Iterative solvers for Stochastic Galerkin discretizations of PDEs with random data

Elisabeth Ullmann

Institut für Numerische Mathematik und Optimierung
TU Bergakademie Freiberg, Germany

Joint work with: Catherine Powell (Manchester), Oliver Ernst (Freiberg)

ICMS Workshop on Uncertainty Quantification, May 2010, Edinburgh, UK

Groundwater flow problem: lognormal diffusion

$$
\begin{aligned}
-\nabla \cdot (\kappa \nabla u) &= f && \text{in } \; D \subset \mathbb{R}^2 \\
u &= 0 && \text{on } \; \partial D
\end{aligned}
$$

Uncertainty: $\kappa \colon \; D \times \Omega \to \mathbb{R}$ is a **lognormal** random field

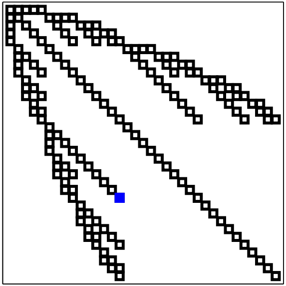
$$
\mathbb{E}\left[\int_D \kappa \nabla u \cdot \nabla v \, dx\right] = \mathbb{E}\left[\int_D f v \, dx\right] \quad \forall v \in H_0^1(D) \otimes L_P^2(\Omega)
$$

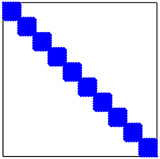
Discretization scheme: Stochastic Galerkin method [GHANEM & SPANOS]

u, v: (Wiener-Hermite) polynomial chaos of order d in M variables

κ: polynomial chaos expansion of order $2d$ in M variables

Structure of Galerkin matrix





$$\widehat{A} = \sum_\alpha G_\alpha \otimes K_\alpha$$

Finite element matrices: $[K_\alpha]_{i,k} = \int_D \kappa_\alpha \nabla \phi_i \cdot \nabla \phi_k \, dx$

Stochastic Galerkin matrices: $[G_\alpha]_{j,\ell} = \mathbb{E}\left[\psi_\alpha \psi_{\iota(j)} \psi_{\iota(\ell)}\right]$

Properties of the Galerkin matrix $\widehat{A} = \sum_\alpha G_\alpha \otimes K_\alpha$

- symmetric, positive-definite (s.p.d.)
- block-dense with sparse blocks [MATTHIES & KEESE]
- # of terms in sum is $\binom{M+2d}{2d} = n_{\boldsymbol{\xi}} \cdot n$, $n \ll n_{\boldsymbol{\xi}}$
- the cost for $\widehat{A}\boldsymbol{x}$ is at least $O(n_{\boldsymbol{x}} n_{\boldsymbol{\xi}}^2)$ / at most $O(n n_{\boldsymbol{x}} (n_{\boldsymbol{\xi}}^2 + n_{\boldsymbol{\xi}}^3))$
- ill-conditioned w.r.t. the FE mesh size h, the degree d, and κ_α

Mean stiffness matrix: $[K_0]_{i,k} = \int_D \mathbb{E}[\kappa] \nabla\phi_i \cdot \nabla\phi_k \, dx$

Mean-based preconditioner: $\widehat{P}_0 = I \otimes K_0$

Result: the spectrum of $\widehat{P}_0^{-1}\widehat{A}$ is independent of h [POWELL & ELMAN]

Properties of the Galerkin matrix $\widehat{A} = \sum_\alpha G_\alpha \otimes K_\alpha$

- symmetric, positive-definite (s.p.d.)
- block-dense with sparse blocks [MATTHIES & KEESE]
- # of terms in sum is $\binom{M+2d}{2d} = n_{\boldsymbol{\xi}} \cdot n$, $n \ll n_{\boldsymbol{\xi}}$
- the cost for $\widehat{A}\mathbf{x}$ is at least $O(n_{\mathbf{x}} n_{\boldsymbol{\xi}}^2)$ / at most $O(n n_{\mathbf{x}} (n_{\boldsymbol{\xi}}^2 + n_{\boldsymbol{\xi}}^3))$
- ill-conditioned w.r.t. the FE mesh size h, the degree d, and κ_α

Mean stiffness matrix: $[K_0]_{i,k} = \int_D \mathbb{E}[\kappa] \nabla \phi_i \cdot \nabla \phi_k \, dx$

Mean-based preconditioner: $\widehat{P}_0 = I \otimes K_0$

Result: the spectrum of $\widehat{P}_0^{-1} \widehat{A}$ is independent of h [POWELL & ELMAN]

Kronecker product preconditioner

Given $\widehat{A}$, K_0, compute $G = \operatorname{argmin}\{H \in \mathbb{R}^{n_\xi \times n_\xi} : \|\widehat{A} - H \otimes K_0\|_F\}$

Closed form solution: $G = \sum_\alpha \frac{\operatorname{tr}(K_\alpha^\top K_0)}{\operatorname{tr}(K_0^\top K_0)} G_\alpha$ [VAN LOAN & PITSIANIS]

Preconditioner: $\widehat{P}_1 = G \otimes K_0$ [U., SISC '10]

Cost for $\widehat{P}_1^{-1}\boldsymbol{x}$: $O((n_\xi + n_\xi^2)n_x)$ Set-up cost: $O(\#\text{ of terms} \cdot n_x)$

Performance: not robust w.r.t. σ (standard deviation of $\log \kappa$) and d (polynomial chaos order) but far more robust than mean-based preconditioner

Groundwater flow problem: lognormal diffusion

$$\begin{aligned} \kappa^{-1}\boldsymbol{q} + \nabla u &= 0 \\ \nabla \cdot \boldsymbol{q} &= f \qquad && \text{in } D \subset \mathbb{R}^2 \\ u &= 0 && \text{on } \partial D \end{aligned}$$

Random field: $\kappa^{-1}(x, \boldsymbol{\xi}) = \sum_\alpha \kappa_\alpha(x)\psi_\alpha(\boldsymbol{\xi})$

Discretization: $RT_0 - P_0$ finite elements & Stochastic Galerkin method

Galerkin matrix: $\begin{bmatrix} I \otimes A_0 + \sum_\alpha G_\alpha \otimes A_\alpha & I \otimes B^T \\ I \otimes B & 0 \end{bmatrix}$

RT_0 mass matrices $\quad A_0, A_\alpha \in \mathbb{R}^{N_{\boldsymbol{q}} \times N_{\boldsymbol{q}}}$:

$$[A_\alpha]_{i,k} = \int_D \kappa_\alpha \boldsymbol{\varphi}_i \cdot \boldsymbol{\varphi}_k \, dx,$$

$$[A_0]_{i,k} = \int_D \mathbb{E}[\kappa]\, \boldsymbol{\varphi}_i \cdot \boldsymbol{\varphi}_k \, dx$$

divergence operator $\quad B \in \mathbb{R}^{N_u \times N_{\boldsymbol{q}}}$:

$$[B]_{i,k} = \int_D \phi_i \, \nabla \cdot \boldsymbol{\varphi}_k \, dx$$

Block-diagonal preconditioners

The Galerkin matrix is a **saddle point matrix.**

$$\widehat{C} := \begin{bmatrix} \widehat{A} & \widehat{B}^\top \\ \widehat{B} & 0 \end{bmatrix}$$

Schur complement

$$\widehat{P}_S = \begin{bmatrix} \widehat{A} & 0 \\ 0 & \widehat{B}\widehat{A}^{-1}\widehat{B}^\top \end{bmatrix}$$

Augmented

$$\widehat{P}_A = \begin{bmatrix} \widehat{A} + \gamma^{-1}\widehat{B}^\top \widehat{W}^{-1}\widehat{B} & 0 \\ 0 & \gamma \widehat{W} \end{bmatrix}$$

Example: Schur complement preconditioners

$$\widehat{P}_S = \begin{bmatrix} \widehat{A} & 0 \\ 0 & \widehat{B}\widehat{A}^{-1}\widehat{B}^\top \end{bmatrix} \quad \Rightarrow \quad \widehat{P}_{L,amg} = \begin{bmatrix} L \otimes D_0 & 0 \\ 0 & L^{-1} \otimes V_0 \end{bmatrix}$$

$\widehat{A} \approx L \otimes D_0,\; D_0 = \mathrm{diag}(A_0),\; L \in \mathbb{R}^{n_\xi \times n_\xi}$ s.p.d.

$\widehat{B}\widehat{A}^{-1}\widehat{B}^\top \approx L^{-1} \otimes BD_0^{-1}B^\top$

$BD_0^{-1}B^\top \approx V_0$ (one AMG V-cycle)

Mean-based: $L = I$ [ERNST ET AL.]

Kronecker product: $L = G = \mathrm{argmin}\{H : \|\widehat{A} - H \otimes D_0\|_F\}$
[POWELL & U.]

Performance of Schur complement preconditioners

$D = (0,1)^2$, $f \equiv 1$, $h = 1/32$, $n_{\boldsymbol{q}} + n_u = 5,184$

$\kappa = \exp(G)$, Gaussian field G: $\mathbb{E}[G] = 1$, $\mathrm{Cov}_G(r) = \sigma_G^2\, r\, K_1(r)$, $r = \|x - y\|_2$,
$M = 5$ random variables, $n_\xi = 6 \ldots 252$, # of terms $= 21 \ldots 3,003$

P-MINRES iteration counts (and timings in sec) for $\|\boldsymbol{r}_k\|_{P^{-1}} < 10^{-8}\|\boldsymbol{r}_0\|_{P^{-1}}$

	σ_G	d=1		d=2		d=3		d=4	
$\widehat{P}_{I,amg}$	0.2	47	(2)	57	(22)	65	(245)	74	(2979)
	0.4	61	(3)	89	(35)	121	(534)	148	(6321)
	0.6	77	(3)	139	(53)	225	(1369)	345	(13794)
	0.8	99	(4)	219	(84)	425	(2604)	747	(29808)
	1.0	126	(5)	339	(129)	785	(4719)	1597	(63911)
$\widehat{P}_{G,amg}$	0.2	40	(2)	42	(13)	45	(169)	48	(1291)
	0.4	45	(2)	53	(15)	60	(229)	68	(1688)
	0.6	51	(2)	66	(18)	81	(307)	99	(2553)
	0.8	56	(2)	82	(24)	111	(419)	145	(3651)
	1.0	64	(3)	102	(28)	152	(575)	210	(5177)