IOP Conference Series: Materials Science and Engineering

PAPER • **OPEN ACCESS**

Sparse modeling for the geotechnical observation data

To cite this article: T Kurita 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **652** 012023

IOP Conf. Series: Materials Science and Engineering **652** (2019) 012023 doi:10.1088/1757-899X/652/1/012023

Sparse modeling for the geotechnical observation data

T Kurita

Tokyo Electric Power Services Co., Ltd., Civil Engineering Department, Tokyo, Japan
Email: kurita@tepsco.co.jp

Abstract. In the geotechnical engineering, it is a significant issue to construct the numerical models for simulation analysis based on the observation data. In this study, sparse modeling used as an effective modeling method in the field of machine learning and image processing was applied to the geotechnical engineering and its effectiveness was examined. Aa a case study based on the field observation data, the fused lasso which is a typical method of the sparse modeling was applied to model the velocity structure of the subsurface ground using the PS logging data. As a result of examination, it was found that simplified models can be obtained by increasing the value of the regularization parameter.

1. Introduction

In order to simulate the seismic ground motions, it is necessary to prepare the velocity structure model of the subsurface ground. Since there is arbitrariness in the manually preparation of analytical models from bowling data, more objectively modeling method is required.

Recently, sparse estimation has been utilised as an effective modeling method in the field of machine learning [1] and image processing [2, 3]. In this study, sparse modeling was applied to the geotechnical engineering and examined its effectiveness. Fused lasso which is a representative method of the sparse modeling was applied to model the velocity structure of the ground using PS logging data.

2. Theory

The fused lasso which is a kind of generalized lasso proposed by Tibshirani and Taylor (2011) [4] was referred. The theory of the generalized lasso and the fused lasso is described below.

2.1. The generalized lasso

In this study, following linear regression equation is considered as a basic model.

$$y_i = \beta_1 x_{1,i} + \beta_2 x_{2,i} + \cdots + \beta_m x_{m,i} \qquad (i = 1, 2, \cdots, n) \tag{1}$$

In Eq.(1), y_i is an induced variable, $x_{m,i}$ is an explanatory variable, and β_m is a regression coefficient.

Precede the formulation of generalised lasso, the mathematical notations are defined as follows. $\|\cdot\|_1$ is the L_1 norm. And, $\|\cdot\|_2$ is the L_2 norm. According to Tibshirani and Taylor (2011) [4], the generalized lasso is amounted to the following optimization problem.

$$\hat{\boldsymbol{\beta}} \in \arg\min_{\boldsymbol{\beta}} \frac{1}{2}\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\mathbf{D}\boldsymbol{\beta}\|_1 \tag{2}$$

Where, $\hat{\boldsymbol{\beta}}$ is a generalized lasso solution vector, $\mathbf{y}$ is a vector of induced variable, $\mathbf{X}$ is a matrix of explanatory variables, $\mathbf{D}$ is a penalty matrix, and $\lambda \geq 0$ is a regularization parameter. The solution

IOP Conf. Series: Materials Science and Engineering **652** (2019) 012023 doi:10.1088/1757-899X/652/1/012023

can be obtained by solving the equivalent Lagrange dual problem. The relationship between the primal and dual solution is as follows.

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\left(\mathbf{X}^T\mathbf{y} - \mathbf{D}^T\hat{\mathbf{u}}\right) \tag{3}$$

Therefore, solving for $\hat{\mathbf{u}}$ in Eq.(3) brings the solution $\hat{\boldsymbol{\beta}}$ in Eq.(2).

2.2. The fused lasso

As a simple example of the generalized lasso of Eq.(2), the 1D fused lasso is formulated as follows.

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \frac{1}{2}\sum_{i=1}^{n}\left(y_i - \beta_i\right)^2 + \lambda\sum_{i=1}^{n-1}\left|\beta_{i+1} - \beta_i\right| \tag{4}$$

Here, Eq.(4) corresponds to setting $\mathbf{X}$ and $\mathbf{D}$ in Eq.(2) to be as follows.

$$\mathbf{X} = \mathbf{I} \tag{5}$$

and

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \tag{6}$$

Moreover, y_i indicates an observation data.

3. Results

The sparse modelling was applied to model the ground velocity structure by using the PS logging results of the subsurface ground obtained in the heavily damaged area of the 2016 Kumamoto earthquake.

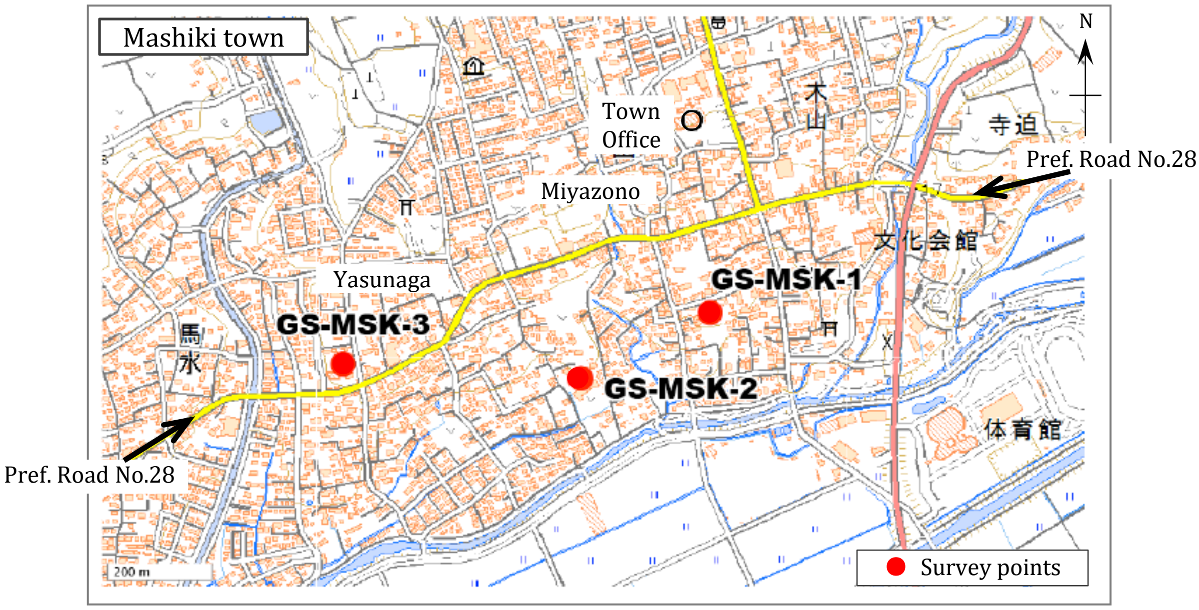

Figure 1. Borehole survey sites in Mashiki town (adapted from Yoshimi *et al.*, 2017 [5])

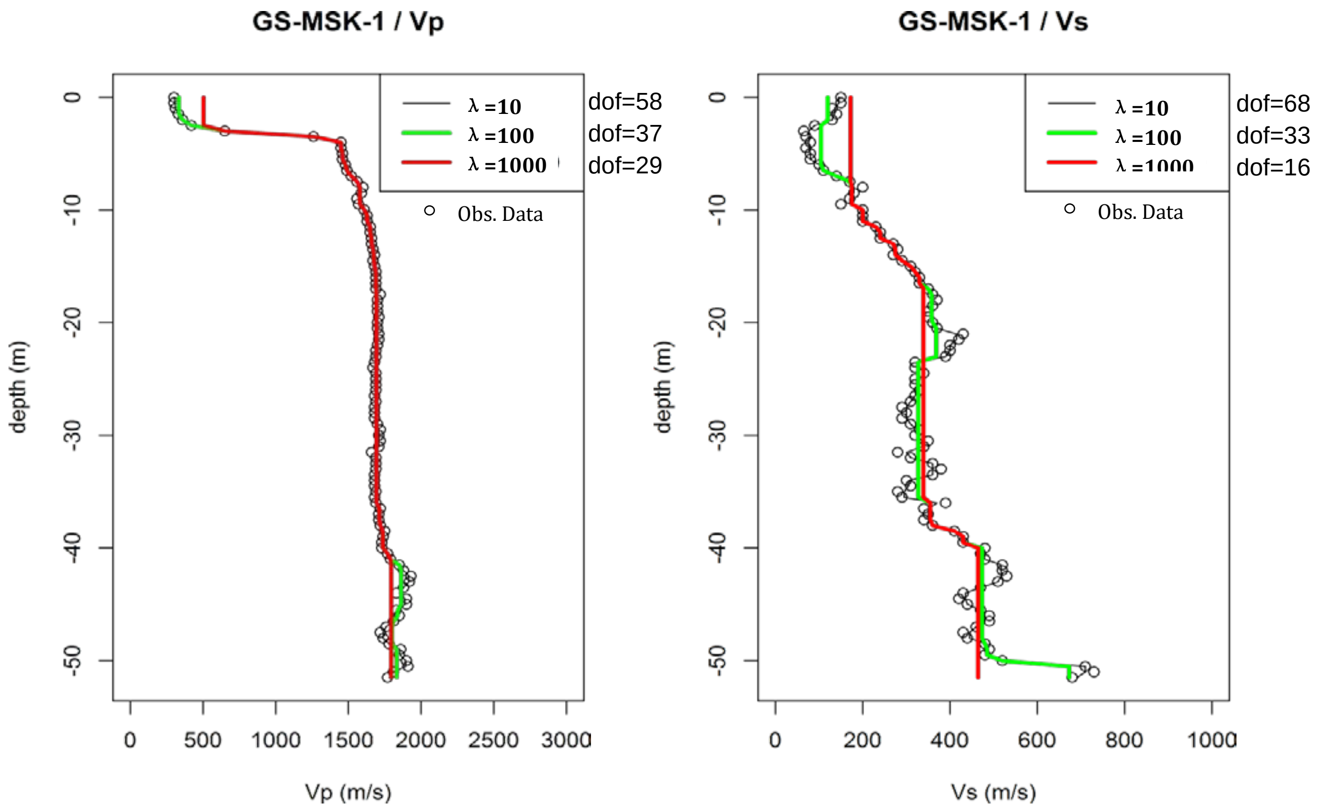

Figure 2. Modeling of the suspension PS logging results at GS-MSK-1 (Number of observed data: 104)

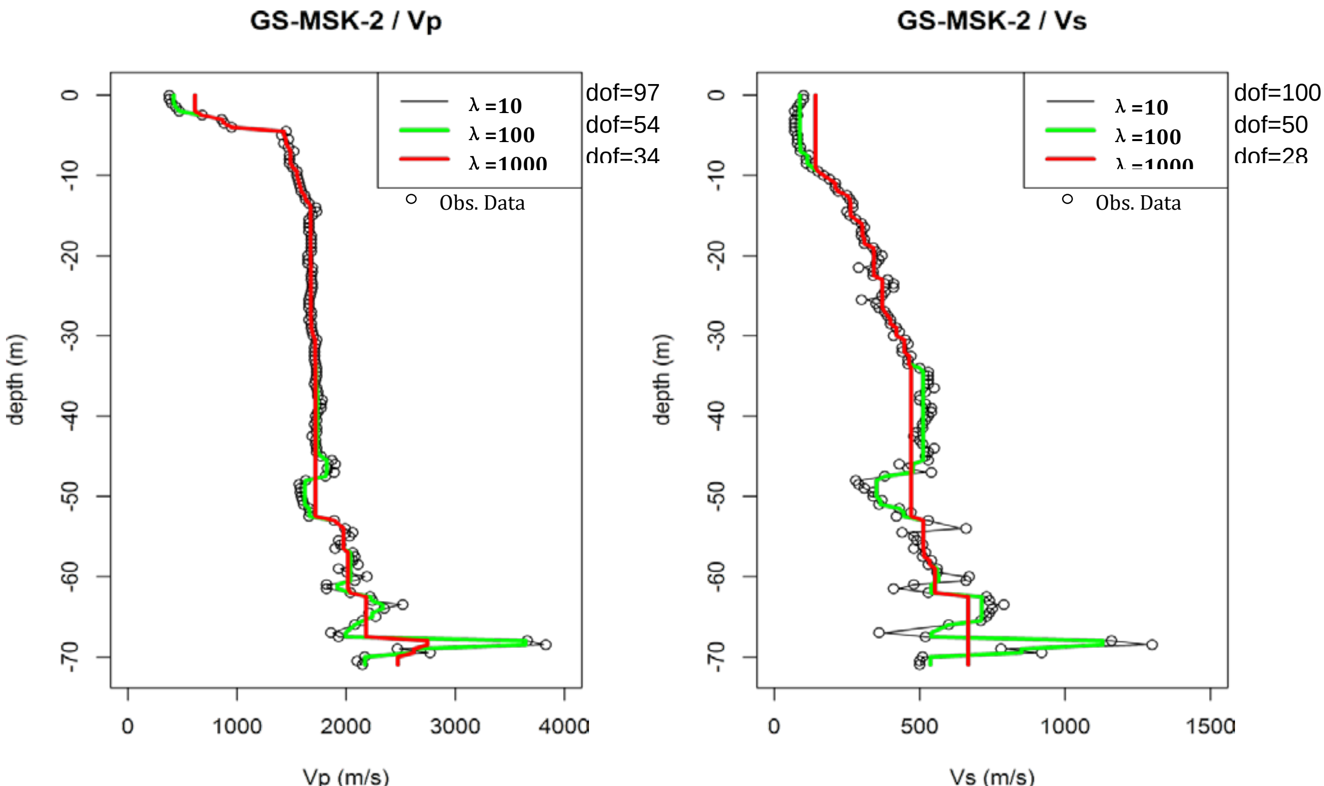

Figure 3. Modeling of the suspension PS logging results at GS-MSK-2 (Number of observed data: 141)

IOP Conf. Series: Materials Science and Engineering **652** (2019) 012023 doi:10.1088/1757-899X/652/1/012023

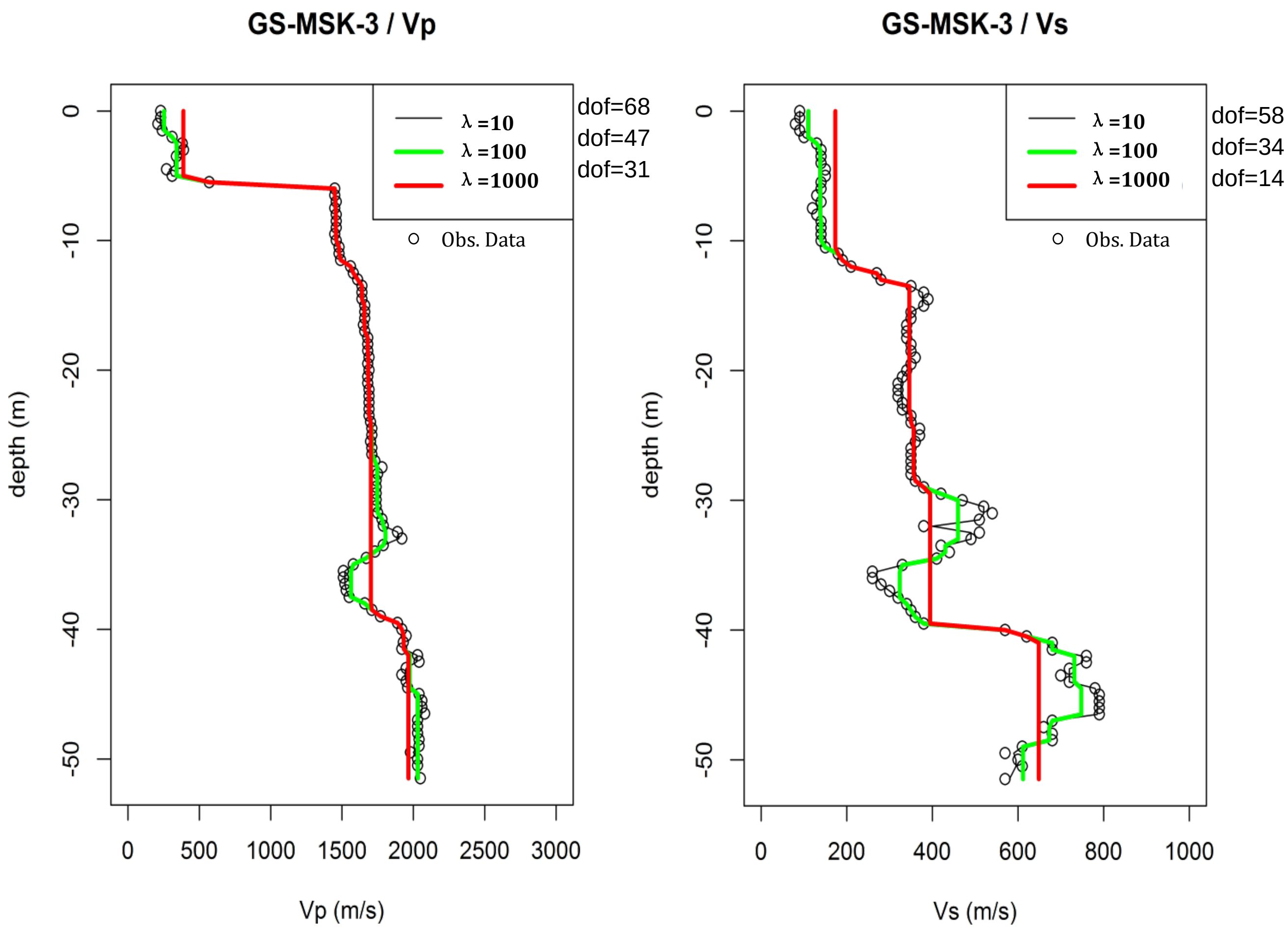

Figure 4. Modeling of the suspension PS logging results at GS-MSK-3 (Number of observed data: 103)

3.1. Outline of the survey data

The PS logging results carried out by Yoshimi et al. (2017, 2016) [5, 6] at three places around Miyazono and Yasunaga areas in Mashiki town, Kumamoto prefecture were used. Miyazono and Yasunaga were heavily damaged areas struck by the 2016 Kumamoto earthquake (Mw 7.0) occurred on April 16. It was a survey conducted to resolve the mechanism of earthquake damage in the area. The ground in the survey area consists of the shallow volcanic ash soil and the deeper pyroclastic flow layer from Mt. Aso. The surveyed sites are shown in Figure 1. Since the suspension type PS logging was carried out, high resolution data was obtained. The step size in the depth direction is 50 cm.
The PS logging data obtained at the three sites are shown Figure 2 to Figure 4. In these figures, open circles indicate observed data.

3.2. The sparse modeling of PS logging data
The sparse modeling of PS logging data is shown in the Figure2 to Figure 4. Each figure shows the fused lasso solutions when the value of the regularization parameters (λ) are changed. The degrees of freedom (dof) of the estimated model are also showed.
Regardless of the surveyed sites, the estimated model follows the fluctuation of the observation data when the value of the regularization parameter is small. On the other hand, with the increasing regularization parameter (λ), it can be seen that the obtained estimation model has many consecutive parts and the inflection points are decreasing. The regularization term of Eq.(4) represents the absolute value of the difference between adjacent model parameters. By minimization of this equation,

estimated values with reduced numbers of discontinuous points have been obtained. Such results are similarly obtained for both P wave velocity (Vp) and S wave velocity (Vs).

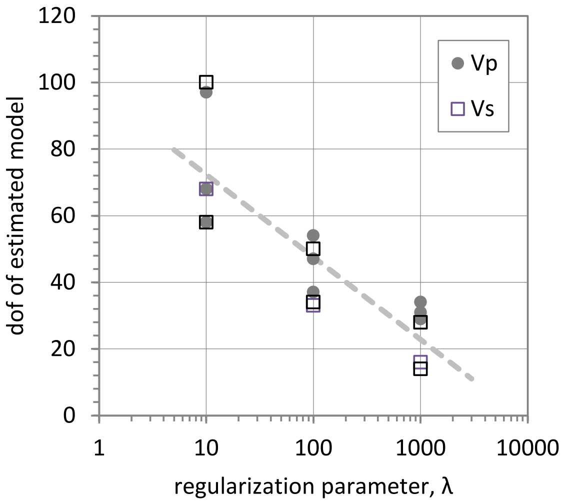

Figure 5. Relationship between the regularization parameter and the dof of estimated model

Figure 5 shows the relationship between the regularization parameter and the dof of estimated model. In this figure, symbols means the results of sparse estimation, and the broken line represents their regression.

4. Conclusions

In this paper, the sparse modeling was applied to the geotechnical observation data and examined its effectiveness. The fused lasso, which is a typical method of sparse modeling, was applied to model the velocity structure of the ground by using the PS logging data. The findings obtained in this research are summarized below.

- ✓ The 1D fused lasso was applied to model the velocity structure of the ground by using the continuous data with large fluctuation. The applicability of this method to the PS logging data was indicated.
- ✓ It was found that simplified models can be obtained by increasing the value of the regularization parameter.
- ✓ A negative correlation was found between the logarithm of the normalization parameter and the dof of the estimation model.

In the future, the applicability of the space modeling to another kind of observation data will be investigated to confirm the validity.

References

[1] Liu B, Wang M, Foroosh H, Tappen M, Penksy M 2015 Sparse Convolutional Neural Networks, *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 806-814

[2] Lusig M, Donoho D, Pauly J M 2007 Sparse MRI: The Application of Compressed Sensing for Rapid MRI Imaging, *Magnetic Resonance in Medicine*, **58**, 1182–1195

[3] Mairal J, Elad M, Sapiro G 2008 Sparse Representation for Color Image Restoration, *IEEE Transactions on Image Proceeding*, **17**, 1, 53-69
[4] Tibshirani R J and Taylor J 2011 The Solution Path of the Generalized Lasso, *The Annals of Statistics*, **39**, 3, 1335-1371
[5] Yoshimi M, Goto H, Hata Y, Yoshida N 2017 Nonlinear site response at the worst hit area of the 2016 Kumamoto earthquakes in the Mashiki town, Kumamoto, Japan, *Document of DPRI Annual Meeting*, *Disaster Prevention Research Institute*, *Kyoto University*, A05 *(in Japanese)*
[6] Yoshimi M, Hata Y, Goto H, Hosoya T, Morita S, Tokumaru T 2016 Borehole exploration in heavily damaged area of the 2016 Kumamoto earthquake, Mashiki town, Kumamoto, *Proceedings of the Japanese Society for Active Fault Studies 2016 Annual Meeting*, P17 *(in Japanese)*