Ingolf Busch and Oliver Ernst
Institut für Numerische Mathematik und Optimierung

Efficient Solution of Large-Scale Covariance Eigenproblems

ICMS Workshop on Uncertainty Quantification
Edinburgh, May 24–28, 2010

Outline

The Covariance Eigenproblem

- KL Expansion
- Asymptotics
- Conditioning

Numerical Solution

- Our Covariance Eigenvalue Code
- Hierarchical Matrix Techniques
- Adapted Quadrature Techniques
- Example

Stochastic UQ relies heavily on computations with random fields

$$a(\boldsymbol{x},\omega), \qquad \boldsymbol{x} \in D \subset \mathbb{R}^d, \quad \omega \in \Omega,$$

with $(\Omega, \mathfrak{A}, P)$ a probability space.

Essential: ability to separate stochastic and deterministic dependence in the form

$$a(\boldsymbol{x},\omega) \approx \sum_j \phi_j(\boldsymbol{x})\, \xi_j(\omega)$$

with deterministic functions ϕ_j and random variables ξ_j.

For random fields with finite variance, the most popular option is Karhunen-Loève (KL) expansion.

Mean : $$\overline{a}(\boldsymbol{x}) := \langle a(\boldsymbol{x},\cdot)\rangle := \int_\Omega a(\boldsymbol{x},\omega)\,dP(\omega), \qquad \boldsymbol{x}\in D,$$

Covariance : $$c(\boldsymbol{x},\boldsymbol{y}) := \big\langle (a(\boldsymbol{x},\cdot)-\overline{a}(\boldsymbol{x}))\,(a(\boldsymbol{y},\cdot)-\overline{a}(\boldsymbol{y}))\big\rangle, \quad \boldsymbol{x},\boldsymbol{y}\in D.$$

Hence

$$a(\boldsymbol{x},\omega) = \overline{a}(\boldsymbol{x}) + \tilde{a}(\boldsymbol{x},\omega) \quad \text{with} \quad \langle\tilde{a}\rangle \equiv 0.$$

KL expansion:

$$a(\boldsymbol{x},\omega) = \overline{a}(\boldsymbol{x}) + \sum_{m=1}^\infty \sqrt{\lambda_m}\,a_m(\boldsymbol{x})\,\xi_m(\omega),$$

where $\{\xi_m\}_{m\in\mathbb{N}}$ are centered, uncorrelated RV and $\{\lambda_m, a_m\}$ the eigenpairs of the covariance operator

$$C : L^2(D) \to L^2(D), \qquad (Cu)(\boldsymbol{x}) = \int_D u(\boldsymbol{y})c(\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y},$$

converges in mean-square and uniformly on D.

For normalized eigenfunctions $a_m(\boldsymbol{x})$,

$$\mathrm{Var}_a(\boldsymbol{x}) = c(\boldsymbol{x}, \boldsymbol{x}) = \sum_{m=1}^{\infty} \lambda_m a_m(\boldsymbol{x})^2,$$

$$\int_D \mathrm{Var}_a(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{m=1}^{\infty} \lambda_m \underbrace{(a_m, a_m)_D}_{=1} = \mathrm{trace}\, C.$$

For constant variance (e.g., stationary RF),

$$\mathrm{Var}_a \equiv \sigma^2 > 0, \qquad \sum_m \lambda_m = |D|\, \sigma^2.$$

For computational purposes, KL expansion truncated after M terms:

$$a^{(M)}(\boldsymbol{x},\omega) = \overline{a}(\boldsymbol{x}) + \sum_{m=1}^{M} \sqrt{\lambda_m}\, a_m(\boldsymbol{x})\, \xi_m(\omega).$$

Truncation error

$$\left\langle \|a - a^{(M)}\|_{L^2(D)}^2 \right\rangle = \sum_{m=M+1}^{\infty} \lambda_m.$$

Choose M to retain sufficient fraction $\delta \in (0,1)$ of total variance, i.e.,

$$\frac{\left\langle \|a - a^{(M)}\|_{L^2(D)}^2 \right\rangle}{\left\langle \|a\|_{L^2(D)}^2 \right\rangle} = \frac{\sum_{m=M+1}^{\infty} \lambda_m}{\sum_{m=1}^{\infty} \lambda_m} = 1 - \frac{\sum_{m=1}^{M} \lambda_m}{|D|\sigma^2} < \delta.$$

$$c(\boldsymbol{x}, \boldsymbol{y}) = c(r), \qquad r = \|\boldsymbol{x} - \boldsymbol{y}\|_2$$

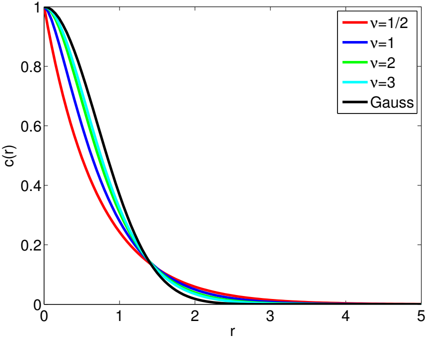
Convenient parametrization: **Matérn class** of covariance kernels:

$$c(r) = \frac{\sigma^2}{2^{\nu-1}\,\Gamma(\nu)} \left(\frac{2\sqrt{\nu}\,r}{\rho}\right)^{\nu} K_\nu \left(\frac{2\sqrt{\nu}\,r}{\rho}\right)$$

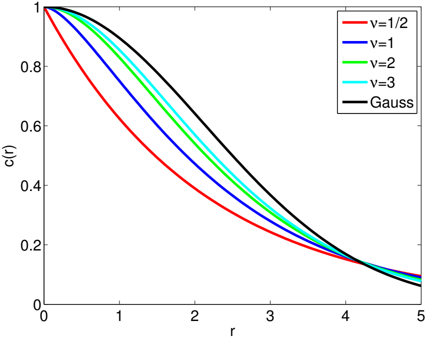
K_ν : modified Bessel function of order ν
ν : smoothness parameter
ρ : correlation length parameter

Special cases:

$\nu = \frac{1}{2}$:	$c(r) = \sigma^2 \exp(-\sqrt{2}r/\rho)$	exponential covariance
$\nu = 1$:	$c(r) = \sigma^2 \left(\frac{2r}{\rho}\right) K_1 \left(\frac{2r}{\rho}\right)$	Bessel covariance
$\nu \to \infty$:	$c(r) = \sigma^2 \exp(-r^2/\rho^2)$	Gaussian covariance



$\rho = 1$



$\rho = 3$

Smoothness of realizations: RF a is s times mean-square differentiable if and only if $\nu > s$.

Eigenvalue Decay: the smoother the kernel, the faster $\{\lambda_m\}_{m\in\mathbb{N}} \to 0$.
More precisely: if $D \subset \mathbb{R}^d$, then if the kernel function c is

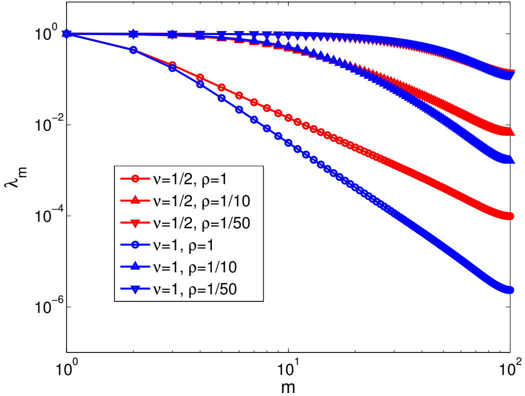
piecewise H^s :	$\lambda_m \le c_1 m^{-s/d}$
piecewise smooth :	$\lambda_m \le c_2 m^{-s}$ for any $s > 0$
piecewise analytic :	$\lambda_m \le c_3 \exp(-c_4 m^{1/d})$

for suitable constants c_1, c_2, c_3, c_4.

Note: piecewise smoothness of kernel also leads to bounds on derivatives of eigenfunctions a_m in $L^\infty(D)$.

Proven e.g. in [Schwab & Todor (2006)], [Todor (2006)]

Before asymptotic decay (determined by the smoothness of the kernel) sets in, there is a preasymptotic **plateau** determined by the correlation length parameter.



Eigenvalue decay, Matérn covariance kernel, $D = [-1, 1]$.

A common variance-reduction technique is to incorporate available measurement data for a given RF, e.g.,

$$a(\boldsymbol{x}_i, \omega) = \eta_i, \qquad i = 1, \dots, k, \quad \forall \omega \in \Omega, \tag{1}$$

at locations $\boldsymbol{x}_i \in D$, i.e., to consider the **RF a conditioned on (1)**.
(Other functionals than point evaluation are also possible.)
This is given by

$$a_k(\boldsymbol{x}, \omega) = a(\boldsymbol{x}, \omega) + \sum_{i,j=1}^{k} c(\boldsymbol{x}, \boldsymbol{x}_i)[\boldsymbol{C}_k^{-1}]_{i,j}(\eta_j - a(\boldsymbol{x}_j, \omega)),$$

where $[\boldsymbol{C}_k]_{i,j} = c(\boldsymbol{x}_i, \boldsymbol{x}_j)$, with covariance function

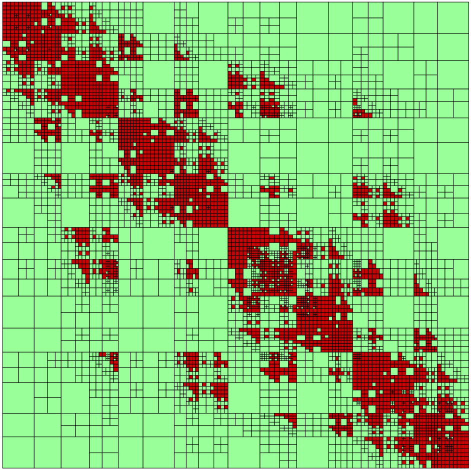
$$c_k(\boldsymbol{x}, \boldsymbol{y}) = c(\boldsymbol{x}, \boldsymbol{y}) - \sum_{i,j=1}^{k} c(\boldsymbol{x}, \boldsymbol{x}_i)[\boldsymbol{C}_k^{-1}]_{i,j} c(\boldsymbol{x}_j, \boldsymbol{y}).$$

- Piecewise constant Galerkin discretization on triangles/tetrahedra.
  Results in generalized matrix eigenvalue problem

$$\boldsymbol{C}\boldsymbol{x} = \lambda \boldsymbol{M}\boldsymbol{x},$$

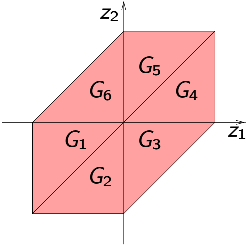
  where M can be chosen diagonal but C is generally dense.
- Construction/storage of and multiplication with Galerkin matrix via hierarchical matrix technique.
- Near-field Galerkin blocks using adapted quadrature.
- Thick-restart Lanczos method for dominant eigenpairs of given order, extended to block variant for multiple eigenpairs.
- Conditioning of covariance on measured data via low-rank correction.

- Algebraic variant of fast multipole method, [Hackbusch et al. (2000)]
- Partition dense matrix into square blocks of 2 types
  - full near field blocks,
  - low-rank far field blocks
- blocks correspond to clusters of degrees of freedom, i.e., clusters of supports of Galerkin basis functions
- yields data-sparse representation of matrix, construction $O(N \log N)$, matrix-vector product in $O(N)$.



Numerical Solution
Adapted Quadrature Techniques

- Slow convergence of quadrature error due to covariance kernels' lack of smoothness near diagonal $\{x = y\}$ of $D \times D$.
- For triangular elements distinguish cases of identical element, common edge, common point and regular case.
- E.g. for identical triangles, difference coordinate $z = x - y$ moves singularity to $z = 0$; integrating over (x, z)-space leaves six triangles in z-plane, which can in turn be each mapped to a reference triangle.



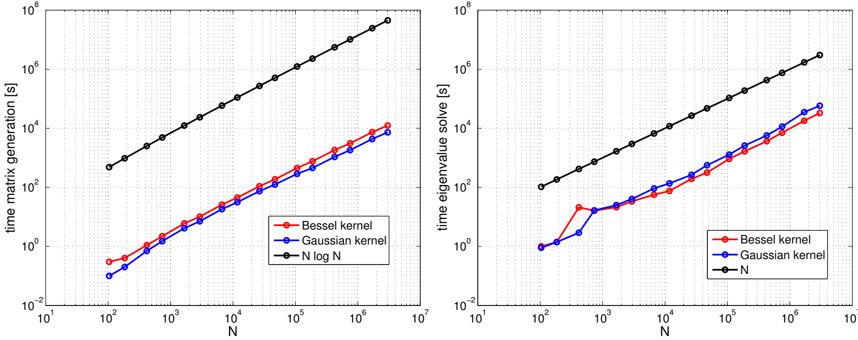
- Change of variables concentrates and weakens singularity.
- Similar techniques often used in boundary element methods, cf. [Sauter & Schwab (2010)].

Example
Performance

Bessel covariance ($\nu = 1$) on $D = [-1,1]^2$, $\rho = 1/4$.

Compute leading 300 eigenpairs on sequence of uniformly refined triangular grids, resulting in $N = 104$ to $N = 3\,014\,656$ DOF.

Used Krylov space of dimension 500.



hierarchical matrix generation | eigenvalue calculation

Concluding Remarks

Summary

- Reliable tool for large-scale covariance eigenvalue problems
- Memory and time requirements scale as $O(N \log N)$
- Mixture of Matlab and C code, uses HLIB 1.4 package from MPI Leipzig

In progress:

- In progress: adapted quadrature in 3D
- In progress: higher order L^2-conforming discretization

References

[1] M. Eiermann, O. G. Ernst, and E. Ullmann. Computational aspects of the stochastic finite element method. *Comput. Visual. Sci.*, 10(1):3–15, 2007.

[2] W. Hackbusch and S. Börm. H^2-matrix approximation of integral operators by interpolation. *Applied Numerical Mathematics*, 43(1–2):129–143, 2002.

[3] C. Schwab and S. Sauter. *Boundary Element Methods*. Springer, 2010.

[4] M. L. Stein. *Interpolation of Spatial Data*. Springer, 1999.

[5] R. A. Todor. Robust eigenvalue computations for smoothing operators. *SIAM J. Numer. Anal.*, 44(2):865–878, 2006.

[6] K. Wu and H. Simon. Thick-restart Lanczos method for large symmetric eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications*, 22(2):602–616, 2000.