Sujet de postdoc

Réduction de dimension semi-supervisée

Miguel MUNOZ ZUNIGA*,
Thomas PERRILLAT-BOTTONET†,
Olivier ZAHM‡

November 17, 2021

Dans le cadre de l'étude et de l'utilisation de codes de calcul, un besoin récurrent concerne la réduction de dimension supervisée des entrées du code afin de réduire les besoins computationnels et améliorer la qualité des analyses. En effet, déterminer les entrées pertinentes à l'explication des sorties d'intérêts permet notamment de réduire la taille des plans d'expériences nécessaire aux analyses (réduction du besoin en simulations couteuses), formuler des analyses de sensibilité sur des entrées influentes, construire des méta-modèles plus efficacement et de meilleure qualité, ou encore de donner de meilleurs taux de convergence des algorithmes d'optimisation.

Dans la littérature il est récurrent d'utiliser plutôt des stratégies non-supervisées (PCA, KL) pour réduire la dimension des entrées avec potentiellement dans un second temps des analyses de sensibilités comme complément supervisé. En effet l'approche non-supervisé a l'avantage de ne pas nécessiter de simulation (approche *a priori*). Les approches supervisées peuvent nécessiter un nombre de simulations relativement conséquent pour obtenir des résultats pertinents ce qui n'est pas toujours accessible en début d'analyse. On peut citer par exemple les cas où les simulations arrivent de manière séquentielle : assimilation de données en temps réelle, enrichissement adaptatif d'un plan d'expériences.

Dans ce context nous proposons une approche semi-supervisée dans laquelle un compromis entre réduction de dimension non-supervisée et supervisée est mise en place en fonction du nombre courant de simulations. Cette approche pourra par exemple se formuler comme un problème d'optimisation du type

$$\min_{B\in\mathbb{R}^{d\times m}} \left(\min_{\substack{f:\mathbb{R}^m\to\mathbb{R} \\ \text{any function}}} \mathbb{E}\left[(Y - f(B^T X))^2\right] + \lambda \min_{\substack{h:\mathbb{R}^m\to\mathbb{R}^d \\ \text{affine function}}} \mathbb{E}\left[\|X - h(B^T X)\|^2\right] \right) \quad (1)$$

avec $\lambda > 0$ un paramètre de régularisation qu'il conviendra de calibrer en fonction du nombre d'évaluation du modèle $Y = u(X)$ disponible. Alors que l'optimisation sur h peut être réalisé analytiquement (PCA), l'optimisation sur f est plus délicate. Nous envisageons d'utiliser des approches à noyau [1,2] ou bien des approches qui utilisent les gradients du modèle $\nabla u(X)$ [3,4,5]. Si ces dernières méthodes s'avèrent être très

*miguel.munoz-zuniga@ifpen.fr
†thomas.perrillat-bottonet@cea.fr
‡olivier.zahm@inria.fr

efficaces, leur utilisation est conditionnée à avoir accès aux gradients du modèle, ce qui n'est pas toujours le cas pour les codes de calcul utilisés. On envisage alors de développer des méthodes dites "gradient-free" qui approchent le gradient soit via des métamodèles, soit via un échantillonnage adapté du modèle comme dans [6]. Notre stratégie pourra également être étendue à la réduction de dimension pour des modèles à entrées mixtes : scalaires, fonctionnelles notamment en remplaçant dans la formulation (1) le terme BX par un opérateur qui sera astucieusement paramétrisé.

Enfin au-delà de la nécessaire validation de ces méthodes sur des cas académiques nous pourrons également appliquer ces stratégies sur à minima un des nombreux simulateurs utilisés et développés à l'IFPEN. Les applications qui nous serviront de fil rouge sont par exemple l'optimisation de la conception et du fonctionnement d'éoliennes marines ou de moteurs électrique, ou encore l'apprentissage de termes d'erreurs dans des codes de calcul CFD.

Ces stratégies seront également pertinentes pour répondre à des applications du LETI tel que le design automatique de filtre nécessitant d'optimiser différents paramètres : longueur d'onde de résonance, largeur de bande, etc. Des simulateurs électromagnétiques de différentes fidélités (avec et/ou sans informations de gradient) pourront être utilisés à ces fins en utilisant des approches de type multi-fidélité [5].

1. Fukumizu K, Bach FR, Jordan MI. Kernel dimension reduction in regression. The Annals of Statistics. 2009 Aug;37(4):1871-905.

2. Fukumizu K, Leng C. Gradient-based kernel dimension reduction for regression. Journal of the American Statistical Association. 2014 Jan 2;109(505):359-70.

3. Constantine PG, Dow E, Wang Q. Active subspace methods in theory and practice: applications to kriging surfaces. SIAM Journal on Scientific Computing. 2014;36(4):A1500-24.

4. Zahm O, Constantine PG, Prieur C, Marzouk YM. Gradient-based dimension reduction of multivariate vector-valued functions. SIAM Journal on Scientific Computing. 2020;42(1):A534-58.

5. Lam RR, Zahm O, Marzouk YM, Willcox KE. Multifidelity dimension reduction via active subspaces. SIAM Journal on Scientific Computing. 2020;42(2):A929-56.

6. Constantine PG, Eftekhari A, Wakin MB. Computing active subspaces efficiently with gradient sketching. In2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) 2015 Dec 13 (pp. 353-356). IEEE.