

Construction de modèles de substitution dynamiques et multifidélité

Matthias De Lozzo (Matthias.De_Lozzo@onera.fr)

Laboratoire : ONERA Toulouse - École : INSA Toulouse - Entreprise : EPSILON

Thèse CIFRE encadrée par Béatrice Laurent (IMT) et Patricia Klotz (ONERA)

Résumé *L'utilisation de modèles de simulation numérique précis en physique est souvent coûteuse en temps de calcul. Une alternative consiste à leur préférer des modèles de substitution à temps de réponse rapide, dont les paramètres sont obtenus par apprentissage statistique d'échantillons « entrées-sorties » du système physique observé ou simulé ou de modèles réduits physiques : c'est l'aspect multifidélité. On a proposé durant la première année un modèle heuristique approchant l'évolution temporelle de la température en des points d'intérêt pour un système dynamique. Ce modèle a été validé sur un cas test académique.*

A. CONTEXTE

Les **modèles de simulation numérique** (différences finies, éléments finis,...) permettent de représenter avec une grande précision les systèmes physiques. Cependant ils ont le défaut d'être coûteux en temps de calcul et dans certaines applications, on leur préfère des **modèles de substitution à temps de réponse rapide**. En particulier en optimisation, en problèmes inverses ou en propagation d'incertitudes, des domaines où le modèle est très sollicité, ces substitutions sont d'une réelle nécessité.

Ces modèles de substitution, aussi appelés modèles réduits ou métamodèles, sont de **basse fidélité** et possèdent une **forme analytique paramétrée** à partir de l'apprentissage d'échantillons « entrées-sorties » de haute fidélité. Ces derniers proviennent de modèles de simulation numérique ou de résultats expérimentaux. Les métamodèles les plus utilisés sont les approximations polynomiales, les processus gaussiens ou encore les réseaux de neurones artificiels. On les qualifie souvent de **boîtes grises ou noires** car leur formulation n'a pas, ou peu, d'interprétation physique.

Le domaine sous-jacent à la construction de ces modèles de substitution est l'**apprentissage statistique** [HF08] qui consiste à identifier les paramètres d'un modèle à partir d'observations composées de sorties associées à des entrées issues de **plans d'expériences** ; ces observations sont regroupées dans des bases d'apprentissage.

B. APPLICATION THERMIQUE

Dans le cadre de cette thèse CIFRE, l'application visée est l'**analyse thermique des composants électroniques d'un avion**, qui consiste à vérifier leur dimensionnement en déterminant la température des points chauds. Les simulations numériques impliquées sont très coûteuses car elles tiennent compte de différentes échelles : l'avion, les équipements qu'il comporte et les composants électriques présents dans ces derniers, ainsi que les interactions entre ces entités et leur environnement. Une approche alternative consiste

à construire des modèles réduits physiques de ces équipements de niveaux différents et de les introduire dans le processus de simulation. On parle alors de **modélisation multi-échelle**.

Les échantillons de simulations thermo-fluidiques ou de résultats expérimentaux, mais aussi de modèles réduits thermiques qui sont des simplifications physiques du modèle numérique haute fidélité. On parle de **modélisation multi-fidélité**.

C. OBJECTIF

On cherche dans cette thèse une **méthodologie pour la construction de modèles de substitution dynamiques à temps de réponse rapide, valides en régimes permanent et transitoire. Les bases d'apprentissage pourront par la suite être de fidélités différentes** : données expérimentales ou issues de modèles de simulation numérique ou de modèles réduits physiques. On pourra s'inspirer de [K000] pour traiter ce problème d'observations multifidélité.

D'un point de vue très général, on cherche à approcher en quelques points de l'espace la solution en temps d'un système d'équations aux dérivées partielles :

$$\begin{cases} \psi(Dy(\underline{x}, t), f(\underline{x}, t)) = 0, (\underline{x}, t) \in \Omega \times]0, \tau] \\ T(\underline{x}, 0) = T_0(\underline{x}), \underline{x} \in \Omega \\ y(\underline{x}, t) = u_1(\underline{x}, t), \underline{x} \in \partial\Omega_1 \\ \nabla y(\underline{x}, t) = u_2(\underline{x}, t), \underline{x} \in \partial\Omega_2 \end{cases} \quad (1)$$

D étant un opérateur différentiel. (1) est la mise en équation d'un modèle distribué. Dans l'exemple de l'équation de la chaleur linéaire, la première ligne devient $\frac{\partial y(\underline{x}, t)}{\partial t} = \alpha \Delta f(\underline{x}, t)$.

D. CONSTRUCTION D'UN MODÈLE DE SUBSTITUTION DYNAMIQUE

Les méthodes classiques pour modéliser une dynamique temporelle se trouvent dans la communauté de l'automatisme et plus précisément, de l'**identification de système** et du contrôle [Lju99]. Les **séries tem-**

porelles bouclées y forment une famille classique de modèles pouvant répondre à notre problématique, avec des versions linéaires et des versions non linéaires faisant intervenir un réseau de neurones.

Réseau de neurones artificiel

En présence d'une relation non linéaire complexe de la forme $\underline{y} = f(\underline{x})$, f peut être approchée par un réseau de neurones artificiel $f_{R.N.}(\underline{x}; \underline{\theta}) = (f_{R.N.}^{(1)}(\underline{x}; \underline{\theta}), \dots, f_{R.N.}^{(n_y)}(\underline{x}; \underline{\theta}))'$ où $f_{R.N.}^{(j)}(\underline{x}; \underline{\theta}) = \sum_{i=1}^{n_x} w_i^{(j)} \sigma \left(\sum_{l=1}^{n_x} w_{il} x_l + w_{i0} \right) + w_0^{(j)}$ avec σ la fonction sigmoïde : $\forall z \in \mathbb{R}, \sigma(z) = \frac{2}{1+e^{-2z}} - 1$. $\underline{\theta}$ renferme l'ensemble des paramètres $w_i^{(j)}$, w_{il} et $w_0^{(j)}$ de $f_{R.N.}$.

Parmi l'ensemble des modèles de substitution existants, les réseaux de neurones de cette forme, issus des perceptrons multicouches (PMC), ont de bonnes propriétés parmi lesquelles celle d'approximation universelle et de parcimonie.

On souhaite plonger ces modèles dans le cadre de l'apprentissage statistique et proposer un algorithme de sélection de modèles (ici le nombre n_n) alternatif à la validation croisée par exemple.

Notre modèle heuristique

Le modèle de substitution dynamique retenu trouve ses sources dans la lecture de [TN05, MT10, WL98] où il est question de modéliser la variation temporelle de la quantité d'intérêt par un réseau de neurones artificiels plutôt que la quantité d'intérêt elle-même. On peut lire selon les auteurs « modélisation de la variation instantanée » ou « modélisation de la dérivée moyenne ».

D'un point de vue modèle par bloc et non pas modèle distribué, la solution du système (1) en n_y points spatiaux d'intérêt peut être vue comme la solution

$\underline{y}(t) = (y^{(1)}(t) \dots y^{(n_y)}(t))'$, $t \in [0, \tau]$, d'une équation différentielle ordinaire tenant directement compte de variables exogènes $\underline{u}(t)$:

$$\begin{cases} \frac{d\underline{y}(t)}{dt} = f(\underline{u}(t), \underline{y}(t)) \\ \underline{y}(0) = \underline{y}_0 \end{cases} \quad (2)$$

où $f : \mathcal{X} \subset \mathbb{R}^{n_u} \times \mathcal{Y} \subset \mathbb{R}^{n_y} \rightarrow \mathcal{Y} \subset \mathbb{R}^{n_y}$ est une application non linéaire. Dans les applicatifs à traiter, les variables exogènes sont essentiellement des conditions aux limites et des puissances injectées.

Après avoir appliqué en t_k la formule de Taylor-Lagrange du premier ordre à $\underline{y}(t)$ et avoir utilisé (2), on aboutit au modèle heuristique suivant :

$$\begin{cases} z_{k+1} = z_k + h_k g(\underline{u}_k, z_k, h_k; \underline{\theta}) \\ z_0 = \underline{y}_0 \\ h_k = t_k - t_{k-1} \end{cases} \quad (3)$$

où g est un modèle de substitution classique approchant la nonlinéarité $f(\underline{u}_k, \underline{y}_k) + \frac{1}{2} \frac{d^2 \underline{y}(t)}{dt^2} \Big|_{t=\xi_k} h_k$, $t_{k-1} \leq \xi_k \leq t_k$.

Ce modèle est dit **bouclé** car il prend en entrée z_k , la sortie du métamodèle, et non pas \underline{y}_k , la sortie observée. La prise en compte de cette récurrence complexifie l'algorithme d'apprentissage tout en facilitant la recherche des paramètres optimaux du métamodèle g .

Par ailleurs, il permet de travailler à **pas de temps variable**, g approchant non seulement la variation instantanée mais aussi le terme de reste issu du développement limité. Ceci est un apport par rapport aux modèles que l'on peut trouver dans [TN05] par exemple. En pratique, le métamodèle g est un réseau de neurones.

RÉFÉRENCES

-
- [HF08] Tibshirani Hastie and Friedman. *The Elements of Statistical Learning (2nd edition)*. Springer-Verlag, 2008.
 - [KO00] Kennedy and O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 2000.
 - [Lju99] Lennart Ljung, editor. *System identification (2nd ed.) : theory for the user*. Prentice Hall PTR, 1999.
 - [Mas07] Pascal Massart. *Concentration Inequalities and Model Selection*. Lecture Notes in Math. Springer, 2007. École d'été de Probabilités de Saint-Flour 2003.
 - [MT10] Regina Paiva Melo Marin and Paulo Marcelo Tasinaffo. Uma metodologia de modelagem empírica utilizando o integrador neural de múltiplos passos do tipo adams-bashforth. *Revista Controle & Automação*, 21(5), 2010.
 - [TN05] Paulo M. Tasinaffo and Atair Rios Neto. Mean derivatives based neural euler integrator for nonlinear dynamic systems modeling. *Learning and Nonlinear Models - Revista da Sociedade Brasileira de Redes Neurais (SBRN)*, 3(2) :98–109, 2005.
 - [WL98] Yi-Jen Wang and Chin-Teng Lin. Runge-kutta neural network for identification of dynamical systems in high accuracy. *IEEE Transactions on neural networks*, 9(2) :294–307, 1998.

À PROPOS DU DOCTORANT ET DE LA THÈSE

Diplômé de l'INSA de Toulouse, j'effectue depuis le mois de novembre 2010 une thèse CIFRE avec l'entreprise de thermique EPSILON au sein du Département de Traitement de l'Information et Modélisation de l'ONERA Toulouse, sous la direction de Béatrice Laurent (IMT) et Patricia Klotz (ONERA). Le sujet porte sur la **construction de modèles de substitution multifidélité avec des applications en thermique** ; l'aspect **dynamique temporelle** du modèle est le premier aspect abordé dans cette thèse. Je travaille actuellement sur la construction d'un **algorithme de sélection de modèles** [Mas07] par moindres carrés pénalisés **pour les réseaux de neurones**.