Advances in Engineering Software 000 (2017) 1–12

Contents lists available at ScienceDirect

Advances in Engineering Software

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

Research paper

Quasi-sparse response surface constructing accurately and robustly for efficient simulation based optimization

Pu Li [a,b], Haiyan Li [a,*], Yunbao Huang [a,*], Kefeng Wang [a], Nan Xia [a]

[a] School of Mechanical and Electrical Engineering, Guangdong University of Technology, Guangzhou, Guangdong, China
[b] School of Physics and Electrical Engineering, Shaoguang University, Shaoguan, Guangdong, China

ARTICLE INFO

Article history:
Received 22 May 2017
Revised 6 July 2017
Accepted 30 July 2017
Available online xxx

Keywords:
Simulation Optimization
Response surface
Quasi-sparse
Uniform sampling

ABSTRACT

Response surface method is often employed in simulation based design and optimization for complex products. The sparsity of response surface on the mathematic basis has been explored to accurately represent the variation between design variables and performance response with only a few design points, which is very beneficial to efficient design optimization. Due to the selected basis, it may lead to a large deviation, or under-fitting of the reconstructed response surface since the number of sampling points is often smaller than its sparseness.

In this paper, a quasi-sparse response surface is presented to trade-off the sparsity and variation of response surface by introducing coefficient shrinkage regularization and uniformly sampling for the design points, which enables more atoms in the basis included to accurately and robustly reconstruct the surface. The group of basis atoms which are correlated with sampling points instead of the most correlated one are all selected to uniform express the sampling points, and the coefficients of basis atoms are shrunk to improve the prediction performance of the model.

Finally, 9 benchmark functions and 1 engineering applications are utilized to demonstrate the significance of the presented approach by comparing with other normally used response surface models, The results shows that the accuracy and robustness of the reconstructed response surface is superior than those of other response surface approaches.

1. Introduction

In the design process of modern electromechanical products and other complicated systems, modeling and simulation technology has been widely employed in simulation analysis and optimization of decision-making to improve the overall performance of the products [1]. In the simulation process of the product, the simulation model is often multidisciplinary, non-linear and with many other significant features, leading to a long time of simulation solution, which may take as many as 30–160 h [2]. The response surface simulation optimization method based on computer experiment design is an effective method [3] to reduce the time of simulation.

In order to meet different application requirements, the following response surface models are mainly used at present [2]: (1) Polynomial Response Surface (PRS), (2)Multivariate Adaptive Regression Splines (MARS), (3) Kriging (Ordinary Kriging (OK) and Blind Kring (BK)), (4) Radial Basis Functions (RBF) and Extended Radial Basis Functions (ERBF), (5)Support Vector Regression (SVR), (6) Sparsity-promoting Polynomial Response Surface (SPPRS). Among them, PRS [4] is a commonly used response surface model, however, the maximum number of polynomials needs to be set, which often results in the problem of under-fitting or over-fitting. MARS is used to express complex response surfaces by using multi-stage low-order polynomials, and it can process the multidimensional response surface by tensor product form [5], but the number of discrete segments and the parametric method are still to be further studied. Local interpolation term on the basis of PRS is added in Kriging, which improved the under-fitting problem of PRS, but it is high in computational complexity and difficult to be adapted to global optimization process [6]. RBF, especially multi-quadratic RBF is easy to be constructed and can achieve high approximation accuracy [7]. SVR can be used to establish the response surface of the structural analysis model and achieves good results [8]. Sparseness plays a big role in the construction of response surface. Sparsity and regularization strategy has been employed by Huong et al [9], Liu and Wang [10–13], Huang [14] to construct the surface from point-cloud. The same strategy is ap-

* Corresponding authors.
E-mail addresses: cathyisme@163.com (H. Li), Huangyblhy@gmail.com (Y. Huang).

http://dx.doi.org/10.1016/j.advengsoft.2017.07.014
0965-9978/© 2017 Published by Elsevier Ltd.

plied in SPPRS [15] provided by Fan from our team in 2014 which is a new high-fidelity surrogate modeling approach based 'sparsity-promoting" regression method. The "sparsity-promoting" regression method is proposed exactly aiming to seek such a sparse representation with limited sampling points. Numerical experiments show its performance exceeds surrogate model technologies [15] mentioned above, however, its performance is limited in the case of the number of sampling points is smaller than the sparseness of response surface on the mathematic basis.

We define sparseness s as the minimize number of basis atom in selected basis to reconstruct the source model. Because of the complexity, many of the models are not sparse enough in specific basis functions, which means $s > n$, where n is the number of sampling points. In this case, the s sparse basis atoms cannot be selected completely and accurately from basis dictionary by response surface pursuing sparsity. In order to be able to accommodate the case of $s > n$, we propose a quasi-sparse response surface (QSRS) model trades off the sparsity and variation of response surface. The group of basis atoms which are correlated with sampling points instead of the most correlated one are all selected to uniform express the sampling points. The coefficients of basis atoms are shrunk to improve the prediction performance of the model [16]. Elastic net regression is employed to achieve above purposes.

The rest of the paper is organized as follow: in Section 2, we present the QSRS, focusing on elastic net regression and solution algorithm; Section 3 gives the simulation test condition; after presenting and discussing the results of tests in Section 4, the conclusion is given.

2. Quasi-sparse response surface

In this section, we first propose the model structure of QSRS, and then introduce the elastic net regression and uniform design sampling method; finally, we give the algorithm process of QSRS.

2.1. Model structure

QSRS can be written as a linear model

$$\hat{\boldsymbol{y}}(\boldsymbol{x}) = \sum_{i=1}^{p} \beta_i \varphi_i(\boldsymbol{x}), \tag{1}$$

where m is the number of variables, $\boldsymbol{x} = [x_1 \cdots x_m]^T$ is a design point, $\{\beta_i\}_{i=1,\,2,\,\cdots p}$ are coefficients, $\{\varphi_i(\boldsymbol{x})\}_{i=1,\,2,\,\cdots p}$ are atoms from a dictionary, p is the number of atoms. Dictionary is a set of basis functions (so-called atoms) such as Legendre polynomial functions. Atom is column vector of the basis functions values in the design point. We use Legendre polynomials as the atoms of dictionary in this paper, and $\varphi_i(\boldsymbol{x})$ can be defined as

$$\varphi_i(\boldsymbol{x}) = L\left(\boldsymbol{x}, \boldsymbol{\eta}^{(i)}\right) = \prod_{j=1}^{m} l_j\left(x_j, \eta_j^{(i)}\right),\ \ i = 1, \cdots, p$$

where $\boldsymbol{\eta}^{(i)} = [\eta_1^{(i)} \cdots \eta_m^{(i)}]^T$ is the exponent vector of $\varphi_i(\boldsymbol{x})$, $L(\boldsymbol{x}, \boldsymbol{\eta}^{(i)})$ is a form of Legendre polynomials, and $l_j(x_j, \eta_j^{(i)})$ is the $\eta_j^{(i)}$-order univariate Legendre polynomial with respect to x_j.

Given a set of sampling points $\boldsymbol{x} = [x^{(1)}, \cdots, x^{(n)}]^T$, $x^{(k)} \in \mathbb{R}^m$, $k = 1, 2, \cdots, n$, and the corresponding actual response $\boldsymbol{y} = [y^{(1)}, \cdots, y^{(n)}]^T$, then the so-called "design matrix" can be defined as

$$\boldsymbol{\Phi} = \begin{bmatrix} \varphi_1\left(x^{(1)}\right) & \cdots & \varphi_p\left(x^{(1)}\right) \\ \vdots & \ddots & \vdots \\ \varphi_1\left(x^{(n)}\right) & \cdots & \varphi_p\left(x^{(n)}\right) \end{bmatrix}.$$

The matrix form of Eq. (1) can be written as

$\hat{\boldsymbol{y}}(\boldsymbol{x}) = \boldsymbol{\Phi}\boldsymbol{\beta}$.

The work QSRS fitting the source model is to solve the problem

$$\boldsymbol{\beta} = arg\min_{\boldsymbol{\beta}} \left\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\beta}\right\|_2 \tag{2}$$

If we want to estimate a sparsity response surface, we should add the constraint to Eq. (2):

$s.t.\ \left\|\boldsymbol{\beta}\right\|_0 = s.$

where s is the sparseness of source model in Legendre polynomial basis functions, and $\|\boldsymbol{\beta}\|_0$ means the nonzero number of coefficients $\boldsymbol{\beta}$. Solving Eq. (2) with the constraint ($\ell 0$ norm) is a NP hard problem, especially s is a unknown variable. The general method is relaxing the sparse constraint condition to $\ell 1$ norm or $\ell 2$ norm.

A good model should be good at both prediction and interpretation. It is achieved by variable selection and coefficient shrinkage in statistical regression field which is just the same as the linear model in the response surface. There are a lot of regression methods to approach the purpose such as elastic net regression, least absolute shrinkage and selection operator (LASSO) and ridge regression. LASSO method, which can provide both variable selection and coefficient shrinkage functions, can be used to generate a sparse model, however, the model performance is limited in the number of atoms selected in the case of $s > n$. As a continuous shrinkage method, ridge regression achieves its better prediction performance through a bias-variance trade-off [17]. However, the model ridge regression produced is not sparsity, for ridge regression always keeps all variables in the model. Elastic net regression combines the characteristic of both LASSO and ridge regression. We use elastic net regression to construct a quasi-sparsity response surface model in this paper.

2.2. Elastic net regression

2.2.1. Definition

Consider the cost function

$$\mathcal{L}(\lambda_1, \lambda_2, \beta) = \left\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\beta}\right\|^2 + \lambda_1\left\|\boldsymbol{\beta}\right\|_1 + \lambda_2\left\|\boldsymbol{\beta}\right\|_2^2, \tag{3}$$

where

$$\|\beta\|_1 = \sum_{j=1}^{p} \left|\beta_j\right|,$$

$$\|\beta\|_2^2 = \sum_{j=1}^{p} \beta_j^2.$$

The Elastic net estimator $\hat{\boldsymbol{\beta}}$is the minimize of Eq. (3):

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{argmin}\{\mathcal{L}(\lambda_1, \lambda_2, \beta)\}. \tag{4}$$

Let $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, we can rewrite Eq. (3) as

$$\mathcal{L}(\lambda, \alpha, \beta) = \|y - \Phi\beta\|^2 + \lambda\left((1-\alpha)\|\beta\|_1 + \alpha\|\beta\|_2^2\right). \tag{5}$$

The parameter α determines the mix of the penalties, and is often pre-chosen on qualitative grounds. It encourages highly correlated features to be averaged, while the first regularization parameter $(1 - \alpha)$ encourages a sparse solution in the coefficients of these averaged features.

The elastic net penalty function $\lambda((1-\alpha)\|\beta\|_1 + \alpha\|\beta\|_2^2)$ is a convex combination of the LASSO and ridge penalty. When $\alpha = 1$, it becomes simple ridge regression. For all $\alpha \in [0, 1)$, due to the lack of first derivative, the elastic net penalty function is singular at 0 and it is strictly convex for all $\alpha > 0$, thus having the characteristics of both the LASSO and ridge regression. Especially when $\alpha = 0$, it becomes LASSO regression and it is convex but not strictly convex.