Numerical Solution Algorithms for Discrete Partial Differential Equations with Random Data

Howard Elman
University of Maryland

Overview

Stochastic diffusion equation $-div\cdot(a\, grad\, u) = f \quad \text{in } \mathcal{D} \subseteq \mathbb{R}^d$

$$a(x,\underline{\xi}) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r}\, a_r(x)\xi_r$$

Our goal: Explore solution methods derived from multigrid for this problem

1. Spatial multigrid applied to stochastic Galerkin system (E. & Furnival)
2. Mean-based preconditioner applied to stochastic Galerkin system (Powell & E.)
3. Mean-based preconditioner applied to stochastic collocation systems (Xiu, Hesthaven, Babuŝka, Nobile, Tempone, Webster)
4. Comparison of performance for latter two strategies (E., C. Miller, E. Phipps & R. Tuminaro)

Problem Statement

Diffusion equation

$$-\nabla\cdot(a\nabla u) = f \quad \text{in } \mathcal{D}\subset R^d$$

$$u = g_D \text{ on } \partial\mathcal{D}_D, \quad (a\nabla u)\cdot n = 0 \text{ on } \partial\mathcal{D}_N = \partial\mathcal{D}\setminus\partial\mathcal{D}_D$$

Assume coercivity: $0<\alpha_1\le a\le\alpha_2<\infty \Longrightarrow$ Well-posed

$$a(x,\underline{\xi}) = a_0(x) + \sigma\sum_{r=1}^{m}\sqrt{\lambda_r}\,a_r(x)\xi_r$$

$\{\xi_r\}$ random parameters, with density functions $\{\rho_r(\xi_r)\}$

uncorrelated, independent, joint density

$$\rho(\underline{\xi}) = \rho_1(\xi_1)\rho_2(\xi_2)\cdots\rho_m(\xi_m)$$

$a_0(x) = \mu(x) = E(a(x,\cdot))$ mean

Stochastic Galerkin: Solve one large system $A\underline{u}=\underline{f}$

$$A = G_0\otimes A_0 + \sum_{r=1}^{m} G_r\otimes A_r$$

Stochastic Collocation: Solve multiple systems of standard structure

Multigrid Solution of Matrix Equation I (E. & Furnival)

Solving ***Au=f***

$$\boxed{A} = G_0 \otimes \boxed{A_0} + \sum_{\mathrm{r}=1}^{\mathrm{m}} G_r \otimes \boxed{A_r}$$

$$[A_r]_{jk} = \sqrt{\lambda_{\mathrm{r}}}\,\sigma \int_D a_r(x)\nabla\varphi_j(x)\cdot\nabla\varphi_k(x)\,dx,$$

$$[G_r]_{lq} = \int_\Omega \psi_l(\xi)\psi_q(\xi)\xi_r\rho(\xi)\,d\xi$$

$A_r = A_r^{(h)}, \quad A = A^{(h)},$ spatial discretization parameter *h*

$A_r = A_r^{(2h)}, \quad A = A^{(2h)},$ spatial discretization parameter 2*h*

Multigrid algorithm (two-grid)

Let $A^{(h)} = Q - N,$ Q = smoothing operator

for $i=0,1,\ldots$

for $j=1:k$ k smoothing steps

$$u^{(h)} \leftarrow (I - Q^{-1}A^{(h)})u^{(h)} + Q^{-1}f^{(h)}$$

end

$r^{(2h)} = \mathcal{R}(f^{(h)} - A^{(h)}u^{(h)})$ Restriction

Solve $A^{(2h)}c^{(2h)} = r^{(2h)}$ Coarse grid correction

$u^{(h)} \leftarrow u^{(h)} + \mathcal{P}c^{(2h)}$ Prolongation

end

Prolongation and restriction:

$\mathcal{P} = I \otimes P,$ induced by natural inclusion in spatial domain

$\mathcal{R} = \mathcal{P}^T = I \otimes R, \quad R = P^T$

Convergence analysis: use "standard" approach

$$e^{(i+1)} = [(A^{(h)})^{-1} - \mathcal{P}(A^{(2h)})^{-1}\mathcal{R})]\,[A^{(h)}(I - Q^{-1}A^{(h)})^{k}\,]\,e^{(i)}$$

Establish ***approximation property***

$$\left\| [\,(A^{(h)})^{-1} - \mathcal{P}(\,A^{(2h)})^{-1}\mathcal{R}\,]\,y \right\|_{A^{(h)}} \le C \|y\|_2 \quad \forall y$$

and ***smoothing property***

$$\left\| [A^{(h)}(I - M^{-1}A^{(h)})^{k}\,]\,y \right\|_2 \le \eta(k) \|y\|_{A^{(h)}} \quad \forall y$$

For smoothing property (Braess):

$M = \theta I$ (Richardson iteration) works with $\theta \ge \max(\lambda(A))$

In experiments:

we use *damped Jacobi*, $M = diag(A)/\omega$

Approximation property

"Standard" MG analysis for deterministic problem:

$$\left\|[(A^{(h)})^{-1}-\mathcal{P}(A^{(2h)})^{-1}\mathcal{R}]y\right\|_{A^{(h)}}=\left\|u^{(h)}-u^{(2h)}\right\|_{A^{(h)}}$$

$$=\left\|u_h-u_{2h}\right\|_a \quad (=a(u_h-u_{2h},u_h-u_{2h})^{1/2})$$

$$\leq\left\|u_h-u\right\|_a+\left\|u-u_{2h}\right\|_a$$

Approximability $\leq\sqrt{\alpha_2}\left(Ch\left\|D^2u\right\|_{L^2(\mathcal{D})}+C2h\left\|D^2u\right\|_{L^2(\mathcal{D})}\right)$

Regularity $\leq Ch\left\|f\right\|_{L^2(\mathcal{D})}$

Property of mass matrix $\leq C\left\|y\right\|_2$

For approximation property in stochastic case

Introduce *semi-discrete* space $H_0^1(\mathcal{D}) \otimes \underbrace{T_p}$

Discrete stochastic space

Weak formulation

$$a(u_p, v_p) = \ell(v_p) \quad \text{for all } v_p \in H_0^1(\mathcal{D}) \otimes T_p$$

Solution u_p

Properties

$$\|u_p - u_{hp}\|_a \equiv a(u_p - u_{hp}, u_p - u_{hp})^{1/2}$$

Approximation (in 2D):

$$\|u_p - u_{hp}\|_a \le Ch \|D^2 u_p\|_{L^2(\mathcal{D}) \otimes L^2(\Gamma)}$$

Established using best approximation property of u_{hp}
and interpolant $\tilde{u}_p(x_j, \xi) = u_p(x_j, \xi) \ \forall \xi$

Similarly for other steps used for deterministic analysis

Comments

- Establishes convergence of multigrid with rate independent of spatial discretization size *h*
- No dependence on stochastic parameters *m*, *p*
- Applies to any basis of stochastic space
- Coarse grid operator: $G = a_0 G_0 + \sigma \sum_{\mathrm{r}=1}^{\mathrm{m}} a_r \sqrt{\lambda_r} G_r,$ size O(N_ξ)

  G_r derives from basis of multivariate polynomials of total degree *p*, orthogonal wrt probability measure $\rho(\xi)d\xi$

  Maximum eigenvalue η = max root of orthogonal polynomial, bounded for bounded measure

  $$\Rightarrow a_0^{1\mathrm{x}1} - \sigma\eta\left(\sum_{\mathrm{r}=1}^{\mathrm{m}} a_r^{1\mathrm{x}1}\sqrt{\lambda_r}\right) \le \lambda(G) \le a_0^{1\mathrm{x}1} + \sigma\eta\left(\sum_{\mathrm{r}=1}^{\mathrm{m}} a_r^{1\mathrm{x}1}\sqrt{\lambda_r}\right),$$

  CG iteration is an option

Multigrid Solution of Matrix Equation II

Solving ***Au=f***

$$A = G_0 \otimes A_0 + \sum_{r=1}^{m} G_r \otimes A_r$$

$$[A_r]_{jk} = \sqrt{\lambda_r}\,\sigma \int_{\mathcal{D}} a_r(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x)\, dx,$$

$$[G_r]_{lq} = \int_{\Omega} \psi_l(\xi) \psi_q(\xi) \xi_r \rho(\xi)\, d\xi$$

Preconditioner for use with CG: $Q = G_0 \otimes A_0$ (Kruger, Pellissetti, Ghanem)

$$A_0 \sim \int_{\mathcal{D}} a_0(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x)\, dx$$

Deterministic diffusion, from mean

$$G_0 = I$$

Analysis (Powell & E.)

Recall $a(x,\omega) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r}\, a_r(x)\xi_r(\omega)$

$$\longrightarrow \qquad A = G_0 \otimes A_0 + \sum_{r=1}^{\mathrm{m}} G_r \otimes A_r$$

$$Q = G_0 \otimes A_0$$

Theorem : For μ constant, the Rayleigh quotient satisfies

$$1 - \tau \le \frac{(w, Aw)}{(w, Qw)} \le 1 + \tau$$

$$\tau = (\tfrac{\sigma}{\mu})\, c(p) \sum_{r=1}^{m} \sqrt{\lambda_r}\, \| a_r \|_\infty$$

Consequence: $\kappa \le \dfrac{1+\tau}{1-\tau}$ dictates convergence of PCG

Sketch of proof $\tau = (\sigma/\mu)\, c(p) \sum_{r=1}^{m} \sqrt{\lambda_r}\, \| a_r \|_\infty$

$$A = G_0 \otimes A_0 + \sum_{\mathrm{r}=1}^{\mathrm{m}} G_r \otimes A_r$$

In spatial domain:

$$(\varphi, A_r\varphi) \sim \sigma \sqrt{\lambda_r} \int_D a_r(x) \nabla\varphi(x)\cdot\nabla\varphi(x)\, dx$$

$$\leq \sigma\sqrt{\lambda_r}\, \| a_r \|_\infty \int_D \nabla\varphi(x)\cdot\nabla\varphi(x)\, dx$$

$$= (\sigma/\mu)\sqrt{\lambda_r}\, \| a_r \|_\infty\, (\varphi, A_0\varphi)$$

From stochastic component: as above

$c(p)$ bounded by largest root of scalar orthogonal polynomial

Multigrid Variant of this Idea

Replace action of A_0^{-1} with multigrid $\longrightarrow$ preconditioner

$$Q_{MG} = G_0 \otimes A_{0,MG}$$ (Le Maitre, et al.)

Analysis: $$\frac{(w,Aw)}{(w,Q_{MG}w)} = \frac{(w,Aw)}{(w,Qw)} \underbrace{\frac{(w,Qw)}{(w,Q_{MG}w)}}_{\in[\beta_1,\beta_2]}$$

Spectral equivalence of MG approximation to diffusion operator

$$\Longrightarrow \quad \kappa \le \frac{(1+\tau)}{(1-\tau)} \frac{\beta_2}{\beta_1}$$

Current Trends: Collocation Methods

Monte-Carlo (sampling) method: find $u \in H^1_E(\mathcal{D})$ s.t.

$$\int_{\mathcal{D}} a(x, \underline{\xi}_k) \nabla u \cdot \nabla v \, dx = \int_{\mathcal{D}} f \, v \, dx$$

for all $v \in H^1_{E_0}(\mathcal{D})$ for a collection of samples $\{\underline{\xi}_k\} \in L^2_P(\Gamma)$

Collocation: (Hesthaven & Xiu, Babuška, Nobile, Tempone, Webster)

Choose $\{\underline{\xi}_k\}$ in a special way, then require discrete solution $u_{hp}(x, \underline{\xi}) \in S^E_h \otimes T_p$ to interpolate $\{u_h(x, \underline{\xi}_k)\}$

Advantages (vs. stochastic Galerkin):

- decouples algebraic system (like MC)
- applies in a straightforward way to nonlinear random terms

Disadvantage: dimensionality $\sim 2^p$ (Galerkin)

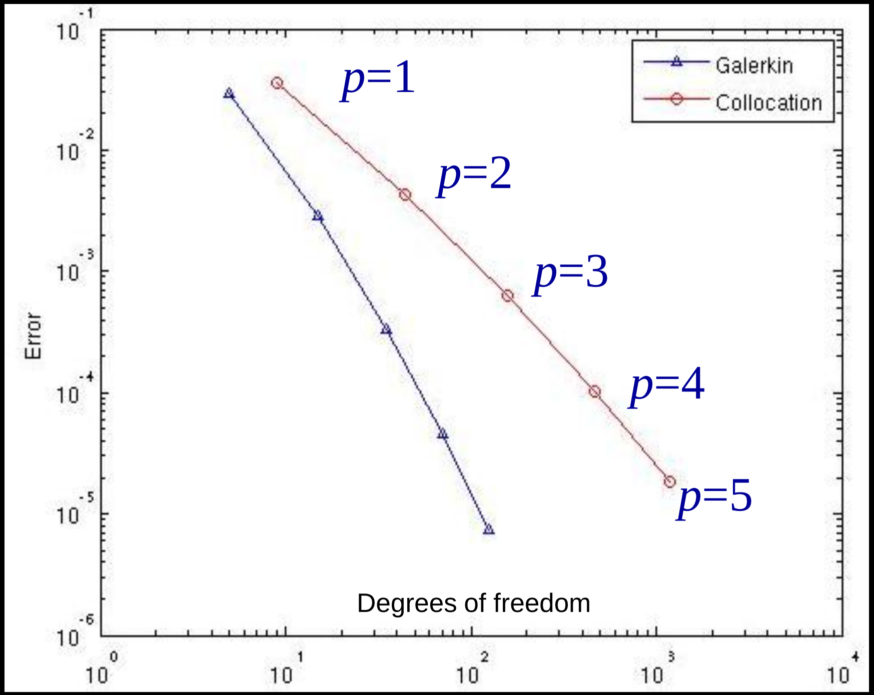
Experimental results

Accuracy:

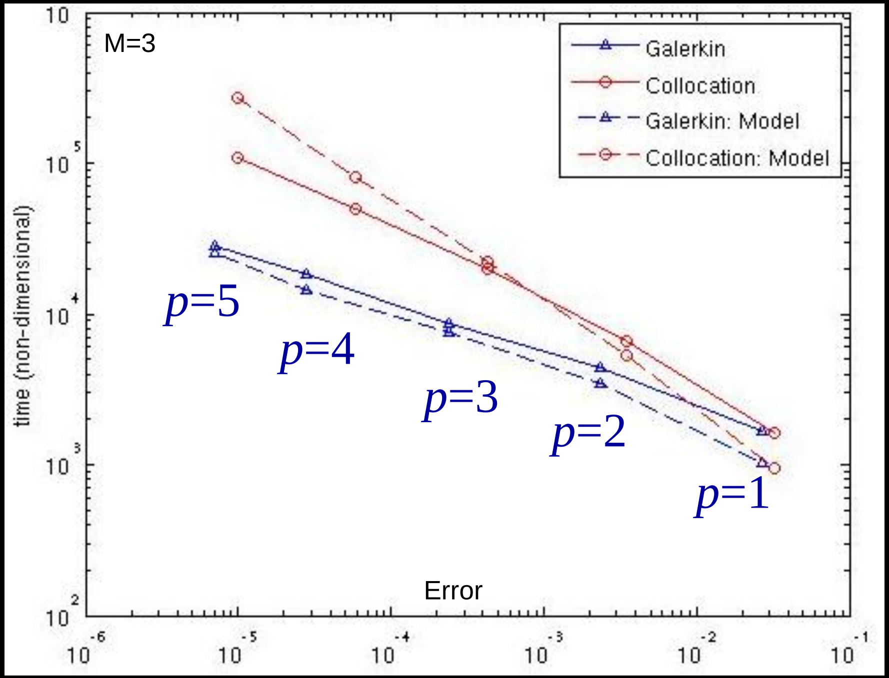
for fixed m=4: similar p=

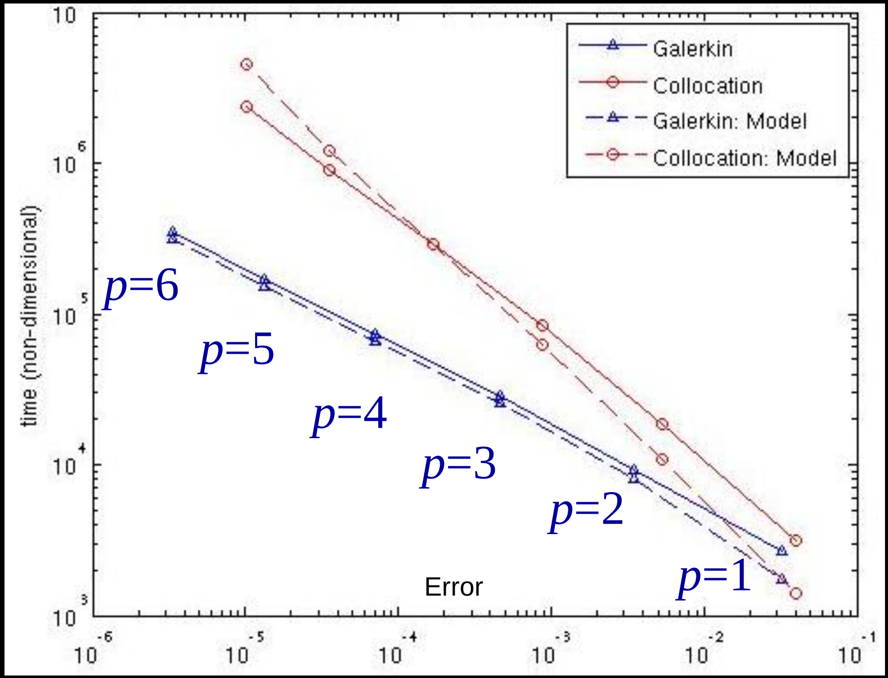
polynomial degree for SG
“level” for collocation

produces comparable errors



Performance:





Experimental results: performance

Performed on a serial machine with C code and CG/AMG code from Trilinos

Observation: Galerkin faster, more so as number of stochastic variables (KL terms) grows

CPU times for larger *m* = #KL terms:

	Galerkin			Collocation		
p	m=5	m=10	m=12	m=5	m=10	m=12
1	.058	.147	.263	.069	.163	.218
2	.269	1.20	2.00	.532	2.13	3.17
3	1.20	13.14	24.50	2.41	16.99	29.31
4	3.50	53.79	121.61	8.31	102.60	200.94
5	6.51	117.73		24.56	515.75	

Conclusions

Several variants of multigrid are effective for solving systems required for stochastic Galerkin

With MG solution: SG is competitive with collocation