Metamodelling based on polynomial chaos expansions

Géraud Blatman

École de Physique des Houches

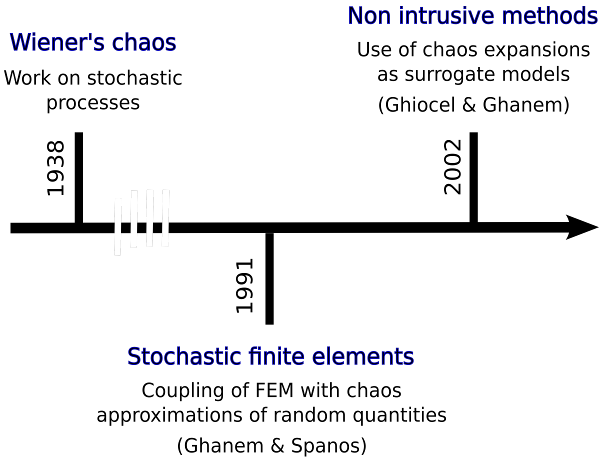
11 Avril 2013

Some popular types of metamodels

- Polynomial response surfaces (PRS)
- Artificial neural networks
- Radial basis functions
- Kriging

A specific PRS method has been widely used in stochastic mechanics since the early 90's : polynomial chaos expansions (PCE).

Some key dates for PCE's



Rationale for using PCE's

- Assumption of a smooth behaviour of the model

  → Good convergence properties

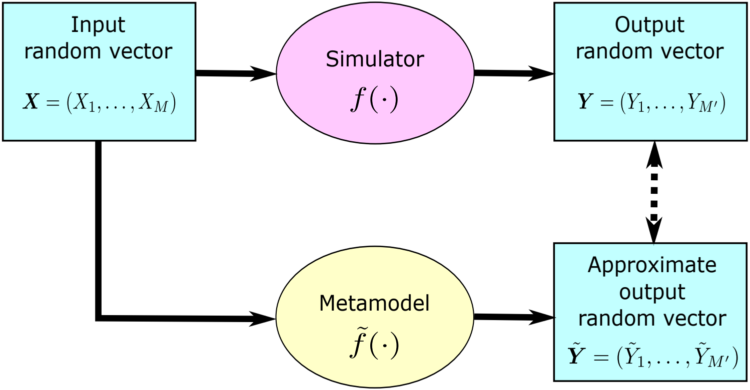
- Metamodel adapted to probabilistic analyses

  → Good approximation and easy derivation of probabilistic quantities (e.g. moments and sensitivity factors)

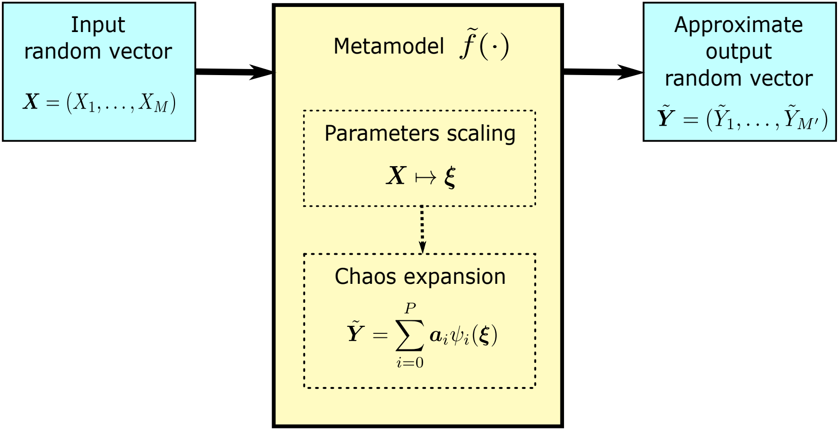
Outline

1. Mathematical formalization
2. Polynomial chaos basis
3. Coefficient estimation
4. Chaos post-processing

Metamodel for uncertainty propagation



Parametric metamodelling based on chaos expansions



PCE for one basic random variable

Simulation model : $Y = h(\xi)$

Assumptions :

- ξ has a continuous PDF
- $\mathbb{E}\left[|\xi^i|\right] < \infty \ , \ \forall\, i \geq 1$
- $\mathbb{V}\text{ar}[Y] < \infty$

There exists a family of orthonormal polynomials $(\pi_i)_{i \in \mathbb{N}}$ such that :

$$Y \ = \ \sum_{i=0}^{\infty} a_i \, \pi_i(\xi) \qquad , \qquad a_i \ = \ \mathbb{E}\left[Y \, \pi_i(\xi)\right]$$

where $\mathbb{E}\left[\pi_i(\xi) \, \pi_j(\xi)\right] \ = \ 1$ if $i = j$ and 0 otherwise

PCE for one basic random variable

Simulation model : $Y = h(\xi)$

Assumptions :

- ξ has a continuous PDF
- $\mathbb{E}\left[|\xi^i|\right] < \infty \ , \ \forall\, i \geq 1$
- $\mathbb{V}\text{ar}[Y] < \infty$

There exists a family of orthonormal polynomials $(\pi_i)_{i\in\mathbb{N}}$ such that :

$$Y \;=\; \sum_{i=0}^{\infty} a_i\, \pi_i(\xi) \qquad , \qquad a_i \;=\; \mathbb{E}\left[Y\, \pi_i(\xi)\right]$$

where $\mathbb{E}\left[\pi_i(\xi)\, \pi_j(\xi)\right] \;=\; 1$ if $i = j$ and 0 otherwise

Correspondance between usual PDFs and polynomials

PDF of ξ	Support	Polynomials
Normal (Gaussian)	$\mathbb{R}$	Hermite
Uniform	$[-1, 1]$	Legendre
Gamma	$(0, \infty)$	Laguerre
Chebyshev	$(-1, 1)$	Chebyshev
Beta	$(-1, 1)$	Jacobi

PCE for several independent basic random variables

Simulation model : $Y = h(\boldsymbol{\xi}) \quad , \quad \boldsymbol{\xi} = (\xi_1, \ldots, \xi_M)$

Assumptions :

- Same as before for each ξ_i
- Independence of the ξ_i's

Upon multiplying the 1D orthonormal polynomials :

$$\psi_{i_1,\ldots,i_M}(\boldsymbol{\xi}) \quad = \quad \pi^{(1)}_{i_1}(\xi_1) \times \cdots \times \pi^{(M)}_{i_M}(\xi_M)$$

one gets :

$$Y \quad = \quad \sum_{i_1=0}^{\infty} \cdots \sum_{i_M=0}^{\infty} a_{i_1,\ldots,i_M} \, \psi_{i_1,\ldots,i_M}(\boldsymbol{\xi})$$

PCE for several independent basic random variables

Simulation model : $Y = h(\boldsymbol{\xi}) \quad , \quad \boldsymbol{\xi} = (\xi_1, \ldots, \xi_M)$

Assumptions :

- Same as before for each ξ_i
- Independence of the ξ_i's

Upon multiplying the 1D orthonormal polynomials :

$$\psi_{i_1,\ldots,i_M}(\boldsymbol{\xi}) \quad = \quad \pi_{i_1}^{(1)}(\xi_1) \times \cdots \times \pi_{i_M}^{(M)}(\xi_M)$$

one gets :

$$Y \quad = \quad \sum_{i_1=0}^{\infty} \cdots \sum_{i_M=0}^{\infty} a_{i_1,\ldots,i_M} \; \psi_{i_1,\ldots,i_M}(\boldsymbol{\xi})$$

Multi-index notation

- Set of tuples of M non negative integers : $\mathbb{N}^M$
- An element of this set : $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M)$

$$Y = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} a_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

where $a_{\boldsymbol{\alpha}} = \mathbb{E}\left[Y \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\right]$, $\mathbb{E}\left[\psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \, \psi_{\boldsymbol{\beta}}(\boldsymbol{\xi})\right] = \delta_{\boldsymbol{\alpha},\boldsymbol{\beta}}$

Example : variables (ξ_1, ξ_2, ξ_3) uniform over [-1,1]

- For each ξ_i, one selects the 1D basis made of normalized Hermite polynomials :

$$\pi_0(X_i) = 1 \quad , \quad \pi_1(\xi_i) = \frac{3\sqrt{2}}{2}\xi_i \quad , \quad \pi_2(\xi_i) = \frac{\sqrt{10}}{4}\left(3\xi_i^2 - 1\right) \quad , \quad \ldots$$

- Examples of chaos multivariate basis polynomials :

$$\begin{array}{rclcc}
\psi_{(0,0,0)}(\boldsymbol{\xi}) & = & \pi_0(\xi_1)\,\pi_0(\xi_2)\,\pi_0(\xi_3) & = & 1 \\
\psi_{(0,2,0)}(\boldsymbol{\xi}) & = & \pi_0(\xi_1)\,\pi_2(\xi_2)\,\pi_0(\xi_3) & = & \frac{\sqrt{10}}{4}\left(3\xi_2^2 - 1\right) \\
\psi_{(1,0,1)}(\boldsymbol{\xi}) & = & \pi_1(\xi_1)\,\pi_0(\xi_2)\,\pi_1(\xi_3) & = & \frac{3}{2}\xi_1\xi_3
\end{array}$$

Truncation of the PCE

- For computational purpose the series has to be truncated.
- Common strategy : retain only the polynomials of total degree p :

$$\tilde{Y} \quad = \quad \sum_{|\boldsymbol{\alpha}| \leq p} a_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \qquad , \qquad |\boldsymbol{\alpha}| = \sum_{i=1}^{M} \alpha_i$$

The number of terms is given by :

$$P \quad = \quad \frac{(p+M)!}{p! \, M!}$$

Scaling the input random variables

- Usually one deals with physical input parameters $\boldsymbol{X}$ expressed in given units.
- These parameters must be transformed into basic, dimensionless variables $\boldsymbol{\xi}$ (cf. table) to construct the PCE.

Denoting by F_{X_i} and the F_{ξ_i} the (known) CDF's of each X_i and ξ_i, an isoprobabilistic transformation is performed :

$$\xi_i \quad = \quad \mathcal{T}_i(X_i) \quad \equiv \quad F_{\xi_i}^{-1}\left[F_{X_i}(X_i)\right]$$

Remark on the convergence rate of PCE's

Surrogate modelling using a truncated PCE :

$$Y = f(\boldsymbol{X}) = f \circ \mathcal{T}^{-1}(\boldsymbol{\xi}) \approx \sum_{|\boldsymbol{\alpha}| \leq p} a_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

The convergence rate depends on the smoothness of $f \circ \mathcal{T}^{-1}$.

→ Choose a scaled random vector $\boldsymbol{\xi}$ ensuring this smoothness.

Guidelines for the choice of the ξ_i's

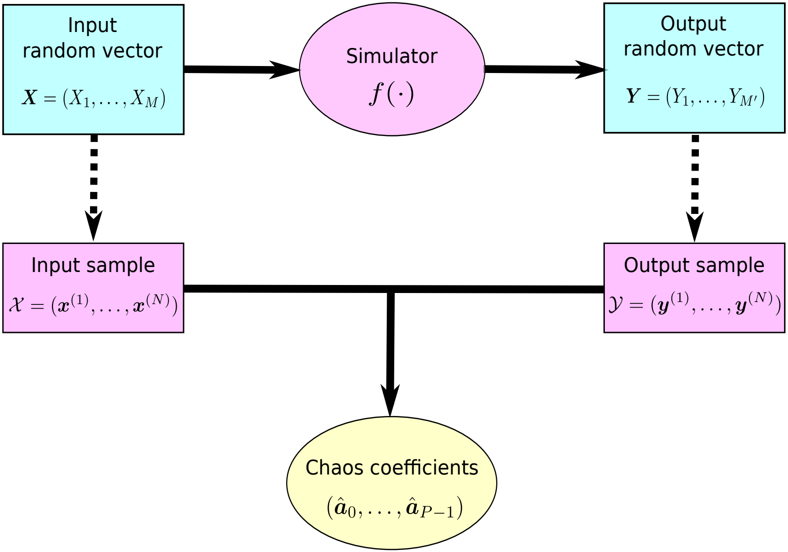
- **Easy case :** The PDF type of X_i is listed in the table.

  → Choose ξ_i as the scaled version of X_i.

  (e.g. $\mathcal{N}(\mu,\sigma) \longrightarrow \mathcal{N}(0,1)$, $\mathcal{U}(a,b) \longrightarrow \mathcal{U}(-1,1)$)

- **Otherwise :** Choose a ξ_i with the same support as X_i.

Estimation of PCE coefficients



Coefficients estimation as a least squares problem

Let us consider the truncated PCE : $\tilde{Y} \;=\; \sum_{|\boldsymbol{\alpha}|<p} a_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$

or in vector notation : $\tilde{Y} \;=\; \boldsymbol{a} \,.\, \boldsymbol{\psi}(\boldsymbol{\xi})$

The PCE coefficients are the solution of the least squares problem :

$$\boldsymbol{a} \quad = \quad \arg\min_{\boldsymbol{b} \in \mathbb{R}^P} \; \mathbb{E}\left[\left(f \circ \mathcal{T}^{-1}(\boldsymbol{\xi}) \;-\; \boldsymbol{b} \,.\, \boldsymbol{\psi}(\boldsymbol{\xi}) \right)^2 \right]$$

Coefficients estimation by least squares

Actual coefficients

$$\boldsymbol{a} \quad = \quad \arg\min_{\boldsymbol{b}\in\mathbb{R}^P} \; \mathbb{E}\left[\left(f\circ\mathcal{T}^{-1}(\boldsymbol{\xi}) \; - \; \boldsymbol{b}\,.\,\boldsymbol{\psi}(\boldsymbol{\xi})\right)^2\right]$$

Estimated coefficients

$$\hat{\boldsymbol{a}} \quad = \quad \arg\min_{\boldsymbol{b}\in\mathbb{R}^P} \; \sum_{i=1}^{N} \left(f\circ\mathcal{T}^{-1}(\boldsymbol{\xi}^{(i)}) \; - \; \boldsymbol{b}\,.\,\boldsymbol{\psi}(\boldsymbol{\xi}^{(i)})\right)^2$$

The $\boldsymbol{\xi}^{(i)}$'s are selected at random

(Monte Carlo or quasi-Monte Carlo sample)

Closed-form least squares solution

Model evaluations

$$\mathcal{Y} = \begin{pmatrix} f \circ \mathcal{T}^{-1}(\boldsymbol{\xi}^{(1)}) \\ \vdots \\ f \circ \mathcal{T}^{-1}(\boldsymbol{\xi}^{(N)}) \end{pmatrix}$$

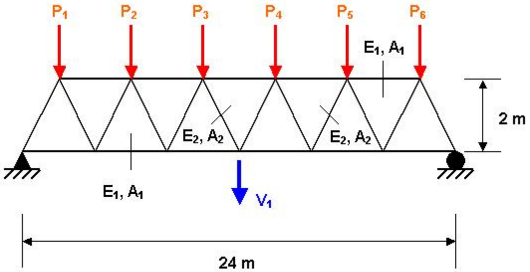
Chaos basis evaluations

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{\boldsymbol{\alpha}_0}(\boldsymbol{\xi}^{(1)}) & \dots & \psi_{\boldsymbol{\alpha}_{P-1}}(\boldsymbol{\xi}^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_{\boldsymbol{\alpha}_0}(\boldsymbol{\xi}^{(N)}) & \dots & \psi_{\boldsymbol{\alpha}_{P-1}}(\boldsymbol{\xi}^{(N)}) \end{pmatrix}$$

$$\hat{\boldsymbol{a}} = (\boldsymbol{\Psi}^{\mathsf{T}}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^{\mathsf{T}}\mathcal{Y}$$

Well-posed problem only if $N > P$

Application to the analysis of a truss structure



Output quantity of interest : Deflection v_1

10 random input variables :

- Young's moduli (lognormal)
- Cross-section areas (lognormal)
- Loads (Gumbel)

Construction and assessment of a PCE

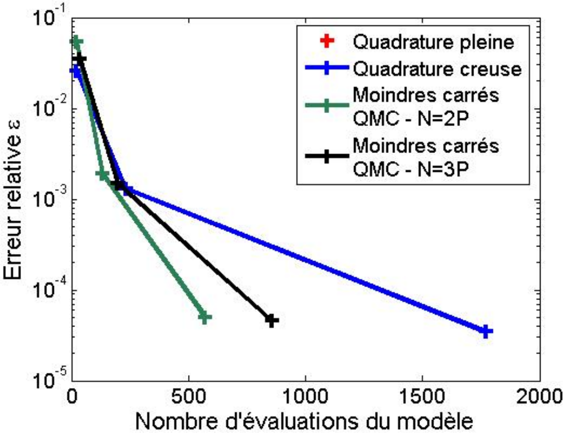
Parameter scaling : $\forall\, i \quad , \quad X_i \;\mapsto\; \xi_i \;\; , \;\; \xi_i \sim \mathcal{N}(0,1)$

Hermite PCE approximation : $v_1 \;\approx\; \sum_{|\boldsymbol{\alpha}| \leq p} a_{\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$

Estimation of approximation error :

$$\varepsilon \;=\; \sum_{i=1}^{\mathcal{N}} \Big(f(\boldsymbol{\xi}^{(i)}) \;-\; \hat{f}(\boldsymbol{\xi}^{(i)}) \Big)^2 \;/\; \tilde{\sigma}^2 \qquad , \qquad \mathcal{N} = 5,000$$

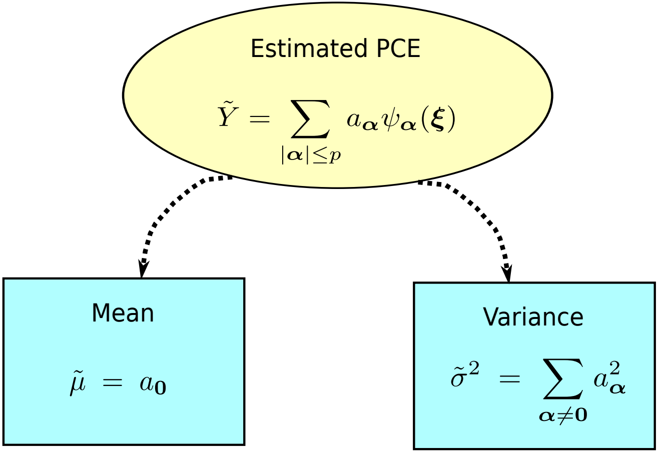
Convergence results (quasi-random samples)



Practical estimation of the approximation error

Use of a validation sample (size $N/3$ - $N/2$)	☺ Relatively accurate estimation ☹ New numerical experiments required
Cross-validation (recycle the already done experiments)	☹ Biased estimation (can be corrected) ☺ No additional computational cost ☺ Closed-form formula for least squares

Derivation of the response moments from PCE coefficients



Reminder : Sobol sensitivity indices

Objective : Quantify the sensitivity of output Y to input X_i

Sobol statistics : Ratio of variances $S[X_i] = \dfrac{\mathbb{V}\mathrm{ar}[\mathbb{E}[Y|X_i]]}{\mathbb{V}\mathrm{ar}[Y]}$

Very costly estimation using a two-sample Monte Carlo method

$(N \propto 10^5 - 10^6)$

Derivation of the Sobol indices from PCE coefficients

Estimated PCE

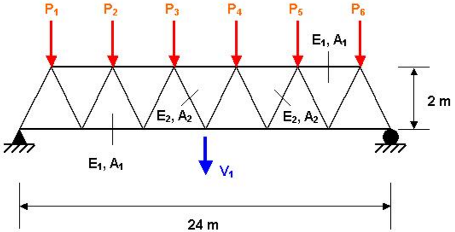
$$\tilde{Y} = \sum_{|\boldsymbol{\alpha}| \leq p} a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

Sobol indices

Polynomials depending only on ξ_i $\quad (\psi_{k_1}, \ldots, \psi_{k_m})$

$$\tilde{S}[X_i] = \frac{\sum_{j=1}^{m} a_{k_j}^2}{\sum_{|\boldsymbol{\alpha}| \neq \mathbf{0}} a_{\boldsymbol{\alpha}}^2}$$

Sensitivity analysis of the truss structure



Retained metamodel : PCE constructed by least squares with $N = 2P$

Sobol indices (%) :

E_1	A_1	P_3	P_4	P_2	P_5	E_2	A_2	P_1	P_6
37	37	8	8	4	4	1	1	0.5	0.5