Spectral Methods

Abstract: First, theoretical introductions to interpolation and spectral expansion are given and a particular emphasis is put on the fast convergence of the spectral approximation. I will present different approaches to solve differential equations, limiting ourselves to the one-dimensional case, with one or several domains.

Content

- Motivation
- Introduction
- Best Polynomial Approximation
- Interpolation
- **Weighted Residual Method**
- Multi-domain
- Non-linear Case
- Conclusion

Motivation

- Non-locality of the spectral representation.
- The fast convergence of the spectral approximation.

Introduction

When doing simulations and solving ODEs in $[-1,1]$, *finite-differences (FD)methods:*

Set up a grid:

$$\{x_i\}_{i=0,\ldots,N} \in [-1,1] \tag{1}$$

Represent f

$$\{f_i = f(x_i)\}_{i=0,\ldots,N} \tag{2}$$

Represent the derivative

$$\forall i < N, f_i' = f'(x_i) \simeq \frac{f_{i+1} - f_i}{x_{i+1} - x_i} \tag{3}$$

Decay as $\Delta x = 1/N$

Introduction

Spectral methods:

$$f(x) \simeq \sum_{i=0}^{N} c_i \Phi_i(x) \tag{4}$$

basis functions or trial functions $\{\Phi_i\}_{i=0,\ldots,N}$.

For a ODE system,

$$\begin{cases} Lu(x) = s(x), x \in U \\ Bu(x) = 0, x \in \partial U \end{cases} \tag{5}$$

L is a **linear** operator;

B is the operator defining the boundary conditions;

s is the source term.

residual $R = L\bar{u} - s$ is small  $\bar{u}$ is a *numerical solution*

Introduction

Q:How to express the smallness of residual?

A:Weighted Residual

$$\forall i = 0, \ldots, N, \quad (\xi_i, R)_\omega = 0 \tag{6}$$

choose a set of *test functions* $\{\xi_i(x)\}_{i=0\ldots N}$.

Best Polynomial Approximation

Polynomials are the functions that a computer can exactly evaluate.

Try to approximate any function by a polynomial.

Given a continuous function f, the best polynomial approximation:$P_N^\star(x)$,

$$\|f - P_N^*\|_\infty = \min\left\{\|f - P\|_\infty, P \in \mathbb{P}_N\right\} \tag{7}$$

If and only if there exist $N+2$ points x_i, s.t.

$$f(x_i) - p_N^\star(x_i) = (-1)^{i+\delta}\|f - p_N^\star\|_\infty, \quad \text{with } \delta = 0 \text{ or } 1 \tag{8}$$

$p_N^\star$ should be the **interpolation approximation**

Interpolation

Continuous function f, grids X with $N+1$ nodes x_i, a unique polynomial of degree N, $I_N^X f$:

$$I_N^X f(x_i) = f(x_i), \quad 0 \le i \le N \tag{9}$$

Alternatively,

$$\begin{aligned} I_N^X f &= \sum_{i=0}^{N} f(x_i)\, \ell_i^X(x) \\ &= \sum_{i=0}^{N} \frac{\omega_{N+1}^X(x)}{(x - x_i)\, \omega_{N+1}'^X(x_i)} f(x_i) \end{aligned} \tag{10}$$

where

$$\ell_i^X(x) = \prod_{j=0, j\neq i}^{N} \frac{x - x_j}{x_i - x_j}, \quad \omega_{N+1}^X(x) = \prod_{i=0}^{N} (x - x_i) \tag{11}$$

Interpolation

Compare to the best approximation:

$$\left\| f - I_N^X f \right\|_\infty \leq (1 + \Lambda_N(X)) \left\| f - p_N^* \right\|_\infty \tag{12}$$

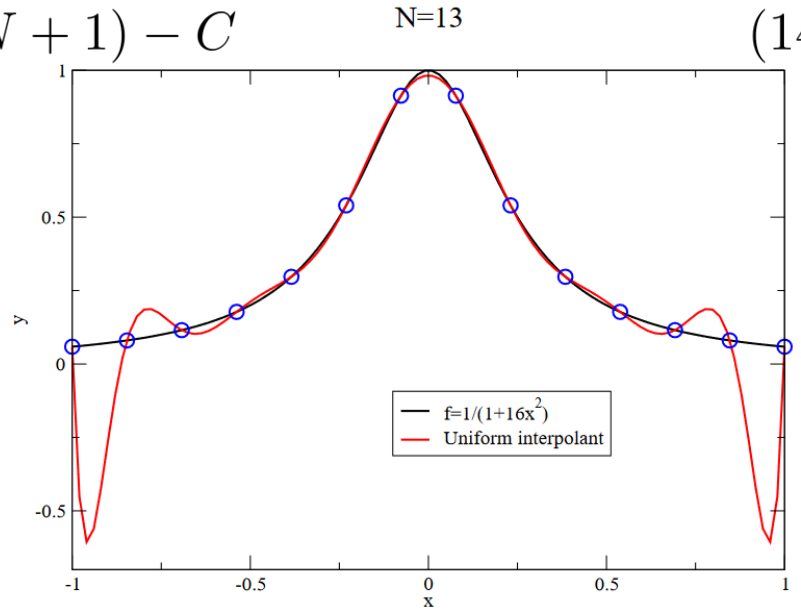
where Λ_N is the *Lebesgue constant* of the grid X,

$$\Lambda_N(X) = \max_{x \in [-1,1]} \sum_{i=0}^{N} \left| \ell_i^X(x) \right|. \tag{13}$$

For any choice of grid X, there exists a constant $C > 0$, such that:

$$\Lambda_N(X) > \frac{2}{\pi} \ln(N+1) - C \tag{14}$$

Runge phenomenon



Interpolation

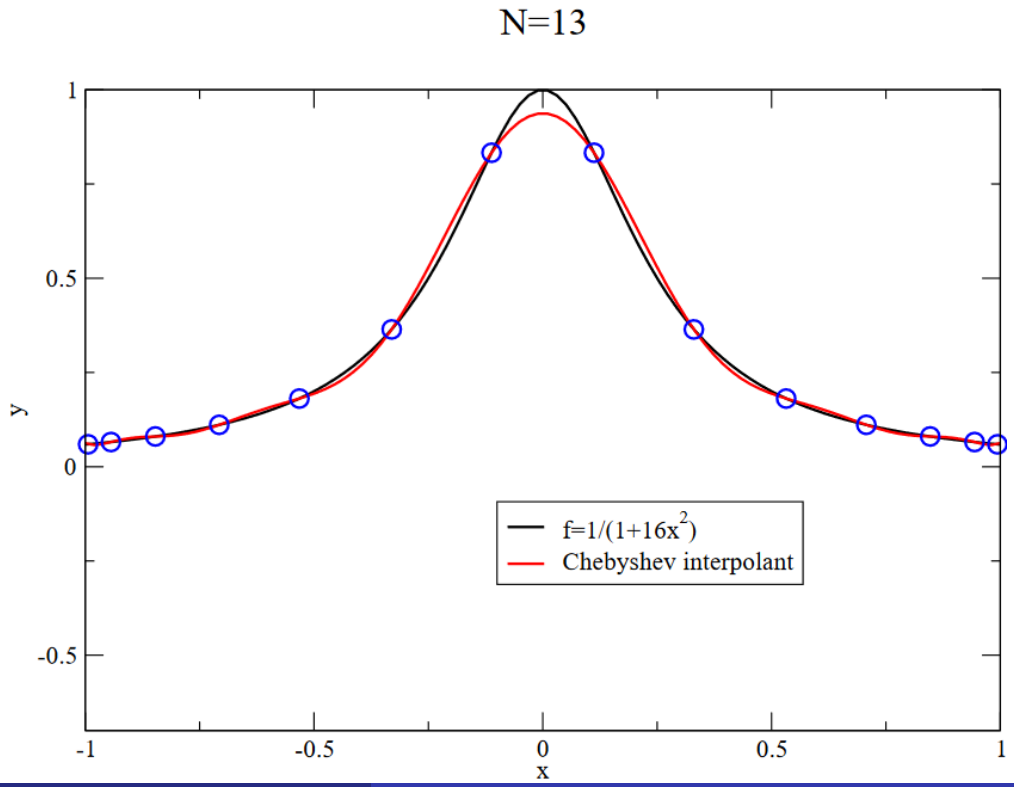
However, one can hope to minimize the interpolation error by choosing a grid such.

$$f(x) - I_N^X f(x) = \frac{f^{N+1}(\epsilon(x))}{(N+1)!}\omega_{N+1}^X(x) \tag{15}$$

For example, we can choose a grid which nodes are **the zeros of the Chebyshev polynomial** $T_{N+1}(x)$.

$$\left\|\omega_{N+1}^X\right\|_\infty = \frac{1}{2^N}$$

Grid points are important



Interpolation

A basis of $\mathbb{P}_N = \{p_n\}_{n=0,\dots N}$. For orthogonal polynomials, $\left(p_i, p_j\right)_\omega = 0, i \neq j$

$$(f,g)_w = \int_{x\in[-1,1]} f(x)g(x)w(x)\mathrm{d}x \tag{16}$$

We can project function f on this basis,

$$P_N f = \sum_{n=0}^{N} \hat{f}_n p_n \tag{17}$$

where the coefficients are

$$\hat{f}_n = \frac{(f,p_n)}{(p_n,p_n)} \tag{18}$$

Which requires the evaluation of f at a great number of points, making the whole numerical scheme impractical.

Gaussian quadrature

Interpolation

Given a measure w, there exist $N+1$ real numbers w_n and $N+1$ real numbers x_n, s.t.

$$\int_{[-1,1]} f(x)\omega(x)dx = \sum_{j=0}^{N} f(x_j)\omega_j + E_N[f], \tag{19}$$

When $f \in \mathbb{P}_{2N+\delta}$, $E_N[f] = 0$; x_j are the collocation points; w_j can be determined by,

$$w_j = \int_{[-1,1]} \ell_j^X(x)w(x)\mathrm{dx} \tag{20}$$

integer δ depends on the choice of quadrature.

When $\{x_n\}$ are the roots of orthogonal polynomial P_{N+1}, we have $f = rP_{N+1} + q \in \mathbb{P}_{2N+1}, r, q \in \mathbb{P}_N$ and $\delta = 1$.

$N+1$ collocation points enjoy the maximum degree of precision $2N+1$.

There are some example polynomials:

- Legendre polynomials $P_n(x)$: x_i are the nodes of $P_{N+1}(x)$,

$$w_i = \frac{2}{(1-x_i^2)\left[P'_{N+1}(x_i)\right]^2} \tag{21}$$

- Chebyshev polynomials $T_n(x)$: $T_n(x) = 2xT_n(x) - T_{n-1}(x), T_0 = 1, T_1 = x$,

$$x_i = \cos\frac{(2i+1)\pi}{2N+2}, \quad w_i = \frac{\pi}{N+1} \tag{22}$$

Interpolation

By making use of the Gaussian quadrature, define the *interpolant* of a function f by

Spectral interpolation $$I_N f = \sum_{n=0}^{N} \tilde{f}_n p_n(x) \tag{23}$$

where

$$\tilde{f}_n = \frac{1}{\gamma_n} \sum_{i=0}^{N} f(x_i)\, p_n(x_i)\, w_i \quad \text{and} \quad \gamma_n = \sum_{i=0}^{N} p_n^2(x_i)\, w_i \tag{24}$$

$\tilde{f}_n$ is that they are computed by estimating f at the $N+1$ collocation points only.

configuration space:
if the function is described by its value at the $N+1$ collocation points $f(x_i)$.

coefficient space:
if the function is described by $N+1$ coefficients $\tilde{f}_i$

Interpolation

One of the main advantage of spectral method is the very fast convergence of the interpolant $I_N f$. For a C^m function f

- For Legendre:

$$\|I_N f - f\|_{L^2} \leq \frac{C_1}{N^{m-1/2}} \sum_{k=0}^{m} \left\| f^{(k)} \right\|_{L^2} \tag{25}$$

- For Chebyshev:

$$\|I_N f - f\|_{L^2_w} \leq \frac{C_2}{N^m} \sum_{k=0}^{m} \left\| f^{(k)} \right\|_{L^2_w} \tag{26}$$

When $m \to \infty$, the difference decays faster than any power of N.

Weighted Residual Method

More about Weighted residual method:

Let us consider a differential equation,

$$Lu(x) = S(x), \quad x \in [-1, 1] \tag{27}$$

u is an admissible solution ⟹ Fulfills the boundary conditions &the residual $R = Lu - S$ is small.

For a set of $N + 1$ test functions $\{\xi_n\}_{n=0\ldots N}$ on $[-1, 1]$

$$(R, \xi_k) = \sum_{i=0}^{N} R(x_i)\xi_k(x_i)w_i = 0, \forall k \leq N \tag{28}$$

Depending on the choice of the test functions and the way the boundary conditions are enforced, there are different solvers.

Weighted Residual Method

1. Tau method

the test functions are the **basis used for the spectral expansion**.

$$\sum_{i=0}^{N} L_{ni}\tilde{u}_i = \tilde{s}_n, \quad \forall n \leq N \tag{29}$$

Impose the boundary conditions:

In the Tau-method, this is done by relaxing the last two equations in (29)(i.e. for $n = N - 1$ and $n = N$) and by replacing them by the boundary conditions at $x = -1$ and $x = 1$.

2. Collocation method

Test functions are zero at each but one collocation point.
They are indeed **the Lagrange cardinal polynomials**, i.e. $\ell_i^X(x_j) = \delta_{ij}$

$$\sum_{i=0}^{N}\sum_{j=0}^{N} L_{ij}\tilde{u}_j T_i(x_n) = S(x_n), \quad \forall n \le N \tag{30}$$

Impose the boundary conditions:

In this case, the relaxed conditions are the two associated with the outermost points, i.e. $n = 0$ and $n = N$, which are replaced by appropriate boundary conditions to get an invertible system.

3. Galerkin method

Test polynomials G_i **individually verify the boundary conditions**.
And solution u as a sum of polynomials G_i

$$u(x) = \sum_{k=0}^{K} \tilde{u}_k^G G_k(x), \text{ with } \quad G_i = \sum_{j=0}^{N} M_{ji} T_j \tag{31}$$

Galerkin system reads:

$$(Lu, G_n) = (S, G_n)$$

$$\sum_{k=0}^{N-2} \tilde{u}_k^G \sum_{i=0}^{N} \sum_{j=0}^{N} M_{in} M_{jk} L_{ij} (T_i|T_i) = \sum_{i=0}^{N} M_{in} \tilde{s}_i (T_i|T_i), \quad \forall n \leq N-2. \tag{32}$$

Multi-domain

Spectral methods are very efficient dealing with C^∞ functions.

If not?

Multi-domain techniques: in each domain, do same thing and only deal with C^∞ functions.

For simplicity the physical space is split into two domains

- First domain: $x \leq 0 \Rightarrow x_1 = 2x + 1, x_1 \in [-1,1]$:

$$I_N u(x) = \sum_{i=0}^{N} \tilde{u}_i^1 T_i\left(x_1(x)\right)$$

- Second domain: $x \geq 0 \Rightarrow x_2 = 2x - 1, x_2 \in [-1,1]$

Multi-domain

- Multi-domain tau method: when $x \leq 0$,

$$(T_n, R) = 0 \Longrightarrow \sum_{i=0}^{N} L_{ni}\tilde{u}_i^1 = \tilde{s}_n^1 \quad \forall n \leq N \tag{33}$$

- Multi-domain collocation method : when $x \leq 0$,

$$\sum_{i=0}^{N}\sum_{j=0}^{N} L_{ij}\tilde{u}_j^1 T_i\left(x_{1n}\right) = S\left(x_{1n}\right) \quad \forall n \leq N \tag{34}$$

- $N-1$ residual equations in the first domain;
- $N-1$ residual equations in the second domain;
- 2 boundary conditions;
- 2 matching conditions: solution and its first derivative at $x = 0$.

$2N+2$ equations = $2N+2$ undetermined coefficients

Non-linear Case

If L is a non-linear differential operator, we can separate L into the linear part $\hat{L}$ and non-linear part N as,

$$\begin{aligned} Lu(x) &= s(x) \\ (\tilde{L} + N)u(x) &= s(x) \\ \tilde{L}u(x) &= S(x) \end{aligned} \tag{35}$$

where $\mathrm{S}(x) = \mathrm{s}(x) - \mathrm{Nu}(x)$. It is difficult to evaluate the $\tilde{S}_i$.

And with the collocation method, the residual is

$$\begin{aligned} R &= \tilde{L}u(x_n) - S(x_n) = 0 \\ \sum_{i=0}^{N}\sum_{j=0}^{N} \tilde{L}_{ij}\tilde{u}_j T_i\left(x_n\right) &= S\left(x_n\right), \quad \forall n \leq N \end{aligned} \tag{36}$$

which is equivalent to

$$M\mathbf{u} = \mathbf{S} \tag{37}$$

Non-linear Case

We can use the Newton method to solve this system.

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} - \mathbf{J}^{-1} R(\mathbf{u}^{(k)}) \tag{38}$$

where $\boldsymbol{J} = R'\left(\boldsymbol{u}^{(k)}\right)$ is Jacobian matrix.

relaxation method:
FD+NM

Pseudo-spectral method

Conclusion

- ✓Introduce the concepts in one dimension
- ✓Weighted Residual Method to minimize error in a certain way and spectral methods in one dimension.
- ✓ Spectral method for non-linear system

Thank you for your attention and the questions