Surface de réponse de type chaos polynomial

Géraud Blatman

EDF R&D, Département Matériaux et Mécanique des Composants

Journée Utilisateurs Open TURNS – 7 juin 2011

CHANGER L'ÉNERGIE ENSEMBLE

Plan

1. Développement sur la base du chaos polynomial

2. Estimation des coefficients

3. Stratégies de troncature de la base

Développement par chaos fonctionnel

Soit le modèle : $\mathbf{Y} = \mathcal{M}(\mathbf{X})$

$$\dim(\mathbf{X}) = M \qquad \dim(\mathbf{Y}) = M'$$

Si $\mathbf{Y}$ est un vecteur aléatoire de second ordre, i.e. $\mathbb{E}\left[\|\boldsymbol{Y}\|^2\right] < \infty$,
alors il admet la décomposition suivante, dite par chaos (Soize, 2004) :

$$\boxed{\mathbf{Y} = \sum_{j=0}^{\infty} \boldsymbol{a}_j \, \phi_j(\mathbf{X})}$$

(convergence L^2)

où $(\phi_j(\mathbf{X}))_{\mathbb{N}}$ est une base orthonormale de l'espace $L^2_{\mathbf{X}}(\mathbb{R}^M, \mathbb{R})$

et les $\boldsymbol{a}_j$ sont des vecteurs déterministes de $\mathbb{R}^{M'}$

Remarque

Les fonctionnalités du chaos polynomial de la version actuelle d'Open TURNS ne s'appliquent qu'à des réponses scalaires. Dans le cas d'une réponse vectorielle, il faut alors travailler composante par composante.

La prochaine version permettra cependant de traiter les réponses vectorielles.

Construction de la base – variables indépendantes (1)

On se donne une base orthonormée $\left(\phi_j^i(\mathsf{X}_i)\right)_{j\in\mathbb{N}}$ pour chacun des espaces $L^2_{\mathsf{X}_i}(\mathbb{R},\mathbb{R})$ (e.g. polynômes orthogonaux, ondelettes)

$$\mathbb{E}\left[\phi_j(\mathsf{X}_i)\ \phi_k(\mathsf{X}_i)\right] \;=\; \delta_{j,k}$$

Puis, en remarquant que : $L^2_{\mathbf{X}}(\mathbb{R}^M,\mathbb{R}) \;=\; \bigotimes_{i=1}^{M} L^2_{\mathsf{X}_i}(\mathbb{R},\mathbb{R})$

on construit une base orthonormée de cet espace en multipliant les bases 1D entre elles :

$$\phi_{\boldsymbol{\alpha}}(\mathbf{X}) \;=\; \prod_{i=1}^{M} \phi^i_{\alpha_i}(\mathsf{X}_i) \quad , \quad \boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_M) \in \mathbb{N}^M$$

Construction de la base – variables indépendantes (2)

Exemple. Soit X_1, X_2, X_3 i.i.d selon $\mathcal{U}([-1,1])$.

Pour chaque v.a., on choisit une base (1D) constituée des polynômes de Legendre normalisés :

$$\phi_0(x) = 1\ ,\ \ \phi_1(x) = \sqrt{\frac{3}{2}}x\ ,\ \ \phi_2(x) = \frac{\sqrt{5}}{2\sqrt{2}}\left(3x^2-1\right)\ ,\ \ldots$$

On écrit quelques fonctions de la base tensorisée :

$$\phi_{(0,0,0)}(\mathbf{X}) = 1 \qquad \phi_{(1,0,0)}(\mathbf{X}) = \sqrt{\tfrac{3}{2}}X_1$$

$$\phi_{(0,2,0)}(\mathbf{X}) = \tfrac{\sqrt{5}}{2\sqrt{2}}\left(3X_2^2-1\right) \qquad \phi_{(1,0,1)}(\mathbf{X}) = \tfrac{3}{2}X_1X_3$$

Construction de la base – variables indépendantes (3)

Cet usage de bases polynomiales est très courant en présence de variables indépendantes (chaos polynomial).

En pratique, on sait associer des familles de polynômes orthogonaux à bon nombre de variables aléatoires “standard” :

Distribution	Support	Polynôme
Normale $\mathcal{N}(0,1)$	$\mathbb{R}$	Hermite
Uniforme $\mathcal{U}([-1,1])$	$[-1,1]$	Legendre
Gamma $\mathcal{G}(k,1,0)$	$]0,\infty[$	Laguerre
Beta $\mathcal{B}(\alpha,\beta,-1,1)$	$]-1,1[$	Jacobi

Construction de la base

En pratique le modèle considéré dépend de variables aléatoires non “standardisées” $\{\mathsf{Z}_1, \ldots, \mathsf{Z}_\mathsf{M}\}$.

On reparamètre alors le problème par des variables de distribution « standard » via une transformation iso-probabiliste :

$$\mathsf{X}_i \;=\; F_{\mathsf{X}_i}^{-1}\left[F_{\mathsf{Z}_i}(\mathsf{Z}_i)\right] \qquad , \quad i = 1, \ldots, M$$

où F_{X_i} et F_{Z_i} désignent respectivement les fonctions de répartition de X_i et Z_i .

Remarque. Variables dépendantes. On peut développer la réponse sur un chaos fonctionnel faisant intervenir la copule. Sinon, pour certaines copules, on peut se ramener au cas indépendant (e.g. transfo. de Nataf).

Post-traitement du chaos

Moments de la réponse :

$$\mathbb{E}[\mathbf{Y}] = \boldsymbol{a}_0 \qquad \mathbb{V}\mathrm{ar}[\mathsf{Y}_\mathsf{i}] = \sum_{\boldsymbol{\alpha} \neq \mathbf{0}} a_{i,\boldsymbol{\alpha}}^2$$

$$\mathbb{C}\mathrm{ov}[\mathbf{Y}] = \boldsymbol{A}\,\boldsymbol{A}^T - \boldsymbol{a}_0\,\boldsymbol{a}_0^{\mathsf{T}} \qquad \boldsymbol{A} = (\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{P-1})^{\mathsf{T}}$$

Indices de sensibilité (Sudret, 2008 ; Crestaux, 2009):

$$S_j^i = \frac{\mathbb{V}\mathrm{ar}[\mathbb{E}[\mathsf{Y}_\mathsf{i}|\mathsf{X}_\mathsf{j}]]}{\mathbb{V}\mathrm{ar}[\mathsf{Y}_\mathsf{i}]} = \frac{\sum_{\boxed{\boldsymbol{\alpha} \in \mathcal{I}_j}} a_{i,\boldsymbol{\alpha}}^2}{\sum_{\boldsymbol{\alpha} \neq \mathbf{0}} a_{i,\boldsymbol{\alpha}}^2}$$

Multi-indices dont la j-ème composante est non nulle

Du développement théorique à l'approximation

En pratique, on doit tronquer le développement en série. Autrement dit, on doit choisir un sous-ensemble Λ fini et non vide de $\mathbb{N}^M$ t.q. :

$$\mathbf{Y} \approx \mathbf{Y}_\Lambda = \sum_{\boldsymbol{\alpha}\in\Lambda} \boldsymbol{a}_{\boldsymbol{\alpha}} \, \phi_{\boldsymbol{\alpha}}(\mathbf{X})$$

D'autre part, il faut choisir une procédure d'estimation des coefficients du chaos :

$$\mathbf{Y} \approx \widehat{\mathbf{Y}}_\Lambda = \sum_{\boldsymbol{\alpha}\in\Lambda} \widehat{\boldsymbol{a}}_{\boldsymbol{\alpha}} \, \phi_{\boldsymbol{\alpha}}(\mathbf{X})$$

Estimation des coefficients (1)

On génère *N* réalisations $\boldsymbol{X} = (\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(N)})$ de $\mathbf{X}$ (e.g. nombres pseudo-aléatoires, quasi-aléatoires), et on évalue le modèle en ces points :

$$\boldsymbol{Y} = (\boldsymbol{y}^{(1)} = \mathcal{M}(\boldsymbol{x}^{(1)}), \dots, \boldsymbol{y}^{(N)} = \mathcal{M}(\boldsymbol{x}^{(N)}))$$

Notations vectorielles/matricielles :

$$\boldsymbol{A} \;=\; (\boldsymbol{a}_0, \dots, \boldsymbol{a}_{P-1})^{\mathsf{T}} \qquad \text{matrice } (M' \times P)$$

$$\boldsymbol{\phi}(\boldsymbol{x}) \;=\; \left(\phi_{\boldsymbol{\alpha}_0}(\boldsymbol{x}), \dots, \phi_{\boldsymbol{\alpha}_{P-1}}(\boldsymbol{x})\right)^{\mathsf{T}}$$

$$\boldsymbol{\Phi} \;=\; \left\{\phi_{\boldsymbol{\alpha}_j}(\boldsymbol{x}^{(i)}) \;,\; i = 1, \dots, N \;,\; j = 0, \dots, P-1\right\}$$

Estimation des coefficients (2)

Méthode des moindres carrés. On cherche les coefficients qui minimisent l'erreur quadratique (Berveiller & Sudret, 2006) :

$$\widehat{\boldsymbol{A}} = \arg \min_{\boldsymbol{A} \in \mathbb{R}^{M' \times P}} \sum_{i=1}^{N} \left\| \boldsymbol{y}^{(i)} - \boldsymbol{A}\, \phi(\boldsymbol{x}^{(i)}) \right\|^2$$

La solution s'écrit : $$\widehat{\boldsymbol{A}}^{\mathsf{T}} = \left(\boldsymbol{\Phi}^{\mathsf{T}} \boldsymbol{\Phi} \right)^{-1} \boldsymbol{\Phi}^{\mathsf{T}} \boldsymbol{Y}$$

Pour que le problème soit bien posé, il faut que la matrice à inverser soit de rang plein, donc que $N \geq P$.

Troncature du chaos

Rappel : choisir un ensemble d'indices Λ judicieux pour construire une approximation par chaos :

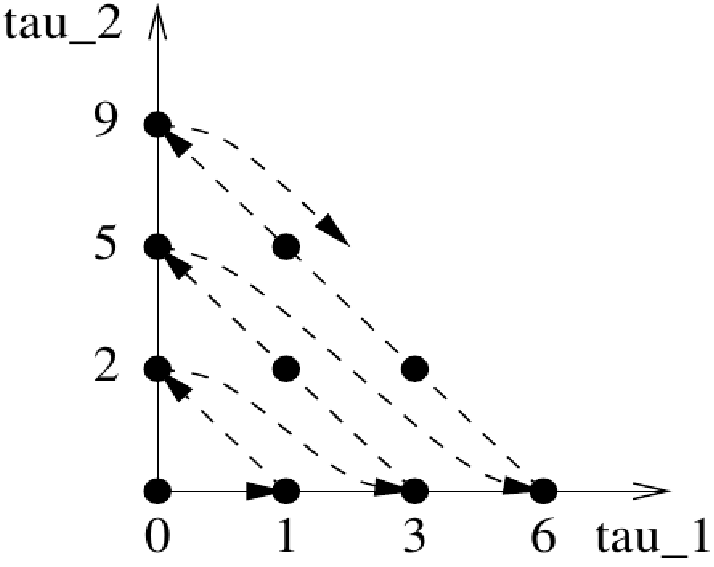
$$\mathbf{Y} \approx \mathbf{Y}_\Lambda = \sum_{\boldsymbol{\alpha} \in \Lambda} \boldsymbol{a}_{\boldsymbol{\alpha}} \, \phi_{\boldsymbol{\alpha}}(\mathbf{X})$$

Stratégie d'énumération de la base

Tronquer la série du chaos exige d'adopter une stratégie d'énumération de la base, autrement dit d'une bijection :

$$\begin{array}{rcccl} \tau & : & \mathbb{N} & \rightarrow & \mathbb{N}^M \\ & & j & \mapsto & \boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_M\} \end{array}$$

La stratégie implémentée dans Open TURNS est illustrée ci-contre en dimension M=2 :

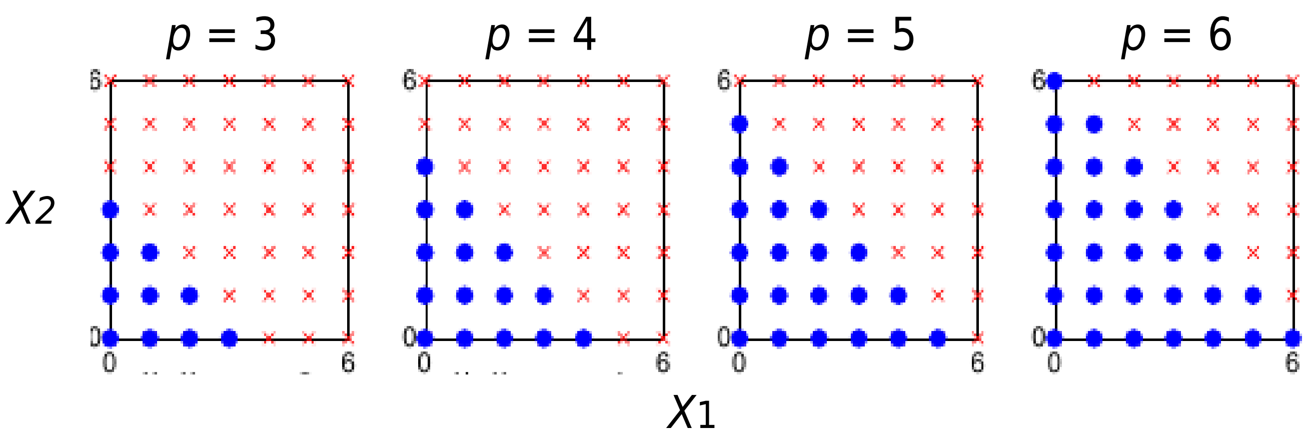


Fixed Strategy (1)

Principe : tronquer la série de sorte à ne conserver que les *P* premiers termes. Typiquement on garde tous les polynômes de la base dont le degré total est inférieur ou égal à *p* :

$$\Lambda = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^M : \sum_{i=1}^{M} \alpha_i \leq p \right\}$$

Exemple. Monômes retenus dans un cas à *M*=2 variables.



Fixed Strategy (2)

Limitations :

1. Choix du degré arbitraire

2. Forte inflation de termes quand p et M augmentent :

$$\mathrm{card}(\Lambda) = P = \frac{(M+p)!}{M!\,p!}$$

Comme le nombre N de simulations du modèle doit être supérieur à P pour résoudre le problème de moindres carrés, le coût de calcul peut se révéler inabordable

Sequential Strategy – cas d'une réponse scalaire

Principe : Enrichir la base jusqu'à atteindre une précision-cible.

Algorithme :

1. La base initiale ne comporte que le premier terme :

$$\Lambda := \{\boldsymbol{\alpha}_0\}$$

2. Si l'erreur d'apprentissage est inférieure à une erreur-cible ε_c , on stoppe l'algorithme.

3. On ajoute le terme suivant à la base courante :

$$\Lambda := \Lambda \cup \{\boldsymbol{\alpha}_k\}$$

et on retourne à l'étape 2.

Erreur d'apprentissage

Il s'agit de l'erreur entre le modèle et le chaos polynomial, aux points ayant servi à construire ce dernier :

$$\varepsilon = \frac{\sum_{i=1}^{N} \left\| \boldsymbol{y}^{(i)} - \boldsymbol{A}\, \phi(\boldsymbol{x}^{(i)}) \right\|^2}{\sum_{i=1}^{N} \left\| \boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}} \right\|^2}$$

où :

$$\bar{\boldsymbol{y}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{y}^{(i)}$$

Cleaning Strategy – cas d'une réponse scalaire

Principe : Trouver la meilleure approximation à *P* termes.

Algorithme :

1. La base initiale est composée des *P* premiers termes :

$$\Lambda := \{\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_{P-1}\}$$

2. On estime les coefficients par moindres carrés
3. On élimine tous les coefficients tels que :

$$|a_{\boldsymbol{\alpha}_j}| \;\leq\; \varepsilon \cdot \max_{\boldsymbol{\alpha} \in \Lambda} |a_{\boldsymbol{\alpha}}|$$

   et on actualise Λ .

4. Si $\mathrm{card}(\Lambda) = P$ on stoppe l'algorithme. Sinon on ajoute le terme suivant : $\Lambda := \Lambda \cup \{\boldsymbol{\alpha}_P\}$, et on retourne à l'étape 2.