Journées Incertitudes
CEA Bruyères le Châtel
3-4 Octobre 2007

Apprentissage de modèles - Plans d’expériences

Jean-Marc Martinez, Thierry Crestaux
CEA DEN/DANS/DM2S Saclay

Introduction

- En simulation numérique, certaines études nécessitent de nombreuses évaluations du modèle numérique
- Coût de calcul prohibitif, recours à des représentations simplifiées $\simeq$ modèle(s) numérique(s) dans le domaine d'étude
- Les représentations simplifiées peuvent être obtenues :
  - à partir du modèle physique/numérique par des techniques de réduction (POD)
  - par régression, plans d'expériences et surfaces de réponse (méta-modèles) ou apprentissage statistique

Fil conducteur ...

- **Méta-modèles linéaires** : intérêt d'une base orthogonale (comme les polynômes de chaos) pour les analyses de sensibilité globale (Anova fonctionnelle, indices de Sobol)
- **Analyse de l'erreur fonctionnelle** $\rightarrow$ intérêt des plans d'expériences optimaux pour réduire a priori l'erreur quadratique moyenne (G-Optimalité)

Intérêt des bases orthogonales pour méta-modèles linéaires

Modèle générique $x \in R^d \rightarrow f(x)$, densité $p(x) = \Pi_{i=1}^{d} p_i(x)$, $f \in \mathcal{H} \subset L^2(p)$, base hilbertienne $\Psi_k, k = 0, 1, \ldots$:

$$\begin{aligned} < \Psi_k, \Psi_l > &= \delta_{kl} \\ f(x) &= \sum_{k=0}^{\infty} \alpha_k \Psi_k(x) \ \underline{\texttt{vrai modèle}} \end{aligned}$$

Sous espace $\{\Psi_k, k = 0, \ldots, p\}$ ensemble des méta-modèles

$$h(x) = \sum_{k=0}^{p} \beta_k \Psi_k(x) \ \underline{\texttt{méta-modèle}}$$

<u>Rappel</u> : la meilleure approximation quadratique : $\beta_k = \alpha_k$.
<u>Intérêt</u> : Analyse de sensibilité globale (ANOVA fonctionnelle de $f()$, indices de Sobol) estimée directement via les β_k.

Polynômes de chaos versus Régression

Les fonctions de bases peuvent être représentées par des polynômes de chaos multidimensionnels et normalisés. Les coefficients sont approchés par intégration numérique (méthode non intrusive) :

$$\begin{array}{rcl} & \alpha_k = & \int f(x)\Psi_k(x)p(x)dx \\ \simeq & \beta_k = & \sum_i f(x^i)\Psi_k(x^i)\omega^i \end{array}$$

sous forme matricielle :

$$\beta = Z\Omega Y$$

... moindres carrés via un n-échantillon ($n > p + 1$)

$$\beta = (Z'Z)^{-1}Z'Y$$

$\Rightarrow$ choix de l'échantillon ?

Analyse de l'erreur

L'erreur fonctionnelle se décompose :

$$\begin{aligned} R(f) &= \int (f(x) - h(x))^2 p(x) dx \\ &= \underbrace{\sum_{k=0}^{p} (\beta_k - \alpha_k)^2}_{\texttt{estimation}} + \underbrace{\sum_{k=p+1}^{\infty} \alpha_k^2}_{\texttt{troncature}} \end{aligned}$$

$\Rightarrow$ Pour une base fixée, on ne peut que réduire l'erreur d'estimation.

Erreur d'estimation

En notant y_t le vecteur de composantes $=$ partie tronquée :

$$y_{ti} = \sum_{k=p+1}^{\infty} \Psi_k(x^i), i = 1, \ldots$$

l'erreur d'estimation est :

$$R_e(f) = y_t' Z (Z'Z)^{-1} Z' y_t$$

Décomposition en Valeurs Singulières $Z = USV'$

$$\boxed{R_e(f) = ||S^{-1}U'y_t||_2^2}$$

Pour réduire a priori l'erreur d'estimation, sélectionner une *bonne* matrice d'information (Fischer) $\rightarrow$ stratégie des plans d'expériences optimaux.

Plans d'expériences optimaux

En apprentissage statistique, hypothèse de régression (O. Chapelle, S. Gazut) : erreurs de troncature y_t *iid* (moyenne nulle, variance σ^2).

$$R_e(f) = \sigma^2 \ \mathtt{Trace}[(Z'Z)^{-1}] \Rightarrow \ \textbf{A-Optimalité}$$

En approximation, hypothèse non probabiliste $||y_t||_2^2 \leq \sigma^2$ et en notant ρ le rayon spectral :

$$R_e(f) \leq \sigma^2 \rho[(Z'Z)^{-1}] \Rightarrow \ \textbf{E-Optimalité}$$

$\Rightarrow$ intérêt de l'analyse spectrale de la matrice de dispersion $(Z'Z)^{-1}$ même dans le cas *non probabiliste*

Pour conclure ...

Plans d'expériences optimaux, outils d'analyse pour la construction de modèles simplifiés même si en simulation numérique

- les expériences numériques sont déterministes
- le méta-modèle postulé n'est pas le *vrai modèle*

à comparer aux méthodes d'échantillonnage : aléatoires, pseudo-aléatoires, déterministes (NISP), ...

Une alternative $\rightarrow$ simplifier le modèle "physique" par des techniques de réduction de modèles (projection/bases) :

- la réduction de modèles (Proper Orthogonal Decomposition) : Ecole d'Eté CEA-EdF-INRIA 2008, Patera (MIT), Made (P.M. Curie), Masmoudi (P. Sabatier), Kunisch (Univ. Graz Autriche), Gerbeau (INRIA)
- les méthodes spectrales intrusives (Galerkin)