Stochastic Galerkin method

Ivana Pultarová

Fakulta stavební ČVUT

PANM 2020, June 2020

Outline

Motivation example.

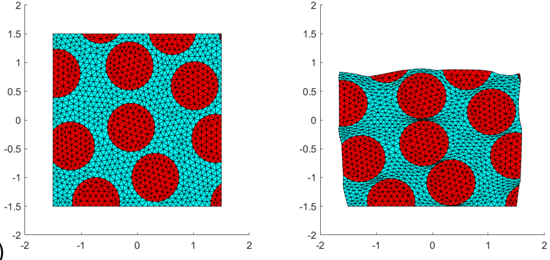
Composite material with two material characteristics $\xi = (\xi_1, \xi_2)$.

We can solve:

– dependence of displacements u on ξ for known distribution of ξ (for reliability determination),

– inverse problem - find ξ if u is given - images of original and loaded states are given

(fig. by courtesy of L. Gaynutdinova)



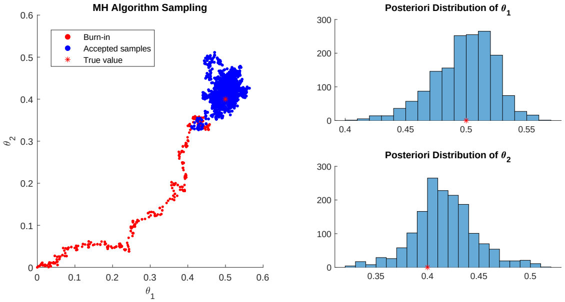
Metropolis-Hastings algorithm:

For $k = 1, 2, \ldots$

1. choose (random) parameters ξ^k
2. solve PDE with ξ^k
3. compare the error (between the solution and the goal) with the previous error
4. either accept ξ^k or accept the previous "successful" ξ^j

Post-processing: get probability distribution from the all accepted ξ's.

Motivation example - cont.
(figures by courtesy of L. Gaynutdinova)



The PDE must be solved for many parameters ξ!
Better to get *a direct dependence* of the solution $u = u(x, \xi)$ on ξ,

$$u(x,\xi) = \sum_{j=1}^{N_{\text{pol}}} \sum_{s=1}^{N_{\text{FE}}} u_{js} \psi_s(x) \phi_j(\xi)$$

(then no solution of PDE in MH algorithm).

The problem

Boundary value problem with random coefficient $a(x,\omega)$ (random field)

$$-\nabla \cdot a(x,\omega)\nabla u(x,\omega) = f(x) \qquad \text{for } (x,\omega) \in D \times \Omega,$$

$u(x,\omega) = 0$ on $\partial D \times \Omega$, where (Ω, Σ, P) is a complete probability space.

Numerical computation: discrete finite random field,
Doob–Dynkin lemma: measure space $L^2_\rho(\Gamma)$, $\rho(\xi) = \mathrm{d}P/\mathrm{d}\xi$,

$$a(x,\omega) := a(x,\xi(\omega)), \qquad \xi(\omega) = (\xi_1(\omega),\ldots,\xi_{N_\xi}(\omega)),$$

where the random variables $\xi_i(\omega)$ are iid with the joint probability density

$$\rho(\xi) = \prod_{i=1}^{N_\xi} \rho_i(\xi_i) \quad \text{and} \quad \Gamma = \prod_{i=1}^{N_\xi} \Gamma_i = \prod_{i=1}^{N_\xi} \mathrm{Im}(\xi_i).$$

R. Blaheta, M. Béreš, S. Domesová, P. Pan (Wuhan), A comparison of deterministic and Bayesian inverse with application in micromechanics, Applications of Mathematics, 2018

A. J. Crowder, Adaptive and Multilevel Stochastic Galerkin Finite Element Methods, Ph.D. Thesis, 2020

Methods:

- Monte Carlo solves

$$\int_D a(x,\xi)\nabla u(x,\xi)\nabla v(x)\,\mathrm{d}x = \int_D f(x)v(x)\,\mathrm{d}x$$

  + non intrusive; Multilevel Monte Carlo
  - large set of samples/problems; no exact error estimation

- Collocation method
  + non intrusive; sparse grids can be used
  - non-nested sets of test samples for Gauss q. or nested for Clenshaw–Curtis q.

- Stochastic Galerkin method (SGM) solves (in a subspace of $L^2_\rho(\Gamma, H^1_0(D)))$

$$\int_\Gamma \int_D a(x,\xi)\nabla u(x,\xi)\nabla v(x,\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi = \int_\Gamma \int_D f(x)v(x,\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi$$

  + error estimates available; the matrix is not constructed
  - intrusive (but double-orthogonal polynomials); large matrices; coupled problems

Variational form of a deterministic problem with N_ξ parameters:
find $u \in V = H_0^1(D) \times L_\rho^2(\Gamma) = L_\rho^2(\Gamma, H_0^1(D))$ such that for any $v \in V$

$$\int_\Gamma \int_D a(x,\xi) \nabla u(x,\xi) \nabla v(x,\xi) \rho(\xi) \mathrm{d}x \mathrm{d}\xi = \int_\Gamma \int_D f(x) v(x,\xi) \rho(\xi) \mathrm{d}x \mathrm{d}\xi,$$

where $a(x,\xi)$ is defined

- on subdomains:

$$a(x,\xi) = \sum_{k=1}^{N_\xi} \xi_k(\omega) \chi_{D_k}(x), \qquad \text{subdomains } D_k,$$

- or by Karhunen-Loéve expansion:

$$a(x,\xi) = \mu(x) + \sum_{k=1}^{\infty} \xi_k(\omega) \sqrt{\lambda_k} a_k(x),$$

where ξ_k are uncorelated random variables with zero mean and unit variance, $a_k(x)$ are eigenfunctions of the covariance operator with

$$C(g)(x) = \int_D c(x,y) g(y) \mathrm{d}y, \qquad c(x,y) = \mathrm{cov}(a(x,\omega), a(y,\omega)).$$

Two frequent forms of the coefficient function (discrete finite random field):

$$\int_\Gamma \int_D a(x,\xi)\nabla u(x,\xi)\nabla v(x,\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi \;=\; \int_\Gamma \int_D f(x)v(x,\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi .$$

or

$$\int_\Gamma \int_D \mathrm{e}^{a(x,\xi)}\nabla u(x,\xi)\nabla v(x,\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi \;=\; \int_\Gamma \int_D f(x)v(x,\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi .$$

where $a(x,\xi)$ is in the form (e.g. Karhunen-Loéve expansion):

$$a(x,\xi) \;\approx\; a_0(x) + \sum_{i=1}^{N_\xi} \xi_i a_i(x),$$

or

$$\mathrm{e}^{a(x,Y)} \;\approx\; \mathrm{e}^{a_0(x)+\sum_{i=1}^{N_\xi}\xi_i a_i(x)} \;\approx\; \sum_{k=1}^{m} \Phi_k(\xi_1,\dots,\xi_{N_\xi})\tilde{a}_k(x),$$

$\Phi_k(\xi)$ are N_ξ-variate orthogonal polynomials:

Hermite polynomials for normal distribution $\rho(\xi) \;=\; \frac{1}{\sqrt{\pi}}\mathrm{e}^{-\xi^2/2}$,

Legendre polynomials for uniform distribution $\rho(\xi) \;=\; \frac{1}{2}\chi_{\langle -1,1\rangle}$,

etc.

Discretization of $u(x,\xi) \in H_0^1(D) \times L_\rho^2(\Gamma)$:

FE basis and generalized polynomial chaos expansion (PCE) of the solution:

$$u(x,\xi) = \sum_{j=1}^{N_{\text{pol}}} \sum_{r=1}^{N_{\text{FE}}} u_{jr} \psi_r(x) \Phi_j(\xi),$$

FE basis $\{\psi_r(x)\}_{r=1}^{N_{\text{FE}}}$

N_ξ-variate orthogonal polynomials $\{\Phi_j(\xi)\}_{j=1}^{N_{\text{pol}}}$, $\quad \Phi_j(\xi) = \phi_{j,1}(\xi_1) \dots \phi_{j,N_\xi}(\xi_{N_\xi})$

complete polynomials (CP) of total order $\leq P$, $\quad N_{\text{pol}} = \binom{N_\xi + P}{N_\xi}$

tensor products (TP) pol. of orders P_k in ξ_k, $k = 1, \dots, N_\xi$, $\quad N_{\text{pol}} = \prod_{k=1}^{N_\xi} (P_k + 1)$

$N_{\text{pol}} \times N_{\text{FE}}$ basis functions $\psi_r(x)\Phi_j(\xi)$

unknowns u_{jr}

Regularity and a priori convergence estimates [I. Babuska, F. Nobile, R. Tempone, R. Ghanem, G.E. Zouraris, C.J. Gittelson, ...]

Matrix A and right hand side b:

$$A_{jr,is} = \int_\Gamma \int_D \left(a_0(x) + \sum_{k=1}^{N_\xi} \xi_k a_k(x) \right) \nabla\psi_s(x)\nabla\psi_r(x)\Phi_i(\xi)\Phi_j(\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi$$

or

$$A_{jr,is} = \int_\Gamma \int_D \sum_{k=1}^{m} \Phi_k(\boldsymbol{\xi})\tilde{a}_k(x)\nabla\psi_s(x)\nabla\psi_r(x)\Phi_i(\xi)\Phi_j(\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi$$

$$b_{jr} = \int_\Gamma \int_D f(x)\psi_r(x)\Phi_j(\xi)\rho(\xi)\,\mathrm{d}x\mathrm{d}\xi$$

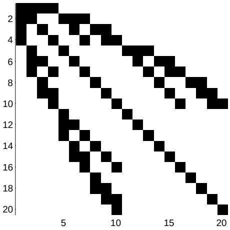
Example of linear $a(x,\xi) = a_0(x) + \xi_1 a_1(x) + \xi_2 a_2(x)$:

$N_\xi = 2$, $P = 2$, uniform distribution, Legendre approximation polynomials, CP:

$$A = \left(\begin{array}{ccc|ccc} K_0 & \frac{1}{\sqrt{3}}K_1 & \frac{1}{\sqrt{3}}K_2 & 0 & 0 & 0 \\ \frac{1}{\sqrt{3}}K_1 & K_0 & 0 & \frac{2}{\sqrt{15}}K_1 & \frac{1}{\sqrt{3}}K_2 & 0 \\ \frac{1}{\sqrt{3}}K_2 & 0 & K_0 & 0 & \frac{1}{\sqrt{3}}K_1 & \frac{2}{\sqrt{15}}K_2 \\ \hline 0 & \frac{2}{\sqrt{15}}K_1 & 0 & K_0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}}K_2 & \frac{1}{\sqrt{3}}K_1 & 0 & K_0 & 0 \\ 0 & 0 & \frac{2}{\sqrt{15}}K_2 & 0 & 0 & K_0 \end{array}\right),$$

where K_0, K_1, and K_2 are the "stiffness matrices" corresponding to $a_0(x)$, $a_1(x)$, and $a_2(x)$, respectively,

$(K_i)_{rs} = \int_D a_i(x)\nabla\psi_r(x)\nabla\psi_s(x)\,\mathrm{d}x$
$(G_i)_{jk} = \int_\Gamma \xi_i \Phi_j(\xi)\Phi_k(\xi)\rho(\xi)\,\mathrm{d}\xi$
$(G_0)_{jk} = \int_\Gamma \Phi_j(\xi)\Phi_k(\xi)\rho(\xi)\,\mathrm{d}\xi = \delta_{jk}$
Jacobi matrices G_i



A is a sum of Kronecker products (which is never built):

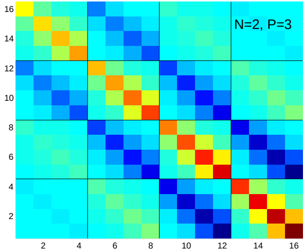
$$A = \sum_{i=0}^{2} G_i \otimes K_i$$

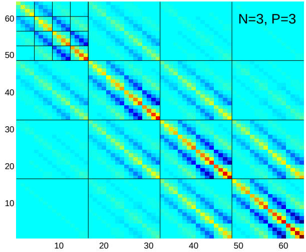
Example of exponential $a(x,\xi)$: A is a sum of Kronecker products:

$$\begin{aligned} A_{ir,js} &= \int_\Gamma \int_D \sum_{k=1}^{m} \Phi_k(\xi)\tilde{a}_k(x)\nabla\psi_s(x)\nabla\psi_r(x)\Phi_i(\xi)\Phi_j(\xi)\rho(\xi)\,\mathrm{d}x\,\mathrm{d}\xi \\ &= \sum_{k=1}^{m} \int_D \tilde{a}_k(x)\nabla\psi_s(x)\nabla\psi_r(x)\,\mathrm{d}x \int_\Gamma \Phi_k(\xi)\Phi_i(\xi)\Phi_j(\xi)\rho(\xi)\,\mathrm{d}\xi \\ &= \sum_{k=1}^{m} (K_k)_{rs}(G_{k,1})_{ij}\cdots(G_{k,N})_{ij} \end{aligned}$$

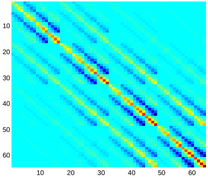
$$(G_{k,t})_{ij} = \int_{\mathscr{R}} \phi_{k,t}(\xi_t)\phi_{i,t}(\xi_t)\phi_{j,t}(\xi_t)\rho(\xi_t)\,\mathrm{d}\xi_t$$

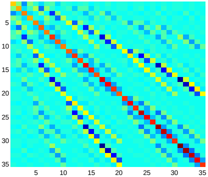
ordering of blocks for TP



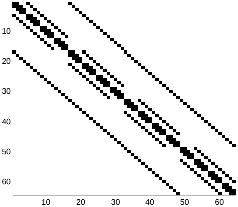


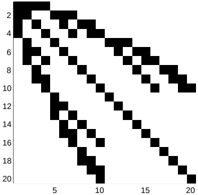
Log-normal $a(x, \xi)$
$\rightarrow$ dense matrix A:





linear $a(x, \xi)$
$\rightarrow$ sparse matrix A:





approximation of $u(x, \xi)$:
tensor product of polynomials in ξ_i (left), complete polynomials in ξ_i (right)

Low rank approximations: vector $u \to$ tensorised matrix U, and

$$v = Au = \left(\sum_{i=0}^{N_\xi} G_i \otimes K_i \right) u \qquad \text{is the same as} \qquad V = \sum_{i=0}^{N_\xi} K_i U G_i^T$$

C. Powell, D. Silvester, V. Simoncini, 2016; H. Matthies et al., 2014

Preconditioning of A:

– with respect to spatial variables:
”multigrid people” - M. Brezina, et al., 2014, H. Elman, D. Furnival, 2007

– with respect to stochastic variables:
only diagonal blocks - C. Powell, H. Elman, 2009
Kronecker product preconditioner - E. Ullmann, 2010
a posteriori error estimates - A. Bespalov, C. Powell, D. Silvester, 2014
two-by-two blocks and Schur complement - B. Sousedik, R. Ghanem, E. Phips, 2013, (numerical tests)

– now: improved theoretical bounds for the condition number of preconditioned A

Patterns of preconditioning matrices M, examples:

$$M^{\mathrm{TP}} = \left(\begin{array}{ccc|ccc|ccc} X & & & & & & & & \\ & X & & & & & & & \\ & & X & & & & & & \\ \hline & & & X & & & & & \\ & & & & X & & & & \\ & & & & & X & & & \\ \hline & & & & & & X & & \\ & & & & & & & X & \\ & & & & & & & & X \end{array}\right), \quad M^{\mathrm{C}} = \left(\begin{array}{c|cc|ccc} X & & & & & \\ \hline & X & & & & \\ & & X & & & \\ \hline & & & X & & \\ & & & & X & \\ & & & & & X \end{array}\right),$$

$$M^{\mathrm{TP}} = \left(\begin{array}{ccc|ccc|ccc} X & X & & & & & & & \\ X & X & X & & & & & & \\ & X & X & & & & & & \\ \hline & & & X & X & & & & \\ & & & X & X & X & & & \\ & & & & X & X & & & \\ \hline & & & & & & X & X & \\ & & & & & & X & X & X \\ & & & & & & & X & X \end{array}\right).$$

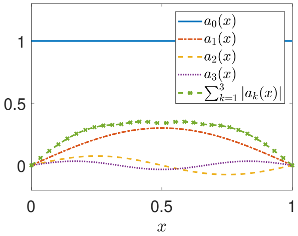
$$M^{\mathrm{TP}} = \left(\begin{array}{ccc|ccc|ccc} X & X & & X & & & & & \\ X & X & X & & X & & & & \\ & X & X & & & X & & & \\ \hline X & & & X & X & & & & \\ & X & & X & X & X & & & \\ & & X & & X & X & & & \\ \hline & & & & & & X & X & \\ & & & & & & X & X & X \\ & & & & & & & X & X \end{array}\right), \quad M^{\mathrm{C}} = \left(\begin{array}{c|cc|ccc} X & X & X & & & \\ \hline X & X & & & & \\ X & & X & & & \\ \hline & & & X & & \\ & & & & X & \\ & & & & & X \end{array}\right).$$

Example. $D = (0,1)$, $N_\xi = 3$, $a(x,\xi) = a_0(x) + \xi_1 a_1(x) + \xi_2 a_2(x) + \xi_3 a_3(x)$,

$$-(a(x,\xi)u(x,\xi)')' = f(x),$$

$\xi_i \in [-1,1]$, $i = 1,2,3$.

Problem 1:



Problem 2:

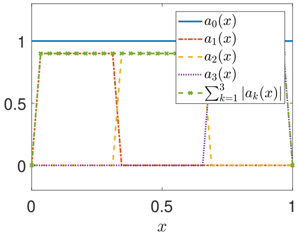


Table: Block-diagonal preconditioning of Problem 1 and Problem 2. New and classical bounds.

$P-1$	$\kappa(A)$	$\underline{c}_{\text{class}}$	$\underline{c}$	$\lambda_{\min}(M^{-1}A)$	$\lambda_{\max}(M^{-1}A)$	$\overline{c}$	$\overline{c}_{\text{class}}$	$\overline{c}/\underline{c}$
Problem 1								
1	458.42	0.76	0.80	0.83	1.17	1.20	1.24	1.51
2	498.47	0.68	0.73	0.76	1.24	1.27	1.32	1.75
...								
6	546.55	0.61	0.67	0.69	1.31	1.33	1.39	2.26
7	550.80	0.61	0.66	0.68	1.32	1.34	1.39	2.29
Problem 2								
1	947.79	-0.65	0.45	0.45	1.56	1.56	2.65	3.43
2	1596.34	-1.21	0.26	0.26	1.74	1.74	3.21	6.57
...								
6	4576.93	-1.71	0.10	0.10	1.90	1.90	3.71	19.34
7	5294.63	-1.74	0.09	0.09	1.91	1.91	3.74	21.80

I.P. AM 2015; I.P. CAMWA 2016; I.P. IJUQ 2017; M. Kubínová, I.P. SIAM/ASA JUQ 2020

Preconditioning of A - estimates of the condition number of $M^{-1}A$ - the main trick:

If

$$c_1 v^T M v \;\leq\; v^T A v \;\leq\; c_2 v^T M v$$

then

$$\kappa(M^{-1}A) \leq \frac{c_2}{c_1}.$$

At the same time,

$$v^T A v = \sum_{j=1}^{n} \int_{E_j} a(x,\xi) \nabla v(x) \cdot \nabla v(x)\, \mathrm{d}x, \quad v^T M v = \sum_{j=1}^{n} \int_{E_j} \widetilde{a}(x,\xi) \nabla v(x) \cdot \nabla v(x)\, \mathrm{d}x,$$

thus c_1 and c_2 can be estimated locally !

$$c_1 \widetilde{a}(x,\xi) \leq a(x,\xi) \leq c_2 \widetilde{a}(x,\xi), \qquad x \in E_j, \quad j = 1, \ldots, n_E.$$

This approach can be used for FEM, stochastic Galerkin, hierarchical bases (AML), ..., for any type of symmetric matrices that are composed "element-wise".

Conclusion

- PDE with (stochastic) parameters - many methods, but all demanding
- Variational approach - stochastic Galerkin method
- Preconditioning - spectral estimates based on local properties of the original and preconditioning operators - joint approach for :

  - condition number estimates for AML preconditioning, V. Eijkhout, P. Vassilevski, O. Axelsson
  - condition number estimates for preconditioned stochastic Galerkin matrices
  - estimates of all eigenvalues of preconditioned matrices:
  T. Gergelits, K.-A. Mardal, B. Nielsen and Z. Strakoš, SINUM 2019,
  T. Gergelits, B. Nielsen, Z. Strakoš, SINUM 2020,
  M. Ladecký, I. P., J. Zeman, AM 2020