An attribute-guided method for reservoir parameter prediction using relevance vector machine

Introduction

Seismic attributes are the geometric shape, kinematic characteristics, dynamic characteristics and statistical characteristics of seismic waves derived from the mathematical transformation of prestack or poststack seismic data (Chen et al., 1997). These characteristics contain a wealth of information about stratum, lithology and fluid. How to obtain more reservoir information from seismic attributes has become a hot spot and difficult point in geophysical research. A variety of statistical methods are used to study the relationship between seismic attributes and reservoir parameters, such as multiple linear regression (e.g., Hart et al., 2000), neural network (e.g., Hampson et al., 2001) and support vector machine (e.g., Cheng et al., 2011). However, multiple linear regression has difficulty in getting the nonlinear relationship between seismic attributes and reservoir parameters. Although the neural network can map arbitrarily complex nonlinear relationships, it does not have a firm theoretical basis. Relevance Vector Machine (RVM) is a sparse probability model under the Bayesian framework (Tipping, 2001; Tipping, 2003). Compared with SVM, RVM leads to a sparser solution and has better generalization capacity. RVM behaves well in both synthetic data and field data, demonstrating a great potential of this method in reservoir prediction (Ji et al., 2019; Pilikos, 2020). Based on the RVM, this paper takes field data as an example to predict porosity. A reasonable porosity distribution of the target layer is obtained. Afterwards, the uncertainty of predicting outcomes is provided by variance in case of uncertainty analysis.

Method and Theory

The RVM is a data-driven probability model. It assumes that the target variable is a linear combination of a series of nonlinear basis functions, as shown in the following formula:

\[ t_i = \sum_{n=1}^{N} \omega_n \phi_n(K_i) + \varepsilon_i = \omega^T \phi(K_i) + \varepsilon_i \]  

where \( t_i \) are objective function, \( K_i \) are kernel function, \( \omega_n \) are the coefficients of the linear combination of the transformed input data, \( \varepsilon_i \in N(0, \sigma^2) \) are independent and identically distributed random noise. The likelihood function is:

\[ p(t|\omega, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp \left\{ -\frac{1}{2\sigma^2} \| t - \phi \omega \|^2 \right\} \]  

where, \( \phi \in \mathbb{R}^{N \times L} \), \( \phi = [\phi_1, \phi_2, ..., \phi_N] \), \( \phi_n = [1, K(x_n, x_1), K(x_n, x_2), ..., K(x_n, x_N)]^T \), \( t = (t_1, t_2, ..., t_N)^T \).

We assume that the constant parameters \( \omega_0 \) are independent of each other, and they correspond to hyperparameter \( \alpha_n \), respectively, the prior disturbance of the parameters follows:

\[ p(\omega|\alpha) = \prod_{n=1}^{N} \sqrt{\frac{\alpha_n}{2\pi}} \exp \left\{ -\frac{\alpha_n \omega_n^2}{2} \right\} \]  

Equation 3 is a N-element Gaussian distribution with expectation of zero. The hyperparameter \( \alpha_n \) and noise variance \( \sigma^2 \) are Gamma distribution:

\[ p(\alpha) = \prod_{n=1}^{N} Gamma(\alpha_n; a, b) = \prod_{n=1}^{N} \frac{1}{\Gamma(a)} b^a \alpha_n^{a-1} \exp(-b\alpha_n) \]  

\[ p(\sigma^2) = Gamma(\sigma^2; c, b) = \frac{1}{\Gamma(c)} d^c (\sigma^2)^{c-1} \exp(-d\sigma^{-2}) \]  

with a, b, c and d are the scale and shape parameters for the gamma distribution. Typically, the values of a, b, c and d are numbers close to zero such as \( a = b = c = d = 10^{-6} \) (Tipping, 2003). According to Bayes’ formula, the posterior probability is:

\[ p(\omega, \alpha, \sigma^2|t) = \frac{p(t|\omega, \alpha, \sigma^2)p(\omega, \alpha, \sigma^2)}{p(t)} \]

From Bayes’ rule and conjugacy, the posterior distribution of the model coefficients \( \omega \) is standard Gaussian distribution with covariance \( \Sigma \) and mean \( \mu \):

\[ \Sigma = (\sigma^{-2} \Phi^T \Phi + A)^{-1} \]  
\[ \mu = \sigma^{-2} \Phi^T \Phi \]  

where \( A = diag(\alpha_1, \alpha_1, ..., \alpha_N) \). Most of the hyperparameter \( \alpha_n \) tend to infinity during calculating, the corresponding weight coefficient \( \omega_n \) are zero. So, the prediction results are only relevant to the remaining samples of non-zero \( \omega_n \), i.e., relevance vectors.
For a new sample \( x \), the predicted value \( t_x \) under the Bayesian framework is:

\[
p(t_x | t) = \int p(t_x | \omega, \alpha_{MP}, \sigma_{MP}^2) p(\omega | t, \alpha_{MP}, \sigma_{MP}^2) \, d\omega
\]

(9)

\[
(\alpha_{MP}, \sigma_{MP}^2) = \arg \max_{\alpha, \sigma^2} p(\alpha, \sigma^2 | t)
\]

(10)

Then, the RVM problem is focusing on solving the problem of maximizing the posterior distribution \( p(\alpha, \sigma^2 | t) \) of \( \sigma_{MP}^2 \) and \( \alpha_{MP} \). When \( p(\alpha, \sigma^2 | t) \propto p(t | \alpha, \sigma^2)p(\alpha)p(\sigma^2) \) and assuming the distribution of \( p(\alpha) \) and \( p(\sigma^2) \) are uniform, this problem is simplified to solve the maximum \( p(t | \alpha, \sigma^2) \) for \( \alpha_{MP} \) and \( \sigma_{MP}^2 \). Usually, it is impossible to obtain analytic solutions of \( p(t | \alpha, \sigma^2) \). Therefore, an iterative method is adopted to solve this problem. The updated form of the hyperparameters \( \alpha_n \) and noise variance \( \sigma^2 \) are:

\[
\alpha_i^{\text{new}} = \frac{\gamma_i}{\mu_i^2}
\]

(11)

\[
(\sigma^2)^{\text{new}} = \frac{\|t - \phi\mu\|^2}{N - \Sigma_{i=0}^N \gamma_i}
\]

(12)

where \( \gamma_i = 1 - \alpha_i \Sigma_{ii} \). \( \mu_i \) is the mean of posterior weight, \( \Sigma_{ii} \) is posterior covariance matrix of diagonal elements.

The new sample’s predicted distribution is the Gaussian distribution:

\[
p(t_x | t) \sim N(t_x, \sigma_x^2)
\]

(13)

\[
t_x = \mu^T \phi(x)
\]

(14)

\[
\sigma_x^2 = \sigma_{MP}^2 + \phi(x)^T \Sigma \phi(x)
\]

(15)

**Feild Application**

The field data is from western China. The dramatically changing velocity and thickness of the low deceleration zone and thin interbeds of sand and mud in hydrocarbon reservoirs in this area make conventional seismic reservoir prediction methods failed to predict reservoir parameters accurately. This paper tries to predict a dependable porosity of the target layer using the Spearman rank correlation coefficient and RVM algorithm.

In this study, 20 seismic attributes are extracted from seismic data, and the Spearman rank correlation coefficient \( \rho \) is used to quantitatively calculate the correlation between seismic attributes and porosity. The absolute value of \( \rho \) for each attribute and porosity is shown in Figure 1. We can see that these five seismic attributes (Amplitude weighted phase, Derivative, Integrate, Quadrature trace and Relative Acoustic Impedance) have significantly higher absolute value of \( \rho \), indicating that they have a strong correlation with porosity. We select these five seismic attributes to predict porosity.

![Figure 1 Absolute value of Spearman rank correlation coefficient](image)

There are 8 wells in the study area. We randomly select 6 wells as training wells and the remaining 2 wells as blind wells. Then, we take the five seismic attributes as input data to predict porosity by using SVM and RVM algorithm, respectively. The results are shown in Figure 2. Table 1 shows the correlation coefficients of the two blind wells predicted by RVM are 0.930 and 0.944 while that of SVM are 0.892.
and 0.928. Meanwhile, the root mean square errors of RVM are 0.054 and 0.046 which are relatively smaller when compared to SVM’s 0.068 and 0.048. The predicting curves are demonstrated in Figure 2, the porosity predicted by RVM algorithm is more consistent with field logging curve, indicating that the accuracy of RVM algorithm is higher than SVM algorithm.

![Figure 2 Porosity prediction results: (a) Well S1 (b) Well S2](image)

<table>
<thead>
<tr>
<th>Well</th>
<th>Correlation coefficient</th>
<th>SVM</th>
<th>RVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.892</td>
<td>0.068</td>
<td>0.054</td>
</tr>
<tr>
<td>S2</td>
<td>0.928</td>
<td>0.048</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Figure 3 shows the cross-well profiles of porosity predicted by RVM algorithm. According to porosity logging curve, the porosity is high at 2114ms of Well S1 and 2106ms of Well S2. High porosity area of RVM profile is consistent with logging data as can be seen in Figures 3a and 3b (marked by purple circle), which proves that this method predicts porosity profile with a better lateral continuity. Figure 4 is the corresponding variance profile to Figure 3. Most of the variance being less than 1 (blue in Figure 4) proves that the prediction is relatively trustable. Figure 5 is a slice map of predicted porosity at the target layer. The porosity of the target layer is higher in the west and lower in the east, which is consistent with the field drilling and logging data. Figure 6 is the corresponding variance slice to Figure 5. Small value of variance (deep blue in Figure 6) shows the uncertainty is low. Therefore, RVM algorithm has a decent stability in porosity prediction.

Conclusions

This paper uses RVM algorithm to predict the porosity of target layer from seismic, logging and geological data. Compared to SVM algorithm, RVM’s prediction has a greater correlation coefficient and a lower RMSE in predicting porosity. Furthermore, RVM algorithm quantifies the uncertainty of prediction results which is helpful for interpreters to evaluate the reliability of prediction results. Cross-well profile of predicted porosity and the porosity slice of target layer show that the porosity predicted by RVM algorithm not only has a good lateral continuity, but also accords well with field drilling and logging data. It can be inferred that this method is capable of predicting the clastic reservoirs parameters.

Acknowledgements

This work is financially supported by NSFC Basic Research Program on Deep Petroleum Resource Accumulation and Key Engineering Technologies (Grant No. U19B6003).
Figure 3 Cross-well profiles of porosity (a) Well S1 (b) Well S2

Figure 4 The variance of cross-well profiles of porosity (a) Well S1 (b) Well S2

Figure 5 Porosity sliced along the layer  
Figure 6 The variance of porosity sliced along the layer

References

Hart, B. S. and Balch, R. S. [2000] Approaches to defining reservoir physical properties from 3-D seismic attributes with limited well control: An example from the Jurassic Smackover Formation, Alabama. Geophysics, 65, 368-376.