Advanced Engineering Informatics 30 (2016) 283–297

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Full length article

An active learning metamodeling approach by sequentially exploiting difference information from variable-fidelity models

Qi Zhou, Xinyu Shao, Ping Jiang *, Zhongmei Gao, Chaochao Wang, Leshi Shu

The State Key Laboratory of Digital Manufacturing Equipment and Technology, School of Mechanical Science and Engineering, Huazhong University of Science & Technology, 430074 Wuhan, PR China

ARTICLE INFO

Article history:
Received 3 September 2015
Received in revised form 8 April 2016
Accepted 12 April 2016

Keywords:
Variable fidelity metamodel
Already-acquired information
Active learning strategy
Metamodel-based design optimization
Kriging

ABSTRACT

Complex system engineering design optimization based on simulation is a very time-consuming, even computationally prohibitive process. To relieve the computational burden, metamodels are commonly used to replace the computation-intensive simulations. In this paper, an active learning variable fidelity (VF) metamodeling approach (AL-VFM) is proposed for the purpose of integrating information from both low-fidelity (LF) and high-fidelity (HF) models. In AL-VFM, Kriging metamodel is adopted to map the difference between the HF and LF models aiming to approach the HF model on the entire domain. Besides, a general active learning strategy is introduced in AL-VFM to make full use of the already-acquired information to guide the VF metamodeling. The already-acquired information represents the location of regions where the differences between the HF and LF models are multi-model, non-smooth and have abrupt changes. Several numerical and engineering cases with different degrees of difficulty verify the applicability of the proposed VF metamodeling approach.

1. Introduction

Computational simulation models have been widely used to explore design alternatives during preliminary design phase. In spite of sustained growths in computer capability and speed, the enormous computational expensive associated with high fidelity engineering simulation codes still makes it impractical to rely exclusively on high fidelity models for design and optimization. Just taking Ford Motor Company as an example, it was reported that it takes the company about 36–160 h to run one crash simulation for a full passenger car [1]. Indeed, it is still impractical to directly use these simulations with an optimizer to evaluate a lot of design alternatives when exploring the design space for an optimum [2,3]. This limitation can be addressed by adopting global metamodel (or surrogate), which can mimic the original system at a considerably reduced computational cost [4,5]. There are a lot of commonly used metamodels, such as Polynomial Response Surface (PRS) models [6], Kriging models [7], Artificial Neural Networks (ANN) models [8,9], Radial Basis Function (RBF) models [10], and Support Vector Regression (SVR) models [11]. A more detailed overview on various metamodeling techniques can refer to [12]. It is important to point out that the quality of the metamodels has a profound impact on the computational cost and convergence characteristics of the metamodel-based design optimization. The quality of the metamodels directly depends on the sample points at which the computer simulation or physical experiments are conducted. Generally, more sample points offer more information of the system, however, at a higher cost [13]. Less sample points require lower expense, while leading to inaccurate metamodels even distorted metamodels. Hence, conflict between high accuracy and low expense seems to be inevitable in building metamodels.

To ease this problem, variable-fidelity (VF) metamodeling approaches based on the interaction of high-fidelity (HF) and low-fidelity (LF) models have been widespread concerned [4,14]. A HF model is one that is able to accurately describing the physical features of the system but with an unaffordable computational expense, e.g., physical experiment, finite element, computational fluid dynamics, etc. A LF model is one that is able to reflect the most prominent characteristics of the system at a considerably less computationally demanding, e.g., numerical empirical formula. Commonly used VF metamodeling approaches are scaling methods, which tune the LF model according to the response values of the HF model. These scaling methods can be divided into two distinct types: local VF metamodeling approaches and global VF metamodeling approaches. In local VF metamodeling approaches, the scaling function is approximated using local metamodels, e.g., linear regression [15], first/s Taylor series [16–18]. The local

* Corresponding author.
E-mail address: jiangping@mail.hust.edu.cn (P. Jiang).

http://dx.doi.org/10.1016/j.aei.2016.04.004
1474-0346/

VF metamodeling approaches are easy to implement and can achieve a relative high accuracy within an appropriate trust region size, e.g., Chang et al. [15] used a multiplicative scaling approach to correct the response values of LF model to match the HF model. An application of this metamodel was tested on a wing-box model of a high-speed civil transport. Alexandrov et al. [16,17] integrated first-order additive and multiplicative scaling modeling method with the convergent techniques of nonlinear programming in engineering analysis and design; and have successfully applied this method to a 3-D aerodynamic wing optimization problem and a 2-D airfoil optimization problem, achieving a threefold savings and twofold savings in computing effort, respectively. The main shortcoming of these approaches is that they are only suitable for local optimization problems [19–21]. While in global VF metamodeling approaches, the scaling function is approximated using global metamodels, e.g., Qian et al. [22] proposed a Bayesian approach to integrate LF model and HF simulation values for engineering design. Xiong et al. [23] put forward a model scaling technique based on Bayesian–Gaussian process to integrate the information from both LF and HF models. Han et al. [24] put forward a gradient-enhanced Kriging to form a generalized corrected based method, which was tested on the design of airfoil. Zheng et al. [25] proposed a hybrid VF global metamodeling method, which a RBF base model and a Kriging linear correction were combined to make full use of LF and HF information. Tyan et al. [26] adopted RBF network as the scaling function to replace Taylor series, making the global VFM approach more efficient for high-dimensional design problems. Compared with the local VF metamodeling approach, the most obvious advantage of these global VF metamodeling approaches is that they are able to cope with multiple optimum situations sophisticatedly on the entire domain. Until now, more researches have been carried out to develop new types of LF model tuning that will further improve the accuracy and reduce the computational effort of VF metamodeling, but little attention has been paid to utilize the already-acquired information of difference characteristics between the HF and LF models. In other words, how to appropriately arrange and make full use of the sample points for HF models to run simulations according to the already-acquired difference information between the HF and LF models during the tuning process should be drawn more attention, especially when the computational cost is limited.

Instead of developing novel types of LF model tuning as in the past, this paper proposes an active learning VF metamodeling approach (AL-VFM), in which the one-shot VF metamodeling process is transformed into an active learning iterative process. The goal of the active learning process is to exploit the already-acquired information from the previous VF data to guide the VF metamodeling. The already-acquired information represents the location of regions where the differences between the HF and LF models are multi-model, non-smooth and have abrupt changes. The approximation performance of AL-VFM approach is demonstrated using some mathematical and engineering cases, and a rough comparison of AL-VFM approach and other metamodeling techniques are made. It is expected that more accurate metamodels can be developed with AL-VFM for the same number of simulation evaluations.

The rest of this paper is organized as follows. In Section 2, the background and several definitions used in this work are put forward. Details of the proposed approach are presented in Section 3. Numerical cases and comparison results are provided in Section 4. Two engineering examples are provided in Section 5 to demonstrate that the proposed VF modeling approach is applicable to complex problems. Conclusions and future work are discussed in Section 6.

2. Background and definitions

In this section, we provide the background and related definitions to the proposed approach, including: Kriging metamodeling, VF metamodeling, difference unstable region (DUR).

2.1. Kriging metamodeling

Kriging is an interpolative Bayesian metamodeling technique. It was originated from geo-statistical and used by Sacks et al. [27] for predicting the unknown response at sample points. Kriging treats the observed response as a combination of a global model and local deviations:

$$\hat{f}(\boldsymbol{x}) = p(\boldsymbol{x}) + Z(\boldsymbol{x}) \tag{1}$$

where $p(\boldsymbol{x})$ is a known polynomial function, $Z(\boldsymbol{x})$ is the realization of a stochastic process with mean zero and nonzero covariance. The nonzero covariance of $Z(\boldsymbol{x})$ is given by:

$$COV\left(Z(x_i), Z(x_j)\right) = \sigma^2 \boldsymbol{R}\left[R(x_i, x_j)\right] \tag{2}$$

where $\boldsymbol{R}$ is the correlation matrix. $\boldsymbol{R}(x_i, x_j)$ is the correlation function between two sample points x_i and x_j. When the Gaussian correlation function is employed, it can be calculated by:

$$\boldsymbol{R}(\theta) = \exp\left[-\sum_{k=1}^{K} \theta_k \left(x_i^k - x_j^k\right)^2\right] \tag{3}$$

where K denotes the dimensions of design space and θ_k are the unknown correlation parameters to be determined. Because Kriging is an interpolative Bayesian metamodeling, the model will have no mean square error (MSE) at all sample points. If the MSE is minimized, the predictor $\widehat{f}(x)$ for unobserved points is expressed as:

$$\hat{f}(x) = \hat{\beta} + \boldsymbol{r}^{\boldsymbol{T}}(x)\boldsymbol{R}^{-1}\left(f - \hat{\beta}\boldsymbol{p}\right) \tag{4}$$

where f is the column vector of length m that contains the sample data of the responses, and $\boldsymbol{p}$ is a column vector of length m that is filled with ones when $p(\boldsymbol{x})$ is taken as a constant. $r^T(x)$ is the correlation vector between an unobserved point x and the sample points.

$$\boldsymbol{r}^{\boldsymbol{T}}(x) = \left[R(x, x^1), R(x, x^2), \ldots, R(x, x^N)\right]^T \tag{5}$$

The scalar $\widehat{\beta}$ is estimated using the following equation:

$$\hat{\beta} = \left(\boldsymbol{p}^T \boldsymbol{R}^{-1} \boldsymbol{p}\right)^{-1} \boldsymbol{p}^T \boldsymbol{R}^{-1} \boldsymbol{f} \tag{6}$$

The estimated variance of the output model can be calculated by:

$$\hat{\sigma}^2 = \frac{(\boldsymbol{f} - \hat{\beta}\boldsymbol{p})^T \boldsymbol{R}^{-1} (\boldsymbol{f} - \hat{\beta}\boldsymbol{p})}{N} \tag{7}$$

The unknown correlation parameters θ_k are founded using maximum likelihood estimation can be formulated as [7]:

$$\begin{aligned} \max \quad & \Phi(\Theta) = -\frac{[N \ln(\hat{\sigma}^2) + \ln|\boldsymbol{R}|]}{2} \\ \text{s.t.} \quad & \Theta > 0 \end{aligned} \tag{8}$$

where Θ denotes the vector of θ_k, and both $\hat{\sigma}$ and $\boldsymbol{R}$ are the function of Θ.

2.2. VF metamodeling

The VF metamodeling technology is based on the assumption that, apart from a HF model that is sufficiently accurate but