next steps, we truncate the expansions to $-N/2+1 \le k,\ell,m,n \le N/2$ and impose the normalisation condition $\widehat{u}_{0,0} = 0$. This results in a system of N^2-1 linear algebraic equations in the unknowns $\widehat{u}_{m,n}$, where $m,n = -N/2+1 ... N/2$, and $(m,n) \neq (0,0)$:

$$\sum_{m,n=-N/2+1}^{N/2} (km+\ell n)\, \widehat{a}_{k-m,\ell-n}\, \widehat{u}_{m,n} = -\frac{1}{\pi^2}\, \widehat{f}_{k,\ell}\,,\;\; k,\ell = -N/2+1 ... N/2\,.$$

Analyticity and periodicity

The fast convergence of spectral methods rests on two properties of the underlying problem: analyticity and periodicity. If one is not satisfied the rate of convergence in general drops to polynomial. However, to a certain extent, we can relax these two assumptions while still retaining the substantive advantages of Fourier expansions.

Relaxing analyticity: In general, the speed of convergence of the truncated Fourier series of a function f depends on the smoothness of the function. In fact, the smoother the function the faster the truncated series converges, i.e., for $f \in C^p(-1,1)$ we receive an $\mathcal{O}(N^{-p})$ order of convergence.

Spectral convergence can be recovered, once analyticity is replaced by the requirement that $f \in C^\infty(-1,1)$, i.e., $f^{(m)}(x)$ exists for all $x \in (-1,1)$ and $m = 0, 1, 2, \ldots$. Consider, for instance, $f(x) = e^{-1/(1-x^2)}$. Then, $f \in C^\infty(-1,1)$ but cannot be extended analytically because of essential singularities at ± 1. Nevertheless, one can show that $|\widehat{f_n}| \sim \mathcal{O}(e^{-cn^\alpha})$, where $c > 0$ and $\alpha \approx 0.44$. While this is slower than exponential convergence in the analytic case, it is still faster than $\mathcal{O}(n^{-m})$ for any integer m and hence, we have spectral convergence.

Analyticity and periodicity – Relaxing periodicity

Relaxing periodicity: Disappointingly, periodicity is necessary for spectral convergence. Once this condition is dropped, we are back to the setting of Theorem 3.3, i.e., Fourier series converge as $\mathcal{O}(N^{-1})$ unless $f(-1) = f(1)$. One way around this is to change our set of basis functions, e.g., to Chebyshev polynomials.

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,

$$T_n(x) := \cos n\theta, \quad x = \cos\theta, \quad \theta \in [0, \pi]. \tag{17}$$

The three-term recurrence relation

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

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

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 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, by setting $x = \cos\theta$,

$$(T_n, T_m)_w = \int_{-1}^{1} T_m(x)\,T_n(x)\,\frac{dx}{\sqrt{1-x^2}} = \int_0^{\pi} \cos m\theta \cos n\theta \, d\theta$$

$$= \begin{cases} \pi, & m = n = 0\,, \\ \frac{\pi}{2}, & m = n \geq 1\,, \\ 0, & m \neq n\,. \end{cases} \tag{18}$$

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

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

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

$$\begin{aligned} (f, T_n)_w = \breve{f}_n (T_n, T_n)_w \quad \Rightarrow \quad \breve{f}_n &= \frac{(f, T_n)_w}{(T_n, T_n)_w} \\ &= \frac{c_n}{\pi} \int_{-1}^{1} f(x) T_n(x) \frac{dx}{\sqrt{1-x^2}}, \end{aligned} \tag{19}$$

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

https://en.wikipedia.org/wiki/Stone-Weierstrass_theorem
https://en.wikipedia.org/wiki/Chebyshev_polynomials