Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2020-0195

Vol. **28**, No. 5, pp. 1937-1969
November 2020

High-Dimensional Nonlinear Multi-Fidelity Model with Gradient-Free Active Subspace Method

Bangde Liu[1] and Guang Lin[2,*]

[1] *School of Mechanical Engineering, Purdue University, 585 Purdue Mall, West Lafayette, IN 47906, USA.*
[2] *Department of Mathematics, School of Mechanical Engineering, Department of Statistics (Courtesy), Department of Earth, Atmospheric, and Planetary Sciences (Courtesy), Purdue University, West Lafayette, IN 47907, USA.*

Received 1 October 2020; Accepted (in revised version) 21 October 2020

Abstract. In scientific and engineering applications, often sufficient low-cost low-fidelity data is available while only a small fractional of high-fidelity data is accessible. The multi-fidelity model integrates a large set of low-cost but biased low-fidelity datasets with a small set of precise but high-cost high-fidelity data to make an accurate inference of quantities of interest. Under many circumstances, the number of model input dimensions is often high in real applications. To simplify the model, dimension reduction is often used. The gradient-free active subspace is employed in this research for dimension reduction. In this work, a novel predictive model for high-dimensional nonlinear problems by integrating the nonlinear multi-fidelity Gaussian progress regression and the gradient-free active subspace method is put forward. Numerical results demonstrated that the proposed approach can not only perform effective dimension reduction on the original data but also obtain accurate prediction results thanks to the effective dimension reduction procedure.

AMS subject classifications: 60G15, 60A99, 62A99

Key words: Gaussian process regression, dimension reduction, multi-fidelity model, active subspace, machine learning.

1 Introduction

The main idea of the multi-fidelity models [1–5] can be constructed by combining computational-expensive high accurate (high-fidelity) data with low-cost less accurate data (low-fidelity). Nowadays, most of the multi-fidelity approaches are set up on the Gaussian process (GP) [6] with the order-one autoregressive model proposed by [7].

*Corresponding author. *Email addresses:* `guanglin@purdue.edu` (G. Lin), `liu1947@purdue.edu` (B. Liu)

GP [8–15] is a suitable method for the multi-fidelity problems, because it has ability to use the prior belief to learn how different fidelities are related. However, the traditional autoregressive multi-fidelity models are only suitable when different fidelities relationship is linear. To cope with the nonlinear relation between fidelities, a method called nonlinear information fusion algorithm [16] that is based on the GP and nonlinear autoregressive scheme had been developed. This method proposes a more general form of the multi-fidelity model structure. Thanks to this method, it can deal with both the linear and nonlinear multi-fidelity problems effectively.

In scientific and engineering applications, people use computer models to study the input and output relations. In most cases, the dimension of the input space is usually very large that makes the models computationally expensive. Reducing dimensions can help otherwise infeasible parameter studies. To enable such studies, people usually use various methods to decrease the dimensions of the input space. There have many popular dimension reduction methods, such as principal component analysis (PCA) [17, 18], forward feature selection, backward feature elimination, and gradient-free active subspace approach [19]. Active subspace is a dimension reduction tool to identify the important directions of the input space. However, the classic active subspace method is based on the gradient information. In most cases, it is hard to get the gradient information. To avoid this shortcoming, a gradient-free active subspace method [19] is proposed.

In this work, a novel nonlinear multi-fidelity predictive model that is based on the nonlinear multi-fidelity scheme and gradient-free active subspace method is built for high-dimensional problems. In particular, the low-fidelity data can be used to calculate the BIC score to determine the active subspace dimensions and as the input of the gradient-free active subspace method to get the dimension reduction matrix. To improve the predictive accuracy, Bayesian active learning method [20] is employed to augment our original data based on the high-fidelity data to enhance accuracy. Bayesian active learning method indicates where a function will be evaluated next under a limited budget. In the proposed model, these new samples are explored according to where the largest variance of function is located. By using the largest variance as the sample location indicator, new samples are added using Bayesian active learning to augment our original data size. Then the dimension reduction matrix is employed to perform the dimension reduction on our new low- and high-fidelity data. Finally, the data after dimension reduction is used as the input of our nonlinear multi-fidelity scheme to build the nonlinear multi-fidelity predictive model.

Our contribution to this paper is two-fold:

1. A new multi-fidelity surrogate model is proposed that can perform gradient-free dimension reduction to the original data and make an accurate model prediction based on data after dimension reduction.

2. The proposed multi-fidelity surrogate model is implemented and a systematic comparison of the proposed model with the standard multi-fidelity model on nonlinear high-dimensional problems is investigated. The feasibility of our method is proved