Die Ressourcenuniversität. Seit 1765.

Institut für Numerische Mathematik und Optimierung

On the Convergence of Generalized Polynomial Chaos Expansions

Oliver Ernst

with E. Ullmann (TU Freiberg),
H.-J. Starkloff, A. Mugler (U Zwickau)

DFG SPP 1324

LMS Durham Symposium: Numerical Analysis of Multiscale Problems
July 5–15, 2010

Outline

Outline

- [Wiener, 1920–1923] Mathematical analysis of Brownian motion. Wiener process, Wiener measure, Wiener integral.
- [Kolmogorov, 1933] Formalization of probability as measure theory
- [Wiener, 1938] *"The Homogeneous Chaos"*
  Attempt at modeling nonlinear phenomena in statistical mechanics, turbulence.
  - "homogeneous chaos": stationary random measure
  - polynomial chaoses through repeated (Wiener) integration
  - general stochastic processes approximated by (nonlinear) functionals of multidimensional Wiener process.
- [Cameron & Martin, 1947] Wiener-Hermite orthogonal expansion of 2nd order random processes
- [Itô, 1953] Connection between Itô Integral, polynomial chaos expansion and expansions with multiple Wiener integrals. See also [Kallianpur, 1980].

- 1980s: Uncertainty Quantification via *Stochastic Finite Element Methods*
  PDEs with random data, spatial part discretized via FE, randomness treated by Monte Carlo method, perturbation expansions, response surface methods
- [Ghanem & Spanos, 1991] *Spectral Stochastic Finite Element Method*
  Seek random field solution to PDE with random input in tensor product space $X \otimes \mathbf{\Xi}$

  X : function space appropriate for deterministic version of PDE

  $\mathbf{\Xi} := L^2(\Omega, \mathfrak{A}, P)$, for probability space $(\Omega, \mathfrak{A}, P)$

  Discretization
  - finite dimensional noise assumption
  - L^2-RV approximated by multivariate Hermite polynomials in finite number of Gaussian RVs, inspired by PC expansions.

Some History
. . . and finally

[Xiu & Karniadakis, 2002-03] *Generalized Polynomial Chaos (GPC)*

Observation: Multivariate polynomials in non-Gaussian basic random variables sometimes have better approximation properties than PC expansions.

Question: When can we expect GPC expansions to converge?

Setting and Notation

$$
\begin{aligned}
(\Omega, \mathfrak{A}, P) &: \text{ probability space} \\
\xi : \Omega \to \mathbb{R} &: \text{ random variable} \\
\langle \xi \rangle &: \text{ expectation} \\
\sigma(S) &: \ \sigma\text{-algebra generated by set of RV } S \\
L^2(\Omega, \mathfrak{A}, P) &: \text{ Hilbert space of real-valued RV w. finite second moments} \\
\|\xi\|_{L^2}^2 = \left\langle \xi^2 \right\rangle &: \text{ associated norm (mean-square convergence)} \\
\mathscr{H} &: \text{ Gaussian linear (Hilbert) space: (complete) subspace of} \\
&\quad\ \text{consisting of centered Gaussian RV } L^2(\Omega, \mathfrak{A}, P)
\end{aligned}
$$

Note: $\mathscr{H}$ cannot contain *all* Gaussian RV in underlying space.

Outline

Background

The Cameron-Martin Theorem

Polynomial Chaos Expansions

Generalized Polynomial Chaos Expansions

Examples

For Gaussian linear space $\mathscr{H}$ and $n \in \mathbb{N}_0$, set

$$\mathscr{P}_n(\mathscr{H}) := \Big\{ p(\xi_1, \dots, \xi_M) : p \text{ an } M\text{-variate polynomial of degree} \le n, \\ \xi_j \in \mathscr{H},\, j = 1, \dots, M, M \in \mathbb{N} \Big\}.$$

$$\begin{aligned} \mathscr{P}_n(\mathscr{H}), \overline{\mathscr{P}}_n(\mathscr{H}) &\subset L^2(\Omega, \mathfrak{A}, P), \\ \mathscr{P}_0(\mathscr{H}) = \overline{\mathscr{P}}_0(\mathscr{H}) &\text{ a.s. constant RV}, \\ \mathscr{P}_1(\mathscr{H}), \overline{\mathscr{P}}_1(\mathscr{H}) &\text{ Gaussian RV}, \\ \{\overline{\mathscr{P}}_n(\mathscr{H})\}_{n \in \mathbb{N}_0} &\text{ strictly increasing subspaces of } L^2(\Omega, \mathfrak{A}, P). \end{aligned}$$

Setting

$$\mathscr{H}_0 := \mathscr{P}_0(\mathscr{H}) = \overline{\mathscr{P}}_0(\mathscr{H}), \qquad \mathscr{H}_n := \overline{\mathscr{P}}_n(\mathscr{H}) \cap \mathscr{P}_{n-1}(\mathscr{H})^\perp, \quad n \in \mathbb{N},$$

we have

$$\overline{\mathscr{P}}_n(\mathscr{H}) = \bigoplus_{k=0}^{n} \mathscr{H}_k.$$

We also set

$$\bigoplus_{n=0}^{\infty} \mathscr{H}_n := \overline{\bigcup_{n=0}^{\infty} \mathscr{P}_n(\mathscr{H})}.$$

Theorem (Cameron & Martin, 1947)

$$\bigoplus_{n=0}^{\infty} \mathscr{H}_n = L^2(\Omega, \sigma(\mathscr{H}), P).$$

In particular, if $\sigma(\mathscr{H}) = \mathfrak{A}$, then

$$L^2(\Omega, \mathfrak{A}, P) = \bigoplus_{n=0}^{\infty} \mathscr{H}_n \ .$$

Note: Condition $\sigma(\mathscr{H}) = \mathfrak{A}$ crucial.
Consider $\xi \sim N(0,1)$, $\mathscr{H} = \text{span}\{\xi\}$, and $\eta \in L^2(\Omega, \mathfrak{A}, P)$, $\langle \eta \rangle = 0$, ξ, η independent. Then all orthogonal projections of η on $\mathscr{H}_n$ vanish a.s., with approximation error $\langle \eta^2 \rangle$.

Outline

$$\mathscr{H}: \text{ Gaussian linear space,}$$
$$P_k : L^2(\Omega, \mathfrak{A}, P) \to \mathscr{H}_k: \text{ orthogonal projection onto } \mathscr{H}_k$$

Polynomial chaos expansion of $\eta \in L^2(\Omega, \sigma(\mathscr{H}), P)$ given by

$$\eta = \sum_{k=0}^{\infty} P_k \eta.$$

Expansion also (mean-square) convergent when $\mathfrak{A} \supsetneq \sigma(\mathscr{H})$, but to orthogonal projection of η onto $L^2(\Omega, \sigma(\mathscr{H}), P)$.

In applications typically have

$$\mathscr{H} = \text{span}\{\xi_j : j \in \mathbb{N}\}, \quad \xi_j \sim N(0,1) \text{ independent basic RV.}$$

Orthonormal basis of $\mathscr{H}$ given by $\{\psi_{\boldsymbol{\alpha}} : |\boldsymbol{\alpha}|_0 < \infty\}$, where

$$\boldsymbol{\alpha} \in \{(\alpha_1, \alpha_2, \dots) : \alpha_j \in \mathbb{N}_0\}, \quad |\boldsymbol{\alpha}|_0 := |\{j : \alpha_j > 0\}|,$$

$$\psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{\alpha_j \neq 0} \psi_{\alpha_j}(\xi_j)$$

where $\{\psi_k\}$ denotes the sequence of normalized Hermite polynomials.

For finitely many basic RV $\xi_1, \dots, \xi_M$ and $\mathscr{P}_n^M(\xi_1, \dots, \xi_M)$ the M-variate polynomials in $\{\xi_j\}_{j=1}^M$ of degree at most n, there holds

$$\eta_n^M := P_n^M \eta \xrightarrow{n,M\to\infty} \eta \qquad \forall \eta \in L^2(\Omega, \sigma(\{\xi_j\}_{j\in\mathbb{N}}), P).$$

Consider a smooth transformation

$$K = K(\boldsymbol{x}, \omega) = f(G(\boldsymbol{x}, \omega)), \qquad \boldsymbol{x} \in D \subset \mathbb{R}^d,$$

of a Gaussian random field $G = G(\boldsymbol{x}, \omega)$ given by its Karhunen-Loève expansion

$$G(\boldsymbol{x}, \omega) = \langle G(\boldsymbol{x}) \rangle + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\, g_m(\boldsymbol{x})\, \xi_m(\omega), \qquad \xi_m \sim N(0,1) \text{ i.i.d.}$$

The coefficients $K_{\boldsymbol{\alpha}}(\boldsymbol{x})$ of the polynomial chaos expansion

$$K(\boldsymbol{x}, \omega) = \sum_{\boldsymbol{\alpha}} K_{\boldsymbol{\alpha}}(\boldsymbol{x})\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}(\omega))$$

satisfy (cf. [Malliavin, 1997])

$$K_{\alpha}(\boldsymbol{x}) = \langle K(\boldsymbol{x}, \omega) \psi_{\alpha}(\boldsymbol{\xi}(\omega)) \rangle = \frac{1}{\sqrt{\boldsymbol{\alpha}!}} \left\langle D^{\boldsymbol{\alpha}} f(G(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \right\rangle.$$

Special case: lognormal random field $K(\boldsymbol{x},\omega) = e^{G(\boldsymbol{x},\omega)}$.

Here we obtain

$$K_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \frac{\langle K(\boldsymbol{x})\rangle}{\sqrt{\boldsymbol{\alpha}!}} \prod_{m=1}^{\infty} \left(\sqrt{\lambda_m}\, g_m(\boldsymbol{x})\right)^{\alpha_m}.$$

Outline

If RV η "far from Gaussian", expand it in polynomials of RV with non-Gaussian distributions.

Many common probability distributions correspond to classical real orthogonal polynomials, e.g.,

Distribution	polynomials	density
Gaussian	Hermite	$\rho(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}$
Gamma(α, λ)	Laguerre	$\rho(\xi) = \frac{\lambda}{\xi(\alpha)} (\lambda\xi)^{\alpha-1} e^{-\lambda\xi}$
Beta(α, β)	Jacobi	$\rho(\xi) = \frac{(1-\xi)^\alpha (1+\xi)^\beta}{2^{\alpha+\beta+1} B(\alpha+1,\beta+1)}$
Uniform(α, β)	Legendre	$\rho(\xi) = \frac{1}{\beta-\alpha}$
Arcsin	Chebyshev	$\rho(\xi) = \frac{1}{\sqrt{1-\xi^2}}$

[Xiu & Karniadakis, 2002–03] Askey family
[Ogura, 1972] Poisson chaos (Charlier polynomials)

Assumption: Basic RV ξ with finite moments $\left\langle |\xi|^k \right\rangle$ of all orders and continuous distribution function F_ξ.

- Then there exists sequence $\{\psi_k\}_{k\in\mathbb{N}_0}$ of polynomials $(\deg\psi_k = k)$ orthonormal with respect to the distribution of ξ, i.e., in $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), F_\xi(dx))$.
- For any $\eta \in L^2(\Omega, \mathfrak{A}, P)$ the coefficients a_k of the expansion

$$\eta \sim \sum_{k=0}^{\infty} a_k \psi_k(\xi), \qquad a_k = \langle \eta \psi_k(\xi) \rangle$$

  are defined.
- **Question:** does the expansion converge to η in quadratic mean?

Equality

$$\eta = \sum_{k=0}^{\infty} a_k \psi_k(\xi) \quad \text{for all} \quad \eta \in L^2(\Omega, \sigma(\xi), P)$$

equivalent with density of polynomials

$$p(\xi) \text{ in } L^2(\Omega, \sigma(\xi), P) \text{ or}$$
$$p(x) \text{ in } L^2(\mathbb{R}, \mathfrak{B}, F_\xi(dx)), \text{ respectively.}$$

Theorem (M. Riesz, 1923)

The polynomials span$\{\xi^k\}_{k \in \mathbb{N}_0}$ are dense in $L^2(\Omega, \sigma(\xi), P)$ if and only if the Hamburger moment problem is uniquely solvable for the distribution of ξ.

Definition

The moment problem is uniquely solvable for a probability distribution on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ or the distribution is *determinate in the Hamburger sense*, if the distribution function is uniquely defined by the sequence of its moments

$$\mu_k := \left\langle \xi^k \right\rangle = \int_{\mathbb{R}} x^k F_\xi(dx), \qquad k \in \mathbb{N}_0.$$

Thus: generalized polynomial chaos expansions in one basic RV ξ converge if and only if the distribution of ξ is determinate.

Some Determinate/Indeterminate Distributions

- determinate distributions:
  - normal
  - uniform
  - beta
  - gamma
  - ...
- indeterminate distributions:
  - lognormal
  - certain powers of Gaussian RV, e.g.

    ξ^{2k+1} for any $k = 1, 2, \ldots$ or

    ξ^{2k} for any $k = 3, 4, \ldots \quad (\xi \sim N(0,1))$
  - certain powers of exponentially distributed RV
  - ...

Generalized Cameron-Martin Theorem

Let $\{\xi_k\}_{k\in\mathbb{N}}$ be independent RV with continuous distributions and possessing moments all orders.
Furthermore let $\{\mathscr{H}_n\}_{n\geq 0}$ be the polynomial subspaces as in the Cameron-Martin theorem.
Then the spaces $\{\mathscr{H}_n\}_{n\geq 0}$ are mutually orthogonal closed subspaces of $L^2(\Omega, \mathfrak{A}, P)$ and there holds

$$\bigoplus_{n=0}^{\infty} \mathscr{H}_n = L^2(\Omega, \sigma(\{\xi_k\}_{k\in\mathbb{N}}), P)$$

if and only if for each basic random variable $\xi_k, k \in \mathbb{N}$, the moment problem for its distribution is uniquely solvable.

Idea of proof:

- For one basic random variable ξ_k orthonormal polynomials yield an orthonormal basis in $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), F_{\xi_k}(dx_k))$.
- For finitely many independent basic random variables tensor products of univariate orthonormal polynomials yield an orthonormal basis in $L^2(\mathbb{R}^M, \mathfrak{B}(\mathbb{R}^M), F_{\xi_1}(dx_1) \times \ldots \times F_{\xi_M}(dx_M))$
- General case: approximation of random variables depending on $(\xi_1, \xi_2, \ldots)$ by random variables depending on a finite number of basic random variables

Note: For nonindependent basic RV the condition is sufficient.

Outline

Consider standard lognormal RV $\xi = \exp(\gamma)$, $\gamma \sim N(0,1)$.

Probability density function given by

$$f_\xi(x) = \begin{cases} \frac{1}{x\sqrt{2\pi}} e^{-\frac{\log^2 x}{2}}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

This distribution is indeterminate in the Hamburger sense.

Associated orthonormal polynomials

$$\psi_0(x) \equiv 1,$$
$$\psi_k(x) = \frac{(-1)^k e^{k(k-1)/4}}{\sqrt{\prod_{i=1}^k (e^i - 1)}} \sum_{j=0}^{k} (-1)^j \begin{bmatrix} k \\ j \end{bmatrix} e^{-j^2+j/2} x^j, \quad k \geq 1,$$

with

$$\begin{bmatrix} k \\ j \end{bmatrix} = \frac{(1-e^{-k})(1-e^{-(k-1)}) \cdot \ldots \cdot (1-e^{-(k-j+1)})}{(1-e^{-j})(1-e^{-(j-1)}) \cdot \ldots \cdot (1-e^{-1})}.$$

Can be derived from Stieltjes-Wigert polynomials (cf. [Szegö, 1939])

Consider lognormal RV ξ and

$g : \mathbb{R} \to \mathbb{R}$ measurable, odd and 1-periodic for which

$$\left\langle (g(\log \xi))^2 \right\rangle < \infty, \qquad \text{e.g. } g(x) = \sin(2\pi x).$$

Then for all $k \in \mathbb{N}$ we have

$$a_k = \langle \psi_k(\xi) g(\log(\xi)) \rangle = \int_0^\infty \psi_k(\xi)\, g(\log(\xi))\, f_\xi(x)\, dx = 0,$$

and therefore, for $\eta = g(\log(\xi)) \in L^2(\Omega, \sigma(\xi), P)$,

$$\eta \neq \sum_{k=0}^\infty a_k \psi_k(\xi).$$

Consider RV $\eta = \frac{1}{\xi}$, ξ lognormal.

Lognormal chaos coefficients of η given by

$$a_0 = \sqrt{e}, \qquad a_k = (-1)^k e^{-(k^2+3k-2)/4} \left(\prod_{j=1}^{k} (e^j - 1) \right)^{1/2}, \quad k \geq 1.$$

Partial sums of chaos expansion $\eta_n := \sum_{k=0}^{n} a_k \psi_k(\xi)$ can be bounded by

$$\|\eta_n\|_{L^2}^2 \leq \frac{e^2}{e-1}.$$

Since $\|\eta\|_{L^2}^2 = e^2$ remainders of partial sums satisfy

$$\|\eta - \eta_n\|_{L^2}^2 = \|\eta\|_{L^2}^2 - \|\eta_n\|_{L^2}^2 \geq e^2 - \frac{e^2}{e-1} > 0.$$

Therefore, for $\eta = \frac{1}{\xi} \in L^2(\Omega, \sigma(\xi), P)$, ξ lognormal, we again have

$$\eta \neq \sum_{k=0}^{\infty} a_k \psi_k(\xi).$$

Consider 1D diffusion problem on $D = (0,1)$

$$-(Ku')' = f, \qquad u(0) = 0, \qquad (Ku')(1) = F,$$

with f, F deterministic and $K = K(x,\omega)$ a random field, with solution

$$u(x,\omega) = \int_0^x \frac{1}{K(y,\omega)} \left(F + \int_y^1 f(z)\, dz \right) dy.$$

For $K(x,\omega) = \xi(\omega)$, ξ lognormal, this becomes

$$u(x,\omega) = \frac{1}{\xi(\omega)} \underbrace{\int_0^x \left(F + \int_y^1 f(z)\, dz \right) dy}_{\text{deterministic}},$$

a random field which cannot be expanded in lognormal chaos.

Same diffusion problem, now

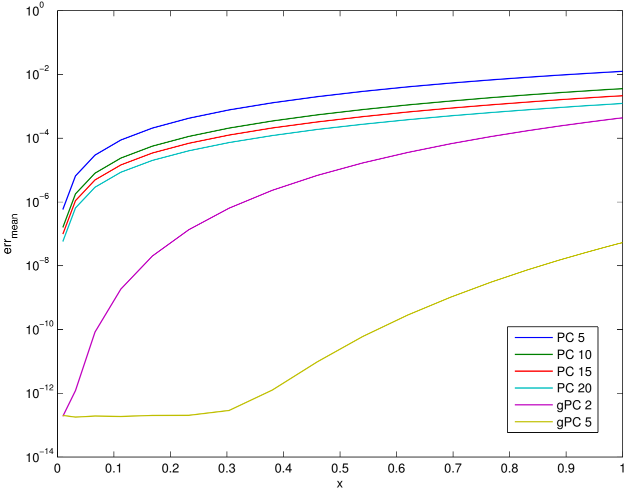
$$K(x,\omega) = \exp(|\xi(\omega)|x), \qquad \xi \sim N(0,1), \quad x \in (0,1),$$

with $f \equiv 1$, $K = 1$.

Distribution of $|\xi|$ (“reflected Gaussian”) is determinate, therefore associated chaos expansion converges to u in mean square.

We thus compare a stochastic Galerkin approximation to the solution of the diffusion problem with spectral element discretization in space combined with

- standard (Hermite) PC approximation vs.
- generalized PC approximation using expansion in orthogonal polynomials associated with distribution of $|\xi|$ (reflected Gaussian).



relative error in mean $\langle u(x) \rangle$

Conclusion

Summary

- Polynomial Chaos expansions are one of many representations for random variables (and random fields) with finite variance.
- PC central to stochastic Galerkin/stochastic collocation methods for discretizing PDEs with random inputs.
- Generalized PC expansions sometimes more efficient, but underlying distribution must be determinate for polynomials to be dense.
- Approximability vs. measurability issues.

Ongoing work:

- Quantify convergence rates of different GPC expansions
- Fewer basic random variables vs. more highly nonlinear transformations.

References

R. H. Cameron and W. T. Martin.
The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals.
Annals of Mathematics, 48(2):385–392, 1947.

O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann.
On the convergence of generalized polynomial chaos expansions.
(submitted), 2010.

Kiyosi Itô.
Multiple Wiener integral.
Journal of the Mathematical Society of Japan, 3(1):157–169, 1951.

Norbert Wiener.
The homogeneous chaos.
American Journal of Mathematics, 60(897–936), 1938.

Dongbin Xiu and George Em Karniadakis.
The Wiener-Askey polynomial chaos for stochastic differential equations.
SIAM J. Sci. Comput., 24(2):619–644, 2002.