Construction d'approximation globale par approche bayésienne pour l'optimisation de problèmes fortement non linéaires

Bruno Soulier

ENS de Cachan
Département Génie Mécanique
61 avenue du Président Wilson 94235 Cachan
bruno.soulier@ens-cachan.fr

Résumé :

Dans le domaine de la Conception Optimale et Robuste de Produits et des Procédés, l'utilisation des surfaces de réponses et de manière générale des méta-modèles permet de lisser les fonctions objectif et contraintes d'optimisation, et d'en calculer explicitement les gradients. Nous présentons dans cet article la formulation de différents modèles d'approximation (polynomiale, krigeage, approximation cumulative). Les limites des techniques classiques de construction de surfaces de réponses polynomiales par régression des moindres carrés dans le cadre de l'exploitation des simulations numériques sont mises en évidence. Par opposition à ces méthodes, nous présentons dans cet article un modèle d'approximation interpolant que l'on peut qualifier d'approximation diffuse. Ce modèle qui s'inspire des méthodes SPH (smooth particle hydrodynamics), présente de très bonnes propriétés de flexibilité et de reproduction de fortes linéarités. Nous montrons les similitudes avec les techniques de krigeage, avant d'illustrer sa construction sur un exemple analytique.

Abstract :

In the optimal and robust design field, the response surface methodology are used to build surrogate approximations with non-random, deterministic computer analysis. The surrogate-based analyse and optimisation provide smooth objective and constraints functions. After reviewing of some response surface models (polynomial, kriging, cumulative approximation), the limits of statistical techniques to approximate numerical experiments are highlighted. In opposition to traditional techniques like fitting polynomial models by least squares regression, the surrogate approximation presented in this paper, based on a cumulative interpolation, takes as a starting point the SPH methods (smooth particle hydrodynamics). This interpolation model generates continuous and derivable surface response on the design domain and makes it possible to restore strong non-linearities of objective and constraints functions.

Mots-clefs :

Méta-modèles ; plans d'expériences ; optimisation

1 Introduction

La conception intégrée - englobant les activités de conception, de fabrication et de contrôle de produits - fait appel de façon massive à la simulation numérique où les phénomènes simulés sont de plus en plus complexes, que ce soit par la physique ou la taille des problèmes envisagés. Pour aider l'ingénieur - concepteur dans ses démarches d'optimisation qui deviennent de plus en plus complexes par leur caractère multi-niveaux ou multidisciplinaires, des méthodes numériques d'exploration de l'espace des conceptions par approximation globale sont développées. Ces méthodes sont basées sur le couplage entre les méthodes des plans d'expériences, la construction des surfaces de réponse, les méthodes numériques d'optimisation, et les approches non-déterministes pour la conception robuste.

Les travaux présentés dans cet article ont pour objectif d'intégrer des méthodes de construction de surfaces de réponses (RSM, méta-modèles), dans les processus de conception optimale de systèmes, de structures ou de procédés Soulier *et al.* (1997), Bosselut *et al.* (1997).

De manière générale, les fonctions objectif et contraintes d'un problème d'optimisation structurale sont implicites par rapport aux variables de conception. Leurs évaluations et le calcul des gradients sont alors synonymes d'analyses éléments finis et constituent, plus particulièrement dans les problèmes d'optimisation fortement non-linéaire, l'opération la plus coûteuse de l'algorithme de minimisation. De ce fait, les méthodes d'optimisation les plus efficaces font appel aux concepts d'approximation où la résolution directe du problème d'optimisation initial est remplacée par une série de problèmes approchés et explicites. Dans le cadre de l'optimisation de problèmes fortement non-linéaires comme en dynamique rapide, les réponses obtenues sont instables et polluées par un bruit numérique. En conséquence, les problèmes d'optimisation qui en découlent sont mal conditionnés et présentent de nombreux minima locaux. Un algorithme d'optimisation (figure 1) adapté à cette problématique reposant sur le couplage d'une méthode de programmation mathématique avec la construction de surfaces d'interpolation est proposé dans Richard *et al.* (2002) et Soulier *et al.* (2003). Cet algorithme comprend deux étapes principales :

- une phase *exploratoire* permettant d'identifier sur le domaine initial différentes zones potentiellement optimales (appelées vallées) ;
- une phase d'*optimisation* au cours de laquelle plusieurs minima locaux sont obtenus.

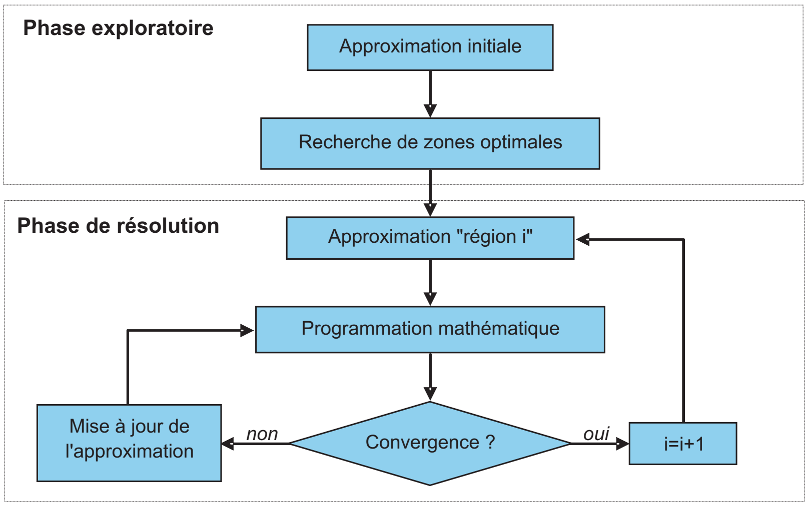

FIG. 1 – Couplage Optimisation/RSM

La précision de la solution et l'efficacité de ce processus itératif sont directement liées à la qualité et à la quantité d'informations fournies par ces approximations. Ces techniques d'approximations sont qualifiées de locales lorsqu'elles reposent sur la linéarisation des fonctions Fleury *et al.* (1984) (approximation d'ordre 1 construites à partir des dérivées des fonctions). Elles sont qualifiées d'approximations globales dans le cas d'identification de modèles de régression polynomiale ou de réseaux neuronaux.

Les trois étapes constitutives des approches par approximations globales sont :

- le choix de la classe ou du type de modèle (polynomial, krigeage, interpolation par approche cumulative, ...) ;

- l'estimation des coefficients du modèle (régression multilinéaire, maximum de vraisemblance, ...) ;
- et *in fine* le choix de l'échantillon conduisant au meilleur estimateur par des techniques de plans d'expériences (orthogonaux, composites, D-optimaux).

2 Limites des surfaces de réponses polynomiales

Dans une démarche d'analyse en surface de réponse de type polynomiale, la réponse y à une entrée multidimensionnelle $\mathbf{x} = (x_1, \ldots, x_k)$ constituée des k variables de conception x_i est généralement modélisée par un polynôme du second degré :

$$y = f(\mathbf{x}) + e \quad avec \quad f(\mathbf{x}) = a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} x_i x_j \tag{1}$$

où les e_i correspondant aux écarts entre les réponses simulées et les réponses du modèle sont supposés suivre une loi normale de moyenne nulle et de variance σ^2 constante sur le domaine d'étude. Les coefficients a_i sont estimés par régression linéaire à partir d'un échantillon de N simulations $\mathbf{y}^t = (y_1, \ldots, y_N)$ aux N points $\mathbf{X}^t = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$:

$$\hat{\mathbf{a}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{y} \quad avec \quad E(\mathbf{e}) = 0, \quad E(\mathbf{e}\mathbf{e}^t) = \sigma^2 \mathbf{I}_N \tag{2}$$

La prédiction de l'approximation en un site $\mathbf{z}$ est définie par l'estimation $\hat{f}(\mathbf{z})$ et la variance $V[\hat{f}(\mathbf{z})]$:

$$\hat{f}(\mathbf{z}) = \mathbf{z}^t\mathbf{a} \quad et \quad V[\hat{f}(\mathbf{z})] = (\mathbf{z}^t(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{z})\sigma^2 \tag{3}$$

Cette démarche n'est pas adaptée à la modélisation de systèmes déterministes pour deux raisons principales :

- les classes de fonction ne sont pas assez flexibles ; même si on augmente le degré des polynômes, on obtient toujours au mieux le meilleur estimateur dans la classe considérée de fonctions ;
- l'hypothèse de normalité et de variance constante sur l'ensemble du domaine d'étude n'a pas vraiment de sens dans le cadre des simulations numériques.

3 Modèle d'interpolation cumulative

Par opposition aux techniques classiques de construction de surfaces de réponses polynomiales par régression linéaire, la modélisation que nous présentons Richard *et al.* (2002), basée sur une approche bayésienne, est inspirée des techniques SPH (Smooth Particle Hydrodynamics) Gingold *et al.* (1977) Monaghan *et al.* (1983) Nayrolles *et al.* (1992), initiées dans les années 80 dans le domaine de l'astrophysique. Rasmussen (1998) propose un modèle d'approximation cumulative (4) qui consiste à construire une approximation de la fonction y étudiée, par une somme pondérée $\tilde{y}$ de fonctions de forme, notées ϕ_i , de type exponentielle ou spline qui engendrent des approximations continues et dérivables sur leur domaine d'étude :

$$\tilde{y}_R(x) = \frac{\sum_i \phi_i(x, x_i, k_i) y_i}{\sum_i \phi_i(x, x_i, k_i)} . \tag{4}$$

avec y_i la valeur mesurée de y au point x_i et $\phi_i(x, x_i, k_i)$ la fonction de forme associée aux points d'observations. L'approximation est par définition non exacte aux points de simulation. En effet, l'exactitude de l'approximation s'écrit :

$$\tilde{y}(x_k) = y_k\,, \quad \forall x_k \in D\,. \tag{5}$$

soit :

$$\frac{\sum\limits_i \phi_i(x_k, x_i, k_i) y_i}{\sum\limits_i \phi_i(x_k, x_i, k_i)} = y_k\,, \quad \forall x_k \in D \Longrightarrow \phi_i(x_k, x_i, k_i) = 0\,. \tag{6}$$

Nous définissons un modèle d'interpolation cumulatif que l'on peut qualifier d'interpolation diffuse par :

$$\tilde{y}(x) = \frac{\sum\limits_i \phi_i(x) a_i}{\sum\limits_i \phi_i(x)} \tag{7}$$

avec ϕ_i des fonctions de forme de type gaussiennes. Les coefficients a_i sont obtenus par une méthode de collocation.

Le modèle proposé permet de restituer les fortes non-linéarités des fonctions objectif et contraintes d'optimisation et engendre des surfaces de réponses continues et dérivables sur le domaine d'étude ce qui permet d'en calculer explicitement les gradients. De plus, la flexibilité de ces modèles facilite la remise à jour des surfaces d'interpolation au cours du processus d'optimisation, ce qui permet d'améliorer localement la qualité des approximations.

4 Analogie avec le krigeage

Les modèles de krigeage $\tilde{y}_{krig}$ sont des modèles statistiques qui reposent sur la réalisation d'un processus stochastique noté Y. Ce processus est régi par les deux hypothèses suivantes :

- Y suit une loi de distribution normale de moyenne μ et de variance σ^2 sur le domaine d'étude ;
- les réalisations du processus en deux sites x et x_0 sont liées par une fonction de corrélation spatiale $r(d)$, où d est la distance séparant les points x et x_0.

Connaissant la réalisation du processus Y_0 au point x_0, le processus primitif permet de calculer la moyenne $\mu_{x|x_0}$ et la variance $\sigma^2_{x|x_0}$ au site x définis par :

$$\left\{ \begin{aligned} \mu_{x|x_0} &= E\left[y\left(x\right)|Y_0\right] \\ \sigma^2_{x|x_0} &= E\left[\left(y\left(x\right) - E\left[y\left(x\right)\right]\right)^2 |Y_0\right] \end{aligned} \right. \tag{8}$$

Soit :

$$\left\{ \begin{aligned} \mu_{x|x_0} &= \mu + r(d)(Y_0 - \mu) \\ \sigma^2_{x|x_0} &= \sigma^2(1 - r(d)) \end{aligned} \right. \tag{9}$$

Les fonctions de corrélation classiquement utilisées pour construire des modèles statistiques sont les suivantes :

- fonction de corrélation linéaire définie par :

$$\left\{ \begin{array}{lll} r(d) = 1 - (1-\rho)|d| & si & |d| \leq \frac{1}{1-\rho} \\ r(d) = 0 & sinon & \end{array} \right. \tag{10}$$

- fonction de corrélation exponentielle définie par :

$$r(d) = \rho^{|d|} \tag{11}$$

- fonction de corrélation gaussienne définie par :

$$r(d) = \rho^{d^2} \tag{12}$$

5 Comparaison des approximations

Nous illustrons la construction de modèles d'approximations globale à partir de la fonction de Rosenbrock. Cette fonction analytique dont l'expression est la suivante est fortement non-linéaire et présente une vallée très allongée en forme d'arc de cercle.

$$f(x_1, x_2) = 100(x_1 - x_2^2)^2 + (1 - x_1)^2 \tag{13}$$

Comme spécifié précedemment, le choix des sites expérimentaux est réalisé suivant une stratégie plans d'expériences. Le modèle d'interpolation cumulative permet d'accroître la qualité du modèle de manière incrémentale par remise à jour du modèle au cours de l'ajout de nouvelles simulations. La figure suivante présente l'évolution du modèle pour un plan d'expérience orthogonal à 3 niveaux par facteurs auquel nous ajoutons successivement 2, 5 et 7 simulations supplémentaires. Nous observons que l'approximation reproduit bien l'allure réelle de la fonction de Rosenbrock à partir des ajouts réalisés dans la zone présentant les plus fortes non-linéarités.

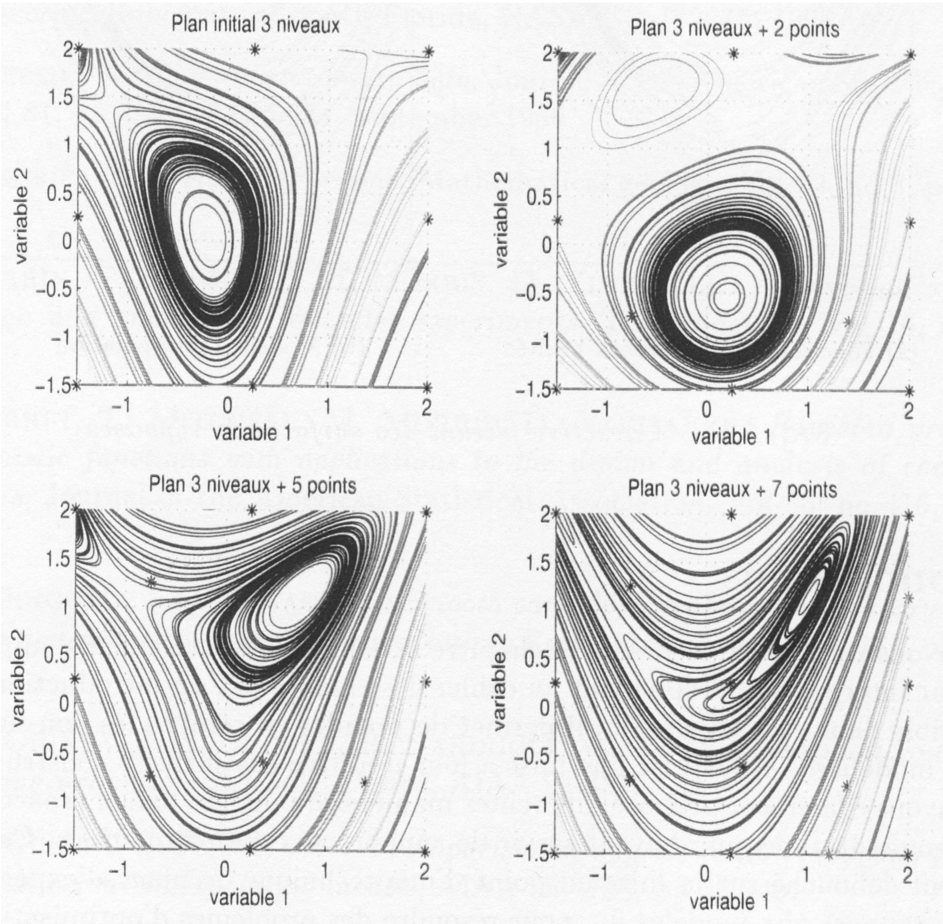

FIG. 2 – Evolution de l'approximation en fonction du nombre d'expériences

6 Conclusions

Partant d'une série d'analyses exactes réalisées dans le cadre d'un plan d'expériences, une approximation explicite globale est construite sous la forme d'une surface de réponse. Le modèle d'interpolation cumulative présenté permet la construction d'approximations globales qui, outre de très bonnes propriétés de flexibilité et de reproduction de fortes linéarités, sont continues et dérivables sur l'ensemble du domaine d'étude. L'utilisation de l'approximation dans les stratégies d'optimisation de problèmes non-linéaires, permet d'identifier plusieurs minima locaux et la remise à jour itérative de l'approximation au cours de la procédure d'optimisation. Cette méthode présente un intérêt dans la construction de méta-modèles pour les procédures d'optimisation multidisciplinaire.

Références

Gingold, R.A., Monaghan, J.J. 1977 Smooth particle hydrodynamics : theory and application to non-sperical stars. *Monthly Notices Roy. Astronm. Soc.* **181** 375-189

Fleury, C., Braibant, V. 1984 Optimization methods, Computer Aided Analysis and Optimization of Mechanical System Dynamics. *E.J. Haug Editor, NATO ASI Series, Computer System Sciences* **9** 637-677

Monaghan, J.J., Gingold, R.A. 1983 Shock simulation by the Particle Method SPH. *Journal of the Computational Physics* **52** 374-389

Nayrolles, B., Touzot, G., Villon, P. 1992 Generalizing the Finite Element Method : Diffuse approximation and diffuse elements. *Computational Mechanics* **10** 307-318

Rasmussen, J. 1998 Non linear programing by cumulative approximation refinement. *Structural Optimization* **15** 1-7

Soulier, B. Richard, L. Hazet, B. Braibant, V. 2003 Crashworthiness optimization using a surrogate approach by stochastic response surface *G. Gogu D. Coutellier P. Chedmail P. Ray Editor, Recent Advances in Integrated Design and Manufacturing in Mechanical Engineering* , 159-168

Richard, L. Soulier, B. 2002 Méthode d'optimisation en dynamique rapide : construction de modèles d'interpolation *Méthodes non-déterministes en conception, Journées Thématique AIP-PRIMECA, Cachan.*

Bosselut, D. Régnier, G. Soulier, B. 1997 Finite Element model of PWR control rod cluster. Parametrical study of vibrating behavior by experimental design method *Proceedings of the 14th International Conference on Structural Mechanics in Reactor Technology - SMIRT 14.*

Soulier, B. Régnier, G. Bosselut, D. 1997 Plans d'expériences numériques : application à l'étude du comportement vibratoire des grappes de commande d'un réacteur à eau pressurisée *Actes du 3ème Colloque National en Calcul des Structures, Giens.*