Part III: Fast Spectral-Galerkin Methods: Algorithms and Analysis

Jie Shen

Department of Mathematics
Purdue University

IMS, National University of Singapore, July 29-30, 2009

General setup

We are interested in solving nonlinear evolution equations of the form

$$u_t + Au + Bu = 0$$

where A is the leading linear differential operator and B is a lower-order, nonlinear operator.

- A can be a nonlinear elliptic operator
- Steady state problems or eigenvalue problems can also be treated.

Some Examples

Korteweg de Vries equation:

$$u_t + u_x + uu_x + \frac{1}{6}u_{xxx} = 0.$$

Cahn-Hilliard equation:

$$u_t - \Delta\{-\Delta u + \frac{1}{\epsilon^2}(|u|^2 - 1)u\} = 0.$$

(incompressible) Navier-Stokes equations:

$$u_t + (u \cdot \nabla u) - \nu \Delta u + \nabla p = f,$$
$$\nabla \cdot u = 0.$$

multiphase flows, liquid crystal flows ...

Nonlinear Schrödinger equation:

$$iu_t + \Delta u + V(x)u + \beta |u|^2 u = 0$$

and its associated eigenvalue problems:

$$\Delta u + V(x)u + \beta |u|^2 u = \lambda u.$$

Time discretization

- Treat A implicitly to avoid stiffness

- Treat B explicitly to avoid solving nonlinear equations.

One needs to solve, at each time step, a linear equation:

$$\alpha u + Au = f.$$

Hence, it is important to develop an accurate and efficient solver for this type of linear problems.

Numerical Approximations

Choose a suitable variational formulation:

Find $u \in X$, s.t.

$$\alpha < u, v > + < Au, v > = < f, v >, \quad \forall v \in Y.$$

Construct finite dimensional spaces X_N and Y_N:

Find $u_N \in X_N$, s.t.

$$\alpha < u_N, v_N >_N + < A_N u_N, v_N >_N = < f, v_N >_N, \quad \forall v \in Y_N.$$

Spectral methods vs. FD/FE methods

FE methods look for $u_h(x) = \sum_{i=0}^{N} \hat{a}_i \hat{\phi}_i^h(x)$ where $\{\hat{\phi}_i^h\}$ are **local** piecewise polynomials.

$$\|u - u_h\|_{L^2} \leq Ch^\alpha, \quad \alpha = \text{ 2, 3 or 4 usually.}$$

Spectral methods look for $u_N(x) = \sum_{i=0}^{N} a_i \psi_i(x)$ where $\{\psi_i\}$ are **global** polynomials.

$$\|u - u_N\|_{L^2} \leq \begin{cases} CN^{-\alpha} \|\partial_x^\alpha u\|_{L^2_\omega} & \text{if } \partial_x^\alpha u \in L^2_\omega(I) \\ Ce^{-\beta N} & \text{if } u \text{ is analytic} \end{cases}.$$

Two Common Misconceptions about Spectral Methods

- Spectral methods always lead to dense, ill-conditioned matrices.

  Solution: Choose appropriate basis functions

- Spectral methods can only be applied to simple geometries.

  Solution:(i) mapping, (ii) perturbation, (iii) spectral-elements

Choice of basis functions

- Fourier series — only appropriate for periodic problems; Gibbs phenomena occurs if applied to non-periodic problems;
- Jacobi polynomials ($J_n^{(\alpha,\beta)}$ with $\alpha, \beta > -1$) — eigenfunctions of singular Sturm-Liouville problem satisfying:

$$\int_{-1}^{1} J_n^{(\alpha,\beta)} J_m^{(\alpha,\beta)} (1-x)^\alpha (1+x)^\beta dx = 0 \quad \text{if} \quad n \neq m.$$

Two important special cases: $\alpha = \beta = 0$ and $\alpha = \beta = -\frac{1}{2}$

Basic Approximation Theorem

Let $\omega^{\alpha,\beta} = (1-x)^{\alpha}(1+x)^{\beta}$.

Let $\pi_N^{\alpha,\beta} : L^2_{\omega^{\alpha,\beta}} \to P_N$ be the orthogonal projector defined by $(u - \pi_N^{\alpha,\beta} u, v_N)_{\omega^{\alpha,\beta}} = 0, \ \forall v_N \in P_N$.

Theorem. (Funaro, Babuska & Guo, etc.)

$$\|\partial_x^l (u - \pi_N^{\alpha,\beta} u)\|_{\omega^{\alpha+l,\beta+l}} \lesssim N^{l-m} \|\partial_x^m u\|_{\omega^{\alpha+m,\beta+m}}.$$

Remark: This result also indicates that the spectral method is particularly suitable for problems with corner singularities and/or boundary layers.

Spectral-collocation method

Consider $Au = f,$ in $\Omega = (-1,1)^d;\ \ Bu|_{\partial\Omega} = 0.$

- Σ_N: a set of collocation (Gauss-type) points;
- $X_N = \{v \in P_N^d : Bv|_{\partial\Omega} = 0\}.$

Spectral-collocation: Find $u_{_N} \in X_N$ such that

$$Au_{_N}(\boldsymbol{x_k}) = f(\boldsymbol{x_k}),\ \boldsymbol{x_k} \in \Sigma_N.$$

- The corresponding basis functions: Lagrange polynomials;
- Easy to implement but leads to full, ill-conditioned matrix A_N: $cond(A_N) \sim N^{2m}$ (m – order of the equation).

Key to efficiency and better conditioning

Use *compact combination* of orthogonal polynomials as basis functions!

Let k be the order of the PDE, we use

$$\phi_i(x) = \sum_{j=i}^{i+k} a_j p_j(x)$$

where $\{a_j\}$ are chosen s.t. $\{\phi_i(x)\}$ satisfy the underlying homogeneous B.C.

Such an approach can be viewed as Finite Elements in frequency space !

Advantages:

- sparse linear systems for problem with constant or polynomial coefficients;
- The resulting linear systems are well-conditioned.

Legendre ($\omega \equiv 1$) Galerkin Method for Even-order Equations

Consider the 1-D equation

$$\alpha u - u_{xx} = f; \;\; a_{\pm} u(\pm 1) + b_{\pm} u_x(\pm 1) = 0.$$

Let us denote $X_N = \{u \in P_N : a_{\pm} u(\pm 1) + b_{\pm} u_x(\pm 1) = 0\}$. We can determine unique (a_k, b_k, c_k) such that

$$\phi_k(x) = a_k L_k(x) + b_k L_{k+1}(x) + c_k L_{k+2}(x) \in X_N, \; 0 \le k \le N-2,$$

$$(D_x \phi_i, D_x \phi_j) = \delta_{ij}.$$

Legendre-Galerkin method:

Find $u_N = \sum_{k=0}^{N-2} \tilde{u}_k \phi_k$ such that

$$\alpha(u_N, \phi_j) + (u_N', \phi_j') = (I_N f, \phi_j), \ j = 0, 1, \cdots, N-2.$$

Let us denote

$$B = (b_{ij}), \quad b_{ij} = (\phi_i, \phi_j) = 0 \text{ for } |j-i| > 2,$$
$$\bar{u} = (\tilde{u}_0, \tilde{u}_1, \cdots, \tilde{u}_{N-2})^T,$$
$$\bar{f} = (f_0, f_1, \cdots, f_{N-2})^T \text{ with } f_i = (I_N f, \phi_i).$$

Then, we have

$$(\alpha B + I)\bar{u} = \bar{f}.$$

Consider now the 2-D case:

$$\alpha u - \Delta u = f \ (x, y) \in \Omega = (-1, 1)^2; \ \ u|_{\partial\Omega} = 0.$$

Setting

$$X_N = \text{span}\{\phi_i(x)\phi_j(y) : i, j = 0, 1, \cdots, N-2\},$$

$$u_N = \sum_{k,j=0}^{N-2} u_{kj}\phi_k(x)\phi_j(y), \ \ f_{kj} = (I_N f, \phi_k(x)\phi_j(y))$$

$$U = (u_{kj}), \ \ F = (f_{kj}), \ \ B = (b_{ij}).$$

Then, the Legendre-Galerkin approximation becomes:

$$\alpha BUB + UB + BU = F$$

Remark: The algebraic system has the same structure as a bilinear finite element method applied to the Helmholtz equation.

- Matrix Decomposition (Lynch, Rice and Thomas 1964):

  Let $E^{-1}BE = \Lambda$ and $U = EV$, the equation becomes:

$$\alpha E\Lambda VB + EVB + E\Lambda V = F.$$

  Since $E^T = E^{-1}$, we have

$$\alpha\Lambda VB + VB + \Lambda V = E^T F \equiv G.$$

  Total operation counts: $2N^3 + O(N^2)$

- Generalized Cyclic Reduction (Swarztrauber & Sweet '74) or divide-and-conquer algorithm (Gu and Eisenstat '96): $O(N^2 \log N)$

N=M=	LG	CLG	FISHPACK	MG (6 digits)
64	0.032	0.028	0.014	0.039
128	0.22	0.14	0.068	0.16
256	1.59	0.74	0.32	1.13

Table 1: CPU time for solving a 2-D Poisson equation

Spectral-Galerkin Method for Odd-Order Equations

Consider the model third-order equation (e.g., needed for solving the KDV equation):

$$\alpha u - \beta u_x + u_{xxx} = f, \quad x \in I = (-1, 1),$$
$$u(\pm 1) = u_x(1) = 0.$$

- Collocation: condition number grows like N^6
- What is an appropriate variational formulation?

Since the leading differential operator is not symmetric, it is natural to use a Petrov-Galerkin method.

Trial space:

$$V_N = \{u \in P_N : u(\pm 1) = u_x(1) = 0\}.$$

How to choose a test space W_N such that

- the Petrov-Galerkin system is well posed;
- it leads to optimal error estimates;
- the resulting linear system is well conditioned and easy to solve.

The natural choice — dual space of V_N:

$$V_N^* = \{u \in P_N : u(\pm 1) = u_x(-1) = 0\}.$$

The dual-Petrov-Galerkin method is: find $u_{_N} \in V_N$ s.t.

$$\alpha(u_{_N}, v_{_N}) - \beta(\partial_x u_{_N}, v_{_N}) + (\partial_x u_{_N}, \partial_x^2 v_{_N}) = (f, v_{_N}), \quad \forall v_{_N} \in V_N^*.$$

Basis functions for V_N and V_N^*

$$\phi_k(x) = L_k(x) - \frac{2k+3}{2k+5}L_{k+1}(x) - L_{k+2}(x) + \frac{2k+3}{2k+5}L_{k+3}(x) \in V_{k+3},$$

$$\psi_k(x) = L_k(x) + \frac{2k+3}{2k+5}L_{k+1}(x) - L_{k+2}(x) - \frac{2k+3}{2k+5}L_{k+3}(x) \in V_{k+3}^*.$$

Therefore,

$$V_N = \text{span}\{\phi_0, \phi_1, \cdots, \phi_{N-3}\};$$

$$V_N^* = \text{span}\{\psi_0, \psi_1, \cdots, \psi_{N-3}\}.$$

Hence, by setting

$$m_{ij} = (\phi_j, \psi_i), \quad p_{ij} = -(\phi_j', \psi_i), \quad s_{ij} = (\phi_j', \psi_i''),$$

the dual-Petrov-Galerkin system reduces to

$$(\alpha M + \beta P + S)\bar{u} = \bar{f}.$$

Thanks to the orthogonality of the Legendre polynomials, we have

$$s_{ij} = (\phi_j', \psi_i'') = (\phi_j''', \psi_i) = -(\phi_j, \psi_i''') = 0, \quad \text{if} \quad i \neq j;$$
$$m_{ij} = 0, \ |i-j| > 3; \ p_{ij} = 0, \ |i-j| > 2.$$

Analysis of spectral-Galerkin method

It is easy to show that

$$\int_I \phi_k \phi_j \omega^{-2,-1} dx = 0, \ k \neq j.$$

i.e., $\{\phi_j\}$ can be viewed as the **generalized Jacobi polynomials** with index $(-2, -1)$.

As we will see later, other equations will lead to generalized Jacobi polynomials with different negative indexes.

Hence, it is useful to develop a general theory for the generalized Jacobi polynomials.

Generalized Jacobi polynomials

We can define the generalized Jacobi polynomials $\{J_n^{k,l}\}$ with k, l being any negative integers such that $\int_I J_n^{k,l}(x) J_m^{k,l} \omega^{k,l}(x) dx = 0 \ \forall\, n \neq m,$ and

$$\partial_x J_n^{k,l}(x) \sim J_{n-1}^{k+1,l+1}(x).$$

Theorem. (Guo, S. & Wang '06) Let k, l, j, m be any integers with $m \geq j \geq 0$, we have

$$\|\partial_x^j (u - \pi_N^{k,l} u)\|_{\omega^{k+j,l+j}} \lesssim N^{j-m} \|\partial_x^m u\|_{\omega^{k+m,l+m}},$$

where $\pi_N^{k,l}$ is the orthogonal projector: $L^2_{\omega^{k,l}} \to P_N \cap L^2_{\omega^{k,l}}$.

Back to the third-order equation

Lemma. ("double orthogonality" and "coercivity")

$$((u - \pi_N^{-2,-1} u)_x, \partial_x^2 v_N) = -((u - \pi_N^{-2,-1} u), \omega^{2,1} \partial_x^3 v_N)_{\omega^{-2,-1}}$$
$$= 0, \quad \forall v_N \in V_N^*.$$

$$\frac{1}{9} \|u_x\|^2_{\omega^{-2,0}} \leq (u_x, (u\omega^{-1,1})_{xx}) \leq 3 \|u_x\|^2_{\omega^{-2,0}}, \ \forall u \in V_N.$$

Theorem. (S. '03)

$$\|u - u_N\|_{\omega^{-1,1}} + N^{-1} \|\partial_x (u - u_N)\|_{\omega^{-1,0}} \lesssim N^{-m} \|\partial_x^m u\|_{\omega^{m-2,m-1}}.$$

Remark. The coercivity result implies that we can scale the basis functions such that the linear system is well-conditioned.

Extensions

- The dual-Petrov-Galerkin method can be used for other odder-order equations: $J_n^{-1,0}$ or $J_n^{0,-1}$ (for first-order) and $J_n^{-3,-2}$ or $J_n^{-2,-3}$ (for fifth-order), etc.

- The proper basis functions of the spectral-Galerkin method for for $2m$-th order equations are $J_n^{-m,-m}$.

- The GJPs can also be defined for non-integer cases which are useful for singular problems and more general Jacobi approximations (Guo, S., Wang '07). The GJPs with indexes $(-\frac{1}{2}, -\frac{3}{2})$,

$(-\frac{3}{2}, -\frac{3}{2})$ have been used recently for Chebyshev approximation to hyperbolic equations (S. & Wang '06).

- For problems with variable coefficients, one can use the spectral solver with constant coefficients as a preconditioner.

- Unbounded domains: Hermite or (generalized) Laguerre functions.

- Multidimensional separable domains: use tensorial basis functions and the matrix decomposition/diagonalization method. A direct, parallel spectral-element solver is developed recently (Y. Kwan & S. '07).

- The spectral-Galerkin methods using GJPs as basis functions enjoy (i) optimal error estimates; (ii) well-conditioned linear systems (sparse with constant or polynomial coefficients).

- Coupled with the fast transforms for general orthogonal polynomials (cf. Roklin & Tygert ’06), the total computational complexity for a large class of PDEs can be made quasi-optimal.

Variable coefficients, complex geometries

We assume that Ω can be mapped onto the reference domain $\hat{\Omega} = (-1,1)^d$, otherwise a proper domain decomposition technique should be employed with fast spectral algorithm as subdomain solvers.

Consider the elliptic equation on the reference domain

$$-\text{div}\,(A(\boldsymbol{x})\nabla u) = f, \text{ in } \hat{\Omega}; \quad u|_{\partial\hat{\Omega}} = 0.$$

We assume

$$0 < \alpha \leq \frac{(A(\boldsymbol{x})\bar{u},\bar{u})}{(\bar{u},\bar{u})} \leq \beta, \quad \forall\, \boldsymbol{x} \in \hat{\Omega},\ \bar{u} \in R^d.$$

Preconditioning in frequency space: Notice that

$$(A\nabla u, \nabla u) \sim (\nabla u, \nabla u)$$

Hence, we can use the fast algorithm for

$$(\nabla u, \nabla v) = (f, v)$$

to precondition the system

$$(A\nabla u, \nabla v) = (f, v).$$

$a(x,y)$	1	$2500(1-x^2)(1-y^2)$	x^2+y^2	$(1+x^2+y^2)^4$
7×7	3	10	8	9
15×15	4	13	9	8
31×31	3	16	11	7
63×63	3	6	12	6
127×127	2	6	20	6
255×255	2	6	18	4

Table 2: Iteration Counts for the PCG Method: using MG preconditioner with variable $a(x,y)$