The Seventh Asian-Pacific Symposium on Structural Reliability and Its Applications (APSSRA2020)
October 4–7 2020, Tokyo, Japan
T. Takada, I. Yoshida & T. Itoi (editors)

A Sparse Polynomial Surrogate Model for Geotechnical Reliability Design

T. Shuku[1], S. Nishimura[2] and T. Shibata[3]

[1]*Okayama University. Email: shuku@cc.okayama-u.ac.jp*
[2]*Okayama University. Email: theg1786@cc.okayama-u.ac.jp*
[3]*Okayama University. Email: tshibata@cc.okayama-u.ac.jp*

Abstract: This paper presents a method for building surrogate model for geotechnical reliability analysis based on sparse modeling, sparse surrogate model. Sparse modeling, which is called least absolute shrinkage statistical operator (lass) in statistics, has the property that some of the parameters in a surrogate model are driven to zero and leads to a simpler model. Building a surrogate model can be divided into two processes, model selection and parameter estimation, and the sparse modeling enables to achieve these two processes at the same time. A polynomial surrogate model was designed to estimate consolidation settlement based on sparse estimation, and its applicability has been investigated by comparing the results by the surrogate model with those by finite element analysis.

Keywords: surrogate model, sparse modeling, geotechnical reliability analysis

1. Introduction

Surrogate model, also called "response surface" or "meta model", is a regression equation that approximates relationships between input and output data of numerical simulations and have been commonly used for parameter identification and reliability-based analysis in many research fields. Applications of surrogate model in civil engineering include Bucher and Bourgund (1990), Tandjiria et al. (2000), Youssef and Soubra (2008), Schoefs et al. (2013), and Zhang et al. (2015).

In surrogate modeling, model selection and parameter estimation need to be dealt with. Model selection is the problem in which how to set basis functions and model complexity (polynomial order) is discussed. Whereas parameter estimation is the problem to determine coefficients of the basis functions. In order to build the "best" surrogate model, all possible combinations of basis functions and their coefficients should be analyzed. This problem, however, is difficult to solve because the computation time to find a solution grows exponentially with problem size and is known as "NP-hard" problem. The methodology to efficiently achieve model selection and coefficient estimation in surrogate modeling.

This study proposes a method for building surrogate models based on sparse modeling or least absolute shrinkage statistical operator (lasso, Tibshirani 1996). The proposed method can solve the both of model selection and parameter estimation problems at the same time. The applicability of the proposed method is investigated through numerical examples of geotechnical reliability analysis. In addition, we compare the results with those by existing method for comparison.

2. Surrogate Modeling based on Lasso

2.1 Basic Model

We use M^{th} order polynomial functions as surrogate models. When the input parameter is x, the polynomial function f is defined by:

$$f = w_0 + w_1 x + w_2 x^2 + \ldots + w_M x^M = \sum_{j=0}^{M} w_j x^j \qquad (1)$$

where, $w_0,\ldots,\ w_M$ are polynomial coefficients. The coefficients are determined by fitting the polynomial function to the training data y_n. This fitting is usually done by minimizing the least squares objective function:

$$J(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{f(x,\mathbf{w}) - y_n\}^2 \qquad (2)$$

where N is the number of training data. If the squared error follows Gaussian distribution and the model function is linear for input parameter, the analytical solution is available.

2.2 Regularization

There remains the problem of choosing the order M of the polynomial, and this is a "model selection" problem. Although higher order polynomials generally give good fits to the data, bias of the model tend to be extremely large. This problem is called "over-fitting" in the context of machine learning.

There is a technique called regularization to control over-fitting, and that is adding a penalty term to the objective functions to discourage the coefficients from reaching large values. The general expression of the modified objective functions including regularization term takes the form

$$\tilde{J}(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{f(x,\mathbf{w}) - y_n\}^2 + \frac{\lambda}{2}\sum_{j=1}^{M}|\ w_j\ |^q \qquad (3)$$

where λ is the regularization parameter which controls relative importance between first and second terms in Eq. (3), q is the parameter controls the regularization term, and q = 2 corresponds to the quadratic regularizer, so-called Ridge regression defined by

$$\tilde{J}_R(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{f(x,\mathbf{w}) - y_n\}^2 + \frac{\lambda}{2}\|\ \mathbf{w}\ \|^2 \qquad (4)$$

where $\|\mathbf{w}\|^2 = \mathbf{w}^{\mathrm{T}}\mathbf{w} = w_0{}^2 + w_1{}^2 + \ldots + w_M{}^2$.

The case of q = 1 is called least absolute shrinkage statistical operator (lasso, Tibshirani 1996), and it takes the form.

$$\tilde{J}_L(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{f(x,\mathbf{w}) - y_n\}^2 + \frac{\lambda}{2}|\ \mathbf{w}\ | \qquad (5)$$

where $|\mathbf{w}| = |w_0|+|w_1|+\ldots+|w_M|$. It has the property that if λ is sufficiently large, some of the coefficients w_j are driven to zero because of the geometry of its regularization term. Figure 1 illustrates the estimation graph of ridge regression and the lasso, and x_1 becomes zero because of the diamond-shaped regularization term. The lasso tends to lead to a sparse model in which the corresponding basis functions play no role. Estimating for surrogate models via the lasso is called "sparse estimation" in this paper.

2.3 Optimization Algorithm

The lasso problem is a convex minimization problem, a quadratic program with a convex constraint. For simplicity, the following problem is used to explain the computational procedure for the lasso solution.

$$\underset{w}{\text{minimize}}\left\{\frac{1}{2}(w-y_n)^2+\frac{\lambda}{2}|w|\right\} \tag{6}$$

The standard approach to this one-dimensional minimization problem is to take the gradient with respect to w and to set it to zero. However, one of the central difficulties in solving Eq. (6) is the presence of a non-smooth L1 norm, $|w|$. In other words, the absolute value function $|w|$ does not have a derivative at $w = 0$. Nevertheless, this problem can be solved by applying a soft-thresholding operator to w, which is defined as

$$S_\lambda(w)=\begin{cases} w-\lambda & (w>\lambda) \\ 0 & (-\lambda\le w\le\lambda) \\ w+\lambda & (w<-\lambda) \end{cases} \tag{7}$$

where S_λ is a soft-thresholding function (Figure 2). This operator translates w toward zero by an amount λ and sets it to zero if $|w| < \lambda$. When $\lambda = 0$, the solution of Eq. (5) becomes the solution for the ordinary least squares problem. The general approach for solving the lasso problem can be summarized as follows:

Step 1: Minimize first term in the objective function

Step 2: Apply the soft-thresholding operator to w

Step 3: Repeat Steps 1 and 2

To minimize the lasso-type objective function, we used Alternative Direction Method of Multipliers (ADMM, Boyd et al., 2010).

3. Numerical Example

3.1 Setup

We built a surrogate model to estimate a value of ground surface settlement due to embankment loading. This section presents the setup of the numerical example.

Figure 3 (a) and (b) show the model ground discretized with finite element mesh and the construction process of the embankment. The model ground is assumed to consist of three layers (sand layer, clay layer, and sandy clay), and the layers were modeled as an linear elastic model and Cam-clay models. An embankment is assumed to be constructed on the model ground following the construction process shown in Figure 3(b), and time-settlement behavior of the ground is observed at five points #1 ~ #5.

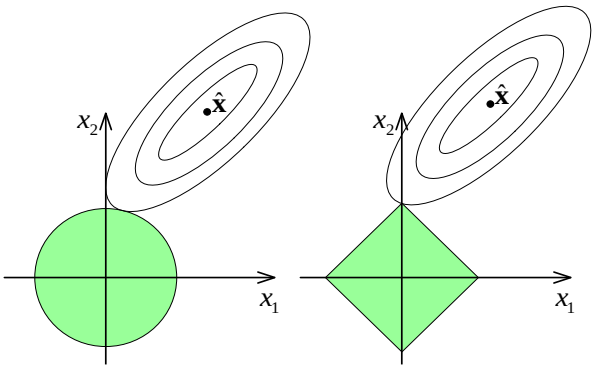


Figure 1. Estimation picture for ridge (left) and lasso (right) regression (modified from Hastie et al. 2015).

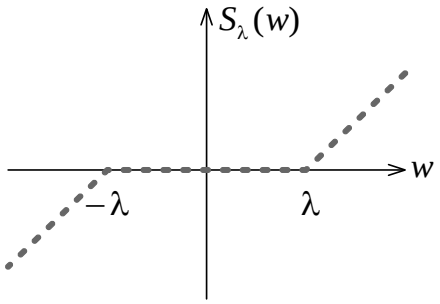


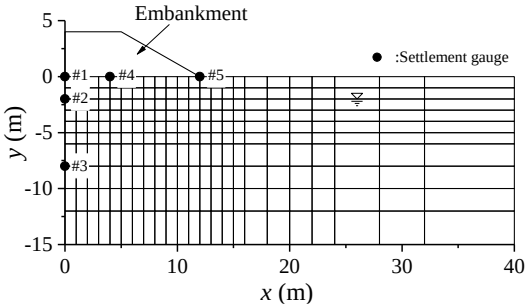
Figure 2. Soft-thresholding function.

The surrogate model was designed to estimate the settlement value at #1 after 2,500 days after construction began, and this settlement value is the output (or the objective variable) in the surrogate models. We assumed that ten parameters, elastic modulus E and the Poisson's ratio ν of the sand layer, and the compression index C_c, the swelling index κ, the critical state parameter M, and the coefficient of permeability k (m/d) of the clay and sandy clay layers, as the input parameters. The total number of input parameters is ten.

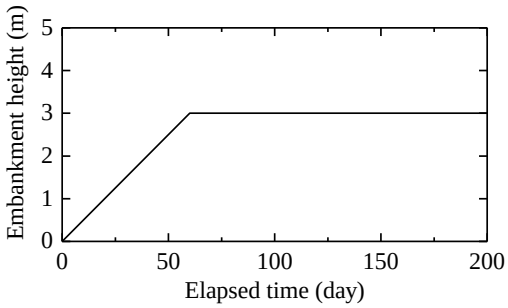
The performance of the surrogate models depends on the value of regularization parameter λ, and we determined the parameter using leave-one-out cross-validation which is commonly and widely used in many research fields. In this study, we built two surrogate models 1) N = 1,000 and 2) N = 50 to investigate the effect of the number of training data on building surrogate models. The performance of the surrogate model was evaluated by comparing the estimated probability density function of the target settlement value by the surrogate model with the true value, i.e., the PDF by finite element analysis.

3.2 Case 1: N=1,000

N = 1,000 was used to build the surrogate model, and the target settlement values were estimated by the lasso-based model and ridge-based model. Figure 4 compares the estimated PDF with the true PDF, and



(a) Finite element mesh.



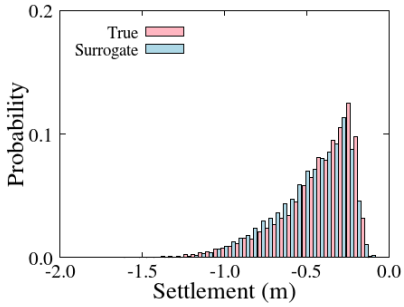
(b) Embankment loading process.

Figure 3. Setup of numerical simulation.

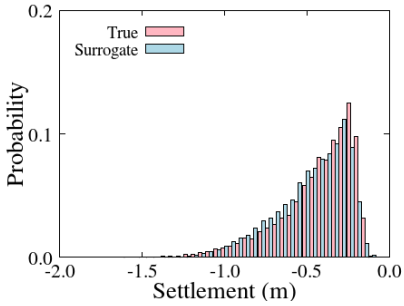
Table 1 summarizes the number of active set and Kullback–Leibler (KL) divergence. The active set means the number of non-zero entry in **x**, and that number is lower, the simpler surrogate model is built. The KL-divergence is a measure of how one probability distribution is different from a reference probability distribution, and we can quantitatively evaluate the performance of the surrogate models. The PDF estimated by two methods, lasso and ridge, are very similar and agree well with the true PDF. The KL-divergence of lasso is a bit smaller than that of ridge, and lasso-based model is more accurate than ridge-based model.

3.3 Case 2: N=100

Only 50 data were used to build the surrogate model in Case 2, and this problem is a typical "underdetermined problem" because the number of unknowns is greater than that of observation data. Figure 5 compares the estimated PDF with the true PDF for ridge and lasso, and Table 2 summarizes the results. The KL-divergence shows that the estimation accuracy of ridge-based model is lower than that of lasso-based model, and the shape of the PDF by ridge is a bit different from the true PDF. The number of active sets in lasso-based model is 34, and most of the coefficients, 32 input parameters, led to "zero". Figure 6(a)(b) shows the solution path of ridge-based and lasso-based models. The vertical lines indicate the best regularization parameter λ determined by the LOOCV.



(a) Ridge regression

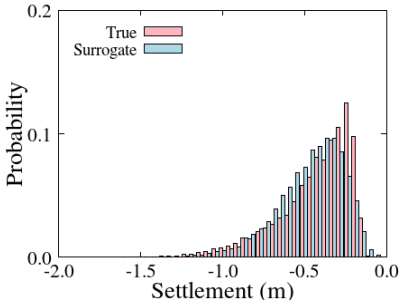


(b) Lasso

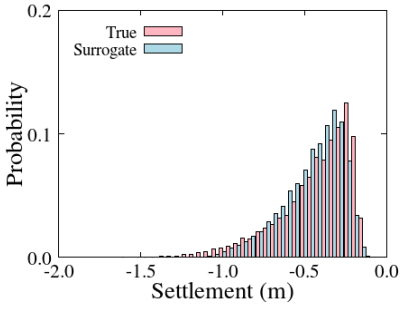
Figure 4. Comparison of PDF (N = 1,000).

Table 1. Summary of Case 1.

	Ridge	Lasso
The number of active sets	96	80
KL-divergence	0.02672	0.02174



(a) Ridge regression

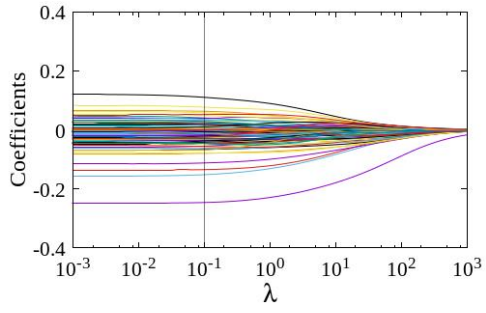


(b) lasso

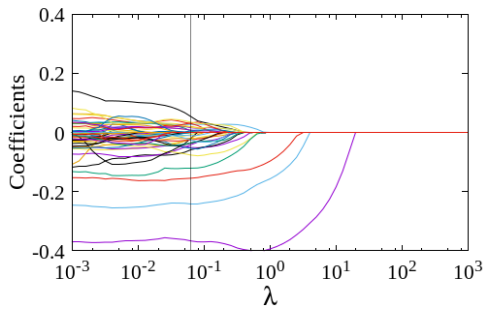
Figure 5. Setup of numerical simulation.

Table 2. Summary of Case 2.

	Ridge	Lasso
The number of active sets	96	34
KL-divergence	0.06578	0.03010



(a) Ridge regression



(b) Sparse estimation (lasso)

Figure 6. Solution path.

In ridge regression, regularization parameter is less sensitive to the shrinkage of the coefficients. In lasso, however, the larger λ is used, the simpler model is estimated. These results demonstrate that the proposed lasso-based method for building surrogate models estimate simpler/less complex models and provide more accurate estimations compared to the existing method.

4. Conclusions

A method for building surrogate models based on lasso was newly proposed. The surrogate model was designed to estimate a value of surface settlement of the ground using the data of the finite element simulations, and the model accuracy was evaluated by comparing the estimated PDF of the settlement value by the surrogate model with the true value. The estimated PDF was agreed well with the true PDF. The proposed method leads to simpler models compared to the existing method, ridge regression, and the lasso-based model can accurately estimate the PDF with small training data.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 18K05880.

References

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. 2010. Distributed optimization and statistical learning via alternating direction method of multipliers, *Foundation Trends Machine Learning*, 3(1), 122p.

Bucher, C. G. and Bourgund, U. 1990. A fast and efficient response approach for structural reliability problems, *Structural Safety*, 7, 57–66.

Schoefs, F., Le, K. T. and Lanata, F. 2013. Surface response meta-models for the assessment of embankment frictional angle stochastic properties from monitoring data: An efficient application to harbor structures, *Computers and Geotechnics*, 53, 122–132.

Tandjiria, V., The, C. I. and Low, B. K. 2000. Reliability analysis of laterally loaded piles using response surface methods, Structural Safety, 22, 335-355.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society Series B*, 58(1), 267–288.

Youssef Abdel Massih, D. S. and Abdul-Hamid Soubra, M. 2008. Reliability-based analysis of strip footing using response surface methodology, *International Journal of Geomechanics*, 8(2), 134–143.

Zhang, J., Chen, H. Z., Huang, H.W. and Lou, Z. 2015. Efficient response surface method for practical geotechnical reliability analysis, *Computers and Geotechnics*, 69, 496–505.