Journal of Applied Geophysics 133 (2016) 1–8

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Frequency-domain sparse Bayesian learning inversion of AVA data for elastic parameters reflectivities

Yongzhen Ji, Sanyi Yuan *, Shangxu Wang, Li Deng

China University of Petroleum, State Key Laboratory of Petroleum Resources and Prospecting, CNPC Key Laboratory of Geophysical Exploration, Beijing 102249, China

ARTICLE INFO

Article history:
Received 6 August 2015
Received in revised form 7 July 2016
Accepted 19 July 2016
Available online 22 July 2016

Keywords:
AVA inversion
Sparse Bayesian learning
Precondition
Hierarchical Gaussian distribution

ABSTRACT

The prestack amplitude variation with angle (AVA) inversion method utilising angle information to obtain the elastic parameters estimation of subsurface rock is vital to reservoir characterisation. Under the assumption of blocky layered media, an AVA inversion algorithm combining prestack spectral reflectivity inversion with sparse Bayesian learning (SBL) is presented. Prior information of the model parameters is involved in the inversion through the hierarchical Gaussian distribution where each parameter has a unique variance instead of sharing a common one. The frequency-domain prestack SBL inversion method retrieves sparse P- and S-wave impedance reflectivities by sequentially adding, deleting or re-estimating hyper-parameters without pre-setting the number of non-zero P- and S-wave reflectivity spikes. The selection of frequency components can help get rid of noise outside the selected frequency band. The precondition of the parameters helps to balance the weight of different parameters and incorporate the relationship between those parameters into the inversion process, thus improves the inversion result. Synthetic and real data examples illustrate the effectiveness of the method.

1. Introduction

The amplitude variation with angle (AVA) inversion method utilising angle information to obtain the elastic parameters of subsurface rock plays a critical role in reservoir characterisation. Migrated prestack seismic data can be transformed into the angle domain as AVA data. Parameters such as P-wave velocity, S-wave velocity and density are then inverted, leading to an AVA inversion problem. AVA inversion in the time domain, which is commonly used, utilises all the frequency information of seismic data. However, information outside ~5 Hz to ~100 Hz in seismic data is relatively inaccurate. The frequency-domain AVA inversion in this paper has an optional frequency range, which can flexibly utilise dominant information and get rid of noise outside the chosen frequency band.

For the sake of non-uniqueness, regularisation in most inversion applications is required to select an optimum model amongst many possible solutions and integrate priori information into the inversion process (e.g., Yuan et al., 2015). In the context of AVA inversion, stabilising inversion algorithms and reducing non-uniqueness with regularisation are also important. Sparse constraint is often chosen to be the regularisation term as sparse solutions are synonymous with high-resolution solutions (Levy and Fullagar, 1981; Alemie and Sacchi, 2011). Therefore, various penalty terms of sparse constraint, such as L1 norm, Cauchy criterion or Huber criterion, have been adopted into the nonlinear objective function by previous researchers (e.g., Levy and Fullagar, 1981; Sacchi et al., 1994; Sacchi, 1997; Zhang and Castagna, 2011; Yuan et al., 2016) to promote a sparse reflectivity estimate. Under the Bayesian framework, the sparse constraint is involved by prior information of model parameters. To obtain a sharp boundary, Theune et al. (2010) analysed two prior models for blocky inversion and found that the differentiable Laplace distribution defines a convex function, whereas the Cauchy distribution does not. Alemie and Sacchi (2011) introduced the trivariate Cauchy distribution which can incorporate correlations between model parameters into the Bayesian inversion and lead to a high-resolution result, but does not allow for computing errors. Under the assumption of blocky layered media, we integrate the prior information of model parameters flexibly into the inversion by a parameterised Gaussian distribution. The hyper-parameters in the priori information which are regarded as specific values inspire sparseness (Wipf and Rao, 2004).

Many methods are presented to solve the AVA inversion problem, such as gradient projection for sparse reconstruction (Figueiredo et al., 2007), fast iterative shrinkage-thresholding algorithm (FISTA) (Beck and Teboulle, 2009), basis pursuit (Zhang and Castagna, 2011) and hybrid FISTA least-squares (FISTA + LS) strategy (Pérez et al., 2013). The current study adopts sparse Bayesian learning (SBL) to solve the AVA inversion problem. SBL is proposed and proven to be an effective and accurate method for regression and classification problems (Tipping, 2001). The SBL paradigm performs parameter learning via type-II maximum likelihood or evidence maximisation (Mackay, 1992), in which marginal likelihood maximisation leads to automatic

* Corresponding author.
E-mail address: yuansy@cup.edu.cn (S. Yuan).

http://dx.doi.org/10.1016/j.jappgeo.2016.07.016

identification of the relevant kernels and makes the model sparse. The SBL model has numerous advantages, such as probabilistic prediction, utilisation of arbitrary basis functions and automatic estimation of nuisance parameters. This model is successful in a wide range of applications, such as image processing (Demir and Ertürk, 2007), positron emission tomography (Peng et al., 2008) and dynamic light scattering (Nyeo and Ansari, 2011). The application of SBL in seismic data, which was introduced by Yuan and Wang (2013), is to identify thin beds below tuning thickness and highlight the stratigraphic boundaries for poststack seismic data under the blocky layer assumption.

When dealing with the correlation amongst parameter sets (Gardner et al., 1974; Castagna et al., 1985; Potter and Stewart, 1998), Alemie and Sacchi (2011) incorporated the relationship amongst model parameters by scale matrix in the prior distribution to replace the role of covariance matrix and link small-weight parameters with large-weight parameters. Zong et al. (2012) adopted decorrelation by the eigenvector analysis of the covariance matrix to stabilise the inversion process. The present study also adopts a precondition method to improve the estimation of model parameters.

The remainder of this paper is organised as follows. Firstly, we describe the related theories, including the AVA spectral reflectivity forward model and the SBL inversion method. Secondly, we use synthetic and field data examples to test the performance of our method. Finally, several conclusions are drawn.

2. Theory

The reflection and transmission coefficients on the boundary between two media changing with different incident angles can be commonly described by the Zoeppritz equation (Zoeppritz, 1919). However, such a description is computationally complex for the AVA inversion of prestack seismic data because of its nonlinear form. For this reason, different linear approximations of the Zoeppritz equation are developed for AVA analysis (e.g., Aki and Richards, 1980; Shuey, 1985; Fatti et al., 1994). The ultimate goal of AVA inversion is to delineate the distributions of lithology and liquid, as well as convert the inverted parameters to the lithology and liquid indicators. Hence, the approximations of the Zoeppritz equation should be selected carefully to reduce accumulative transformation error and enhance the stability of the inversion process (Zong et al., 2012). Fatti's approximation (Fatti et al., 1994) is adopted in the current study because of three reasons: 1) The posterior probability density inferred in the model space is maximally decoupled amongst P-wave impedance, S-wave impedance and density, with the largest error on the density parameter (Debski and Tarantola, 1995; Rabben et al., 2008), and the coefficient matrix of these three parameters has a relatively small condition number. 2) The parameter set can be readily converted to common lithology and liquid indicators (e.g., V_P/V_S, $\mu\rho$, $\lambda\rho$), which have good capabilities for lithology and fluid identification in a wide range. 3) Considering that the large-angle information of prestack data is difficult to acquire and has poor quality, estimation of density term is often deceptive, and hence affects the estimation of the first two terms. However, when Fatti's approximation is adopted, if necessary, the density term can be excluded from inversion process without sacrificing much accuracy.

Fatti's approximation (Fatti et al., 1994) is

$$R_{PP}(\theta) = (1+\tan^2\theta)R_{IP}/2 - 4\gamma^2\sin^2\theta R_{IS} + \left(4\gamma^2\sin^2\theta - \tan^2\theta\right)R_\rho/2, \quad (1)$$

where R_{IP}, R_{IS} and R_ρ are the P-wave impedance, S-wave impedance and density reflectivities at interface, γ represents the ratio of S-wave velocity to P-wave velocity and θ is the incident angle. The equation can be written in a matrix/vector form as

$$R_{PP}(\theta) = [A(\theta)\ \ B(\theta)\ \ C(\theta)][R_{IP}\quad R_{IS}\quad R_\rho]^T, \quad (2)$$

where $A(\theta)=(1+\tan^2\theta)/2$, $B(\theta)=-4\gamma^2\sin^2\theta$, $C(\theta)=(4\gamma^2\sin^2\theta-\tan^2\theta)/2$ and the superscript T represents transpose. This form can be regarded as the linearly weighted superposition of the three parameters and the corresponding weights $A(\theta)$, $B(\theta)$ and $C(\theta)$. A prestack spectral reflectivity forward model is then obtained by taking the Fourier transform for both sides of Eq. (2) as

$$\tilde{\boldsymbol{R}}_{PP}(\theta) = \boldsymbol{F}[\boldsymbol{A}(\theta)\quad \boldsymbol{B}(\theta)\quad \boldsymbol{C}(\theta)][\boldsymbol{R}_{IP}\quad \boldsymbol{R}_{IS}\quad \boldsymbol{R}_\rho]^T, \quad (3)$$

where $\boldsymbol{F} = \begin{bmatrix} \exp(-i2\pi t_1 f_1) & \exp(-i2\pi t_2 f_1) & \cdots & \exp(-i2\pi t_K f_1) \\ \exp(-i2\pi t_1 f_2) & \exp(-i2\pi t_2 f_2) & \cdots & \exp(-i2\pi t_K f_2) \\ \vdots & \vdots & \vdots & \vdots \\ \exp(-i2\pi t_1 f_M) & \exp(-i2\pi t_2 f_M) & \cdots & \exp(-i2\pi t_K f_M) \end{bmatrix}$ represents the discrete version of the Fourier transform, $f_m(m=1,2,3,\cdots,M)$ are the frequencies we select within the limited band, $t_k(k=1,2,3,\cdots,K)$ are the time sample, $\tilde{\boldsymbol{R}}_{PP}$ is the Fourier spectrum of the time-domain P-wave reflectivity that can be obtained by using $\mathbf{S}(f)/\mathbf{W}(f)$ or the spectrum of the deconvolution result, $\mathbf{S}(f)$ represents the Fourier spectrum of seismic angle data and $\mathbf{W}(f)$ represents the Fourier spectrum of the corresponding angle wavelet, $\boldsymbol{A}(\theta)=diag[A(t_1,\theta)\quad A(t_2,\theta)\quad \cdots \quad A(t_K,\theta)]$, $\boldsymbol{B}(\theta)=diag[B(t_1,\theta)\quad B(t_2,\theta)\quad \cdots \quad B(t_K,\theta)]$, $\boldsymbol{C}(\theta)=diag[C(t_1,\theta)\quad C(t_2,\theta)\quad \cdots \quad C(t_K,\theta)]$, $\boldsymbol{R}_{IP}=[R_{IP}(t_1)\quad R_{IP}(t_2)\quad \cdots R_{IP}(t_K)]$, $\boldsymbol{R}_{IS}=[R_{IS}(t_1)\quad R_{IS}(t_2)\quad \cdots R_{IS}(t_K)]$, $\boldsymbol{R}_\rho=[R_\rho(t_1)\quad R_\rho(t_2)\quad \cdots \quad R_\rho(t_K)]$ and $diag[\cdots]$ represents the diagonal matrix. If N angle gathers, i.e., $\theta\in[\theta_1\quad \theta_N]$, exist, we have

$$\left[\tilde{\boldsymbol{R}}_{PP}(\theta_1)\tilde{\boldsymbol{R}}_{PP}(\theta_2)\vdots\tilde{\boldsymbol{R}}_{PP}(\theta_N)\right] = \begin{bmatrix} \boldsymbol{F} & 0 & \cdots & 0 \\ 0 & \boldsymbol{F} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{F} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{A}(\theta_1) & \boldsymbol{B}(\theta_1) & \boldsymbol{C}(\theta_1) \\ \boldsymbol{A}(\theta_2) & \boldsymbol{B}(\theta_2) & \boldsymbol{C}(\theta_2) \\ \vdots & \vdots & \vdots \\ \boldsymbol{A}(\theta_N) & \boldsymbol{B}(\theta_N) & \boldsymbol{C}(\theta_N) \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_{IP}^T \\ \boldsymbol{R}_{IS}^T \\ \boldsymbol{R}_\rho^T \end{bmatrix}, \quad (4)$$

For simplicity, when the noise is considered, the equation is rewritten in the following matrix/vector form:

$$\boldsymbol{d}_{M*N} = \boldsymbol{G}_{M*N\times 3K}\boldsymbol{m}_{3K} + \boldsymbol{n}, \quad (5)$$

where $\boldsymbol{d} = \left[\tilde{\boldsymbol{R}}_{PP}(\theta_1)\tilde{\boldsymbol{R}}_{PP}(\theta_2)\cdots\tilde{\boldsymbol{R}}_{PP}(\theta_N)\right]^T$ is the angle data vector, $\boldsymbol{m} = [\boldsymbol{R}_{IP}\quad \boldsymbol{R}_{IS}\quad \boldsymbol{R}_\rho]^T$ is the parameter vector and $\mathbf{n}$ is the noise vector.

In order to balance the weight of different parameters and consider the relationship amongst such parameters, we build the precondition matrix in two steps. Firstly, we compose the precondition matrix as $\boldsymbol{W}_1 = \begin{bmatrix} \lambda_1\cdot\boldsymbol{I} & 0 & 0 \\ 0 & \lambda_2\cdot\boldsymbol{I} & 0 \\ 0 & 0 & \lambda_3\cdot\boldsymbol{I} \end{bmatrix}$ to balance the weights of different parameters, where $\mathbf{I}$ is the identity matrix and $\lambda_i(i=1,2,3)$ are different precondition coefficients for R_{IP}, R_{IS} and R_ρ respectively. Secondly, the relationship amongst parameters are introduced by covariance analysis, i.e., adding the covariance of each two different parameters to the precondition matrix, to achieve the final precondition matrix $\boldsymbol{W} = \begin{bmatrix} \lambda_1\cdot\boldsymbol{I} & \varepsilon\cdot cov_{R_{IP}R_{IS}}\cdot\boldsymbol{I} & \varepsilon\cdot cov_{R_{IP}R_\rho}\cdot\boldsymbol{I} \\ \varepsilon\cdot cov_{R_{IP}R_{IS}}\cdot\boldsymbol{I} & \lambda_2\cdot\boldsymbol{I} & \varepsilon\cdot cov_{R_{IS}R_\rho}\cdot\boldsymbol{I} \\ \varepsilon\cdot cov_{R_{IP}R_\rho}\cdot\boldsymbol{I} & \varepsilon\cdot cov_{R_{IS}R_\rho}\cdot\boldsymbol{I} & \lambda_3\cdot\boldsymbol{I} \end{bmatrix}$, where cov represents the covariance of two different parameters and ε is the normalisation coefficient. The choice of $\lambda_i(i=1,2,3)$ is according to the available well-log data or the rule of thumb if there are no well-log data available. The choice of ε is according to trial and error. Then, if we define $\boldsymbol{G}_{M*N\times 3K}{}' = \boldsymbol{G}_{M*N\times 3K}\boldsymbol{W}_{3K\times 3K}$ and $\boldsymbol{m}_{3K}{}' = \boldsymbol{W}_{3K\times 3K}^{-1}\boldsymbol{m}_{3K}$, hence Eq. (5) becomes

$$\boldsymbol{d}_{M\times N} = \boldsymbol{G}'_{M*N\times 3K}\boldsymbol{m}'_{3K} + \boldsymbol{n}. \quad (6)$$