Applied Soft Computing 22 (2014) 118–125

Contents lists available at ScienceDirect

Applied Soft Computing

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

Prediction of elastic compressibility of rock material with soft computing techniques

Zaobao Liu[a,b,*], Jianfu Shao[a,b], Weiya Xu[b], Yu Zhang[a], Hongjie Chen[b]

[a] Geotechnical Research Institute, Hohai University, Nanjing 210098, China
[b] Laboratory of Mechanics of Lille, University of Lille I, 59655 Villeneuve d'Ascq, France

ARTICLE INFO

Article history:
Received 21 March 2013
Received in revised form 14 October 2013
Accepted 13 May 2014
Available online 21 May 2014

Keywords:
Soft computing
Relevance vector machine
Mechanical parameter
Porous material
Artificial neural network
Support vector machine

ABSTRACT

Mechanical and physical properties of sandstone are interesting scientifically and have great practical significance as well as their relations to the mineralogy and pore features. These relations are however highly nonlinear and cannot be easily formulated by conventional methods. This paper investigates the potential of the technique named as the relevance vector machine (RVM) for prediction of the elastic compressibility of sandstone based on its characteristics of physical properties. Based on the fact that the hyper-parameters may have effects on the RVM performance, an iteration method is proposed in this paper to search for optimal hyper-parameter value so that it can produce best predictions. Also, the qualitative sensitivity of the physical properties is investigated by the backward regression analysis. Meanwhile, the hyper-parameter effect of the RVM approach is discussed in the prediction of the elastic compressibility of sandstone. The predicted results of the RVM demonstrate that hyper-parameter values have evident effects on the RVM performance. Comparisons on the results of the RVM, the artificial neural network and the support vector machine prove that the proposed strategy is feasible and reliable for prediction of the elastic compressibility of sandstone based on its physical properties.

Introduction

The mechanical and physical properties of sandstone are fascinating scientifically and have great practical significance as well as their relations to its microstructural characteristics [1–4]. These relations between the mechanical parameters and physical properties are however highly nonlinear and cannot be easily formulated by conventional methods. Sandstone is a typical kind of porous media composed of solid particles and pore spaces. Solid particles form the skeleton of sandstone and are surrounded by the pore spaces. The solid particles consist of various kinds of minerals, such as the quartz, the feldspar and the detrital clay. Three types of pores are often measured by the laboratory experiments: (i) the intergranular pore, (ii) the connective pore and (iii) the micro pore. The mechanical behaviours of sandstones are closely related to the physical properties such as the mineral composition and pore properties. The compositions of solid particles and pores of sandstone primarily result in its varying mechanical behaviours. Hence, mechanical behaviours of sandstones may be estimated according to the characteristics of the mineral compositions and the pore features.

Elastic compressibility is a common mechanical parameter of such porous media like the sandstone. Recently, the elastic compressibility of different materials has been studied by various kinds of techniques [5–8]. The physical properties of sandstone have been tested in the laboratory to investigate their relationship with microstructural features and loading pressures [9]. Substantial experiments have been done to identify the mechanical parameters such as the elastic compressibility. The obtained relationships between the elastic compressibility and the previous maximum pressure show that the elastic compressibility at any given pressure is a function of the previous stress history of the sample [10]. However no specific method has been presented to account for these relations. Thus the estimation of the elastic compressibility is difficult based on the loading pressures. Substantial discussions have been made on the compressibility of sandstones including the stress and the compressibility, the pore structure and the compressibility and the laboratory measurements of the compressibility [7]. Also, methods are presented with permission of quantitative predictions of the sandstone compressibility. Some work on prediction of the rock compressibility behaviours has been undertaken using the neural networks and support vector machine based on rock physical properties [11,12]. Some experiments have

* Corresponding author. Tel.: +33 3 20 43 46 30.
E-mail address: zaobao.liu@polytech-lille.fr (Z. Liu).

http://dx.doi.org/10.1016/j.asoc.2014.05.009
1568-4946/

been done to discover the fractal characteristics of the pore structure of the low permeability sandstone [13]. A pore structure model is applied to predict the elastic wave velocities in fluid-saturated sandstones [14]. These results show that pore structures do influence the characteristics of sandstones. Hence, the effects of the mineral compositions as well as the microstructure features should be considered for the estimation of the elastic compressibility.

It is sometimes expensive to measure certain mechanical features of porous materials. Recently, the soft computing techniques including the neural network, the support vector machine [15], the cloud models [16], and the Gaussian process [17] have been incorporated into rock mechanics and engineering to discover the mechanical behaviours of rocks [18–24] as well as the associated hazards [25–30]. In this study, we focus on the estimation of the elastic compressibility of sandstone according to its characteristics of the mineral compositions and the pore structures as well as the loading pressure using the state-of-art soft computing technique named as the relevance vector machine (RVM) [31]. The RVM is a probabilistic approach for learning disciplines in data. The hyper-parameters influence the modelling performance of RVMs [32]. The application of the RVM has been presented in the slope reliable analysis [33]. However, the effect of the hyper parameter has not been discovered yet and usually given manually.

In this paper, we present an iterative strategy to optimize the RVM hyper-parameters adaptively and then apply this approach to estimate the elastic compressibility of sandstones. We aim to: (1) show the qualitative sensitivity of the factors to the elastic compressibility of sandstones; (2) show the effects of hyper-parameters of RVM in the estimation of the elastic compressibility of sandstones; (3) investigate the potential capability of the adaptive RVM in the estimation of the elastic compressibility of sandstone according to its characteristics of properties like the mineral composition and the pore features. The specific mechanical effects of the minerals or the pore structure are not included in this study.

Methodology: adaptive relevance vector machine

Relevance vector machine for multivariable estimation

The RVM is proposed in the Bayesian framework with probabilistic significance [31]. It produces a mapping between the target variables and the associated independent variables, i.e.

$$y = f(x; w) = \sum_{i=1}^{N} w_i K(x, x_i) + w_0 = w^t \phi(x) \quad (1)$$

where N is the total sample number; i is the ith sample number; $y = f(x; w)$ is the mapping; x, x_i denote the associated independent variables; $K(x, x_i)$ is the kernel functions; $\phi(x) = [1, K(x, x_1), K(x, x_2), \ldots, K(x, x_n)]^t$; w_i is the weight of ith sample, $w = [w_0, w_1, \ldots, w_N]^t$.

In regression, the RVM employs Eq. (1) with an additive noise term to link the input x_n and scalar target variable t_n

$$t_n = f(x_n; w) + \epsilon_n \quad (2)$$

where ϵ_n is a zero-mean white noise process with variance σ^2, that is $p(\epsilon_n|\sigma^2) = N(\epsilon_n|0, \sigma^2)$.

Posing $\beta = \sigma^2$, and assuming independence of the samples, the likelihood of the training samples is

$$p(t|X, w, \beta) = (2\pi\beta^{-1})^{-N/2} \exp\left(-\frac{1}{2}\beta\left\|t - \phi w\right\|^2\right) \quad (3)$$

where $t = [t_1, \ldots, t_N]^t$, $X = [X_n]_{n=1}^N$. With more parameters $(N+1)$ than training data samples (N), direct maximum-likelihood estimation of w would lead to over-fitting. In the RVM Bayesian framework, zero-mean Gaussian shrinkage priors are imposed on every w_i and, assuming the independence of the parameters, one can have

$$p(w_i|\alpha_i) = N(w_i|0, \alpha_i^{-1}) \Rightarrow p(w|\alpha) = \prod_{i=0}^{N} N(w_i|0, \alpha_i^{-1}) \quad (4)$$

with $\alpha = [\alpha_0, \alpha_1, \ldots, \alpha_N]^t$, a $N+1$ vector of hyper-parameters representing the precision on the parameters.

Finally uniform hyper-priors are assumed for the precision hyper-parameters, α and β. An interesting property of these hyper-priors is that when the evidence of the model is maximized with respect to the hyper-parameters the corresponding parameters turn to be zero. This is a type of "automatic relevance determination" [34] leading to a sparse set of parameters w. Using Bayes rule and the properties of Gaussian functions, the posterior distribution of the weight can also be described by a Gaussian:

$$p(w|X, t, \alpha, \beta) = N(w|m, \Sigma) \quad (5)$$

where the mean m and the covariance Σ are given by

$$m = \beta\Sigma\Phi^t t; \quad \Sigma = (A + \beta\Phi^t\Phi)^{-1} \quad (6)$$

with $A = diag(\alpha_0, \ldots, \alpha_N)$ a diagonal matrix of precisions.

In practice, the values of α and β are estimated by maximizing the marginal likelihood $p(t|X, \alpha, \beta)$, i.e., using a type-II maximum-likelihood method [35]. Only the most probable values are thus calculated, which is an approximation to estimate their full distribution. With this simplification, the marginal likelihood can be obtained by integrating over the weight parameters

$$p(t|X, \alpha, \beta) = \int p(t|X, w, \beta)p(w|\alpha)dw = N(t|0, \beta^{-1}I + \Phi A^{-1}\Phi^t) \quad (7)$$

Values of α and β that maximizes (the log of) Eq. (7) can then be obtained iteratively using the following updating rules

$$\alpha_i^{new} = \frac{1 - \alpha_i\Sigma_{ii}}{m_i^2}; \quad (\beta^{new})^{-1} = \frac{\left\|t - \Phi m\right\|^2}{N - \sum_{i=1}^{N}(1 - \alpha_i\Sigma_{ii})} \quad (8)$$

where m_i is the ith element of the estimated posterior weight w and Σ_{ii} the ith diagonal element of the posterior covariance matrix Σ from Eq. (6).

Once the iterative procedure has converged to the "most probable values α_{MP} and β_{MP}, the distribution of target value t_* for a new data point x_* is also Gaussian and estimated by

$$p(t_*|X, t, \alpha_{MP}, \beta_{MP}) = \int p(t_*|X, w, \beta_{MP})p(w|X, t, \alpha_{MP}, \beta_{MP})dw = N(t_*|m^t\varphi(x_*), \sigma_*^2) \quad (9)$$

$$\sigma_*^2 = \beta_{MP}^{-1} + \varphi(x_*)^t\Sigma\varphi(x_*) \quad (10)$$

where Σ is given by Eq. (6) with α and β set at their optimal value.

Adaption of hyper-parameters

Unlike the optimization of the weight w, the hyper-parameter r^2 is user-defined before model training. In order to optimize the hyper-parameter value, the target function (TF) should be defined first. The adaption of the kernel parameters goes in the following steps:

(a) Initialize the hyper-parameter r^2 in a proper range and specify the initial r_0^2 and an appropriate step for iteration;
(b) Specify one kernel function;
(c) Train the RVM model with the data set to obtain an optimal α_{MP} and the weight w;