Journal of Computational Physics 321 (2016) 259–278

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Inverse regression-based uncertainty quantification algorithms for high-dimensional models: Theory and practice

Weixuan Li [a], Guang Lin [b,*], Bing Li [c]

[a] Advanced Computing, Mathematics, and Data Division, Pacific Northwest National Laboratory, 902 Battelle Blvd., Richland, WA 99352, USA
[b] Department of Mathematics & School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA
[c] Department of Statistics, 410 Thomas Building, Penn State University, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form 8 May 2016
Accepted 19 May 2016
Available online 27 May 2016

Keywords:
Uncertainty quantification
High dimensionality
Sufficient dimension reduction
Sliced inverse regression
Polynomial chaos
Control variates

ABSTRACT

Many uncertainty quantification (UQ) approaches suffer from the *curse of dimensionality*, that is, their computational costs become intractable for problems involving a large number of uncertainty parameters. In these situations, the classic Monte Carlo often remains the preferred method of choice because its convergence rate $O(n^{-1/2})$, where n is the required number of model simulations, does not depend on the dimension of the problem. However, many high-dimensional UQ problems are intrinsically low-dimensional, because the variation of the quantity of interest (QoI) is often caused by only a few latent parameters varying within a low-dimensional subspace, known as the sufficient dimension reduction (SDR) subspace in the statistics literature. Motivated by this observation, we propose two inverse regression-based UQ algorithms (IRUQ) for high-dimensional problems. Both algorithms use inverse regression to convert the original high-dimensional problem to a low-dimensional one, which is then efficiently solved by building a response surface for the reduced model, for example via the polynomial chaos expansion. The first algorithm, which is for the situations where an exact SDR subspace exists, is proved to converge at rate $O(n^{-1})$, hence much faster than MC. The second algorithm, which doesn't require an exact SDR, employs the reduced model as a control variate to reduce the error of the MC estimate. The accuracy gain could still be significant, depending on how well the reduced model approximates the original high-dimensional one. IRUQ also provides several additional practical advantages: it is non-intrusive; it does not require computing the high-dimensional gradient of the QoI; and it reports an error bar so the user knows how reliable the result is.

1. Introduction

Numerical simulation is a powerful tool for studying physical processes and predicting quantities of researchers' interest. An accurate simulation relies on the knowledge of the parameter values that describe the characteristics of the system being simulated, such as material properties, initial and boundary conditions. However, in many applications, these parameter values are not exactly known to the modeler and thus are subject to uncertainties. In this situation, an uncertainty quantification (UQ) procedure may be implemented to examine how the prediction of the quantity of interest (QoI) is affected by the uncertain parameters.

* Corresponding author.
E-mail address: guanglin@purdue.edu (G. Lin).

http://dx.doi.org/10.1016/j.jcp.2016.05.040
0021-9991/

A straightforward approach to UQ is the Monte Carlo (MC) method, which works by generating a large sample of model simulations based on different possible realizations of the uncertain parameters. Then the statistical properties of the QoI such as mean, variance, and probability density function (PDF) may be estimated from the sample. By the Central Limit Theorem, the MC estimator is root n consistent, that is, the estimation error converges to 0 *in probability* at the order of $n^{-1/2}$, where n is the number of realizations generated [1, Chapter 2]. Since this convergence rate is relatively slow, a large number of model simulations are needed for an accurate estimation, which could be a huge computational burden and thus makes MC impractical except for simple, small-scale problems.

As an alternative to direct MC sampling via simulations, we can construct a response surface (also called a surrogate model, an emulator, or a proxy model in literature), which is an inexpensive approximate model for the relationship between the QoI and uncertain parameters. Examples of response surfaces for UQ include polynomial chaos (PC) expansion [2–4], Kriging interpolation or Gaussian process approximation [5–8]. Once a response surface is built, we can quickly generate realizations of the QoI without resorting to simulations on the original model and thus greatly reduce the computational cost. Moreover, some response surface models, for example the PC expansion, enable convenient analytic solutions of the statistical properties of the QoI.

However, most response surface-based UQ approaches are limited to low-dimensional problems, that is, problems involving only a small number of uncertain parameters, because the computational effort required to build a response surface grows fast as the dimension of the problem increases. For example, when a response surface is built using the full tensor product-based collocation scheme, the computational cost (measured by the number of model evaluations at the collocation points) is an exponential function with respect to the dimension of the problem. Although this growth rate may be reduced with techniques such as the Smolyak-type sparse grid [9–11] or the functional ANOVA decomposition [12,13], the response surface methods could still lose their efficiency advantage over MC, which, in contrast, has a convergence rate that does not depend on the problem's dimension. This difficulty is known as the *curse of dimensionality*. To mitigate this issue, a dimension reduction procedure is needed.

A commonly used dimension reduction method is principle component analysis (PCA) [14]. PCA reduces the dimension of the input of a UQ problem by making use of the correlations among different uncertain parameters. This correlation information is derived from the joint distribution of uncertain parameters, usually given as prior in a UQ problem. When the parameters are highly correlated, PCA is capable of representing the variations of all the uncertain parameters by a much smaller number of random variables, which are linear functions of the original uncertain parameters, called principal components. However, PCA results in no or little dimension reduction when the parameters are uncorrelated or only weakly correlated.

Moreover, PCA does not make use of the knowledge of how the QoI depends on the uncertain parameters, which can be immensely valuable for further reducing the dimension. For instance, different uncertain parameters in a model often have different levels of influence on the QoI. If we are able to identify the important parameters, say using sensitivity analysis, an anisotropic response surface may be built accordingly [15,13,16]. For example, a PC expansion may include more PC terms with respect to the important parameters and exclude the high-order PC terms with respect to the not-so-important parameters. Such a strategy allows for a high-dimensional PC expansion using relatively small number of terms without seriously compromising accuracy. However, sensitivity-analysis-based dimension reduction techniques cannot handle problems where different uncertain parameters have roughly equal impacts on the QoI.

Under these circumstances, one may consider sensitivity analysis in a more general sense, that is, to study the sensitivity of the QoI with respect to linear combinations of the uncertain parameters. Specifically, we rotate the coordinate system of the parameter space such that the QoI varies mainly along the directions of a few synthetic coordinates (linear combinations of original parameters). Then a low-dimensional response surface may be built with this small set of coordinates. Examples of such techniques include active subspace [17,18], basis adaptation [19], and compressive sensing [20]. By allowing coordinate rotation, the generalized sensitivity analysis gives us extra freedom to drop more unimportant directions and thereby achieves further dimension reduction. Finding such generalized sensitivity information is a nontrivial task as it requires more knowledge about the model. For example, if the gradient of the model (i.e., the partial derivatives of the QoI with respect to the uncertain parameters) can be evaluated at a sample of parameter points, then the important directions are extracted by the average outer product of the model gradient [18]. However, evaluating the gradient at the sample points, especially for high-dimensional models, could be even more expensive than evaluating the QoI itself, which is what the MC approach requires for the UQ problem.

The idea of rotating the coordinates for building low-dimensional response surfaces is closely related to the notion of linear sufficient dimension reduction (SDR), which is based on the sufficiency principle and was first introduced in the context of multivariate regression [21]. Various approaches for searching SDRs have been developed [22–29], among which the most widely used is sliced inverse regression (SIR) [22]. This approach estimates SDR subspace by regressing the input parameters against the output QoI, hence the nomenclature inverse regression. During the past two decades, SIR achieved its popularity due to its theoretical soundness, implementational ease and wide applicability.

In this paper, we investigate an inverse-regression-based uncertainty quantification (IRUQ) approach for high-dimensional problems. Specifically, we use SIR to seek a low-dimensional SDR subspace for the QoI under study. After the dimension reduction, building an accurate response surface over the SDR subspace becomes computationally viable. This approach is non-intrusive, which means that its implementation requires only a simulator that evaluates the QoI with specified parameters. Neither the gradient information nor any other manipulation of the model's internal structure is needed. Since