Multichannel Seismic Deconvolution Using Bayesian Method

Li Yanqin, Peng Hongwei, Yu Ruihong

1. Department of Disaster Prevention Equipment, Institute of Disaster Prevention, Beijing 101601, China
2. School of Electrical Engineering & Automation, Tianjin University, Tianjin 300072, China

Tel.: 010-58412954, fax: 01457422953
E-mail: liyanqin-2190@126.com

Received: 9 September 2013 /Accepted: 25 October 2013 /Published: 30 November 2013

Abstract: In this paper, we propose an algorithm for multichannel blind deconvolution of seismic signals, which exploits variational Bayesian method. It is related to the Kullback–Leibler divergence, which measures the independence degree of deconvolved data sequence. We assume that the reflectivity sequence is almost the same for each receiver while the noise level may differ at each channel. Compared to blind deconvolution of a single seismic trace, multichannel blind deconvolution provides an accurate convergence of the estimated parameters and reflectivity sequence.

Keywords: Multichannel, Seismic deconvolution, Bayesian method, Bernoulli–Gaussian process, Blind deconvolution.

1. Introduction

Seismic waves are reflected times on their way from an airguns source to an observation site. By analyzing the properties of different wave types, it is possible to obtain insight in the inner structure of the Earth [1]. Blind deconvolution is a method used to produce an increased resolution estimate of the reflectivity in the situations where the reflectivity sequence, the source wavelet signal and the noise power level are absolutely unknown. Multichannel blind deconvolution is usually better than single-channel blind deconvolution [2]. Different source wavelet sequences are used to cope with the ill-posed nature of the basic blind deconvolution problem, and to improve estimate reflectivity precision.

Since Bernoulli–Gaussian (BG) model introduced by Mendal, a standard tool, Maximum Likelihood approach (ML) for estimating reflectivity and seismic wavelet is the expectation–maximization (EM) algorithm [3]. However, the EM converges to a local optimum. Stochastic versions of the EM algorithm (SEM) are used to bypass the difficulty [4]. Furthermore, the estimation of the BG noise and source wavelet parameters can also be solved by MCMC techniques in Bayesian framework [5].

The most key point of applying Bayesian theory is to choose the form of the prior distribution ignored by the classical approaches [5], that is, the statistical modeling of sparsity. Since the sparsity is important information about the characteristics of a signal, using Bayesian methods for model parameters on estimation can improve performance. Furthermore, because Bayesian approaches are model-based, they can handle observation noise and missing data problems in seismic deconvolution. After parameter estimation, the maximum posterior mode (MPM) method is carried out to enable reflectivity sequence...
deconvolution [2], which also involves MCMC simulation through Gibbs sampling [6]. The reflectivity deconvolution has been extended to multichannel deconvolution [7]. In cases where several recordings are available, a new multichannel blind deconvolution approach based on the variational Bayesian is proposed in this paper, to improve the wavelet and reflectivity estimation.

This paper is organized as follows: In Section 2, the signal, the prior and the unknown parameters (hyperparameters) models are given, and the formulation of the multichannel blind seismic restoration problem is presented. In Section 3, the variational Bayesian approach to distribution approximation for the multichannel blind deconvolution problem. And then parameters estimation can be obtained. In Section 4, we present simulation results.

2. Data Model

2.1. Signal Model

The observed multichannel signal can be represented as

\[
\begin{bmatrix}
  y_1^\prime \\
  y_2^\prime \\
  \vdots \\
  y_d^\prime
\end{bmatrix}
= \begin{bmatrix}
  s_1^\prime \\
  s_2^\prime \\
  \vdots \\
  s_d^\prime
\end{bmatrix} \ast \begin{bmatrix}
  n_1^\prime \\
  n_2^\prime \\
  \vdots \\
  n_d^\prime
\end{bmatrix} + \begin{bmatrix}
  r_1 \\
  r_2 \\
  \vdots \\
  r_d
\end{bmatrix},
\]

(1)

where

\[
y^\prime = \begin{bmatrix} y_1^\prime, y_2^\prime, \ldots, y_d^\prime \end{bmatrix}^T, s^\prime = \begin{bmatrix} s_1^\prime, s_2^\prime, \ldots, s_d^\prime \end{bmatrix}^T,
\]

\[
n^\prime = \begin{bmatrix} n_1^\prime, n_2^\prime, \ldots, n_d^\prime \end{bmatrix}^T
\]

In which \(l = 1, 2, \ldots, d\) is the index of the different channels. The \(s^\prime\) is the finite impulse response wavelet of length \(N\) in the channel \(l\), which is assumed to be zero-mean Gaussian distribution with unknown variance \(\alpha_{s,l}^{-1}\). \(n^\prime\) stands for additive noise of the length \(N\) in the channel \(l\). The \(n^\prime_{(l=1\cdots d)}\) mutually independent and have \(N(0, \alpha_{n,l}^{-1})\) normal distributions with mean zero and an unknown variance \(\alpha_{n,l}^{-1}\). The reflectivity process \(r = (r_1, \ldots, r_M)\) assumed to be a generalized Bemoulli-Gaussian distribution [6]. This distribution is characterized by an underlying state model \(q = (q(k))_{k=0,K}\) with \(q(k) = 0\) at low reflectivity points and \(q(k) = 1\) at high reflectivity points.

Therefore, the corresponding reflectivity \(r\) is distributed by a zero mean Gaussian distribution with and variance \(\alpha_r^{-1}\) if \(q(k) = 1\) and \(\alpha_r^{-1} (\alpha_r^{-1} <\alpha_r^{-1})\) if \(q(k) = 0\) respectively. That is

\[
r(a) \sim \mathcal{N}(0, \alpha_r^{-1}) + (1 - \lambda)N(0, \alpha_r^{-1})
\]

(2)

Equation (2) can also be written as

\[
p(r|\lambda, \alpha_r) \propto \left(2 + (1 - \lambda)^2\right)^{N/2} \alpha_r^{-1}
\]

(3)

\[
\begin{bmatrix}
  r_1 \\
  r_2 \\
  \vdots \\
  r_d
\end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \alpha_r^{-1} \mathbf{I})
\]

In which \(\mathbf{0}\) is a zero mean Gaussian distribution matrix. The intensity rate of the Bernoulli-Gaussian process \(\lambda\) is that

\[
p(\lambda) = Be(\lambda|\alpha_\lambda, \beta_\lambda) \propto \frac{\lambda^{\alpha_\lambda - 1} (1 - \lambda)^{\beta_\lambda - 1}}{\Gamma(\alpha_\lambda, \beta_\lambda)},
\]

(6)

where \(\Gamma(\alpha_\lambda, \beta_\lambda) = \frac{\Gamma(\alpha_\lambda) \Gamma(\beta_\lambda)}{\Gamma(\alpha_\lambda + \beta_\lambda)}\), \(\alpha_\lambda, \beta_\lambda\) are the scale parameter and shape parameter respectively. And the Beta distribution has the following mean

\[
E[\lambda] = \frac{\alpha_\lambda}{\alpha_\lambda + \beta_\lambda}
\]

(7)

The noise in (1) is assumed to be Gaussian with zero mean and variance equal to \(\alpha_{n,l}^{-1}\), the probability of the observation model is regarded to be

\[
p(y|r, s^\prime, \alpha_{s}) \propto \frac{\left(\left\|y - Ss^\prime\right\|^2\right)^{\frac{\alpha_{s}}{2}}}{\left(\frac{1}{2} \alpha_{s}\right)^{N/2}}
\]

(8)
3. Multichannel Variational Bayesian Blind Seismic Deconvolution

We exploit the variational method to the Bayesian formulation of multichannel blind deconvolution problem. If the estimate $\hat{\lambda},\alpha_1,\alpha',\alpha_2$ is calculated, then the mean of this posterior distribution of the parameters can be used to integrated out into estimate the reflectivity and seismic wavelet by solving

$$\hat{r},\hat{s} = \arg\max_{r,s} p(r,s)$$

Also, the mean value of this posterior distribution can be regarded as the estimate of the reflectivity and seismic wavelet.

The key difficulty in obtaining a practical form for the estimator is that the posterior distribution of the hyperparameters and given the observations is unknown. This fact makes a direct application of the EM algorithm impossible [6]. However, the true posterior needs to be approximated by another factorized distribution which facilitates the calculation of the estimators. The primary idea of the multichannel blind seismic deconvolution is to approximate the true posterior by a free-form distribution which minimizes the Kullback–Leibler (KL($\cdot$)) divergence [8, 9]. That is to say, we approximate $p(\Theta | y)$ by $q(\Theta)$ in a way that

$$\tilde{q}(\theta) = \arg\min_{\theta \in \Theta} C_{KL}(q(\theta) < p(\theta | y)),$$

where $\Theta$ denotes the set $\Theta$ with $\theta$ removed from the set. This method results in the following solution.

$$\tilde{q}(\theta) = \text{const} \times \exp\left(\mathbb{E}\left[\log p(\theta)p(y|\theta)\right]_{q(\theta)}\right)$$

With

$$\mathbb{E}\left[\log p(\theta)p(y|\theta)\right]_{q(\theta)} = \int p(\theta)p(y|\theta)q(\theta)\,d\theta,$$

where $\mathbb{E}\left[\cdot\right]_{q(\theta)}$ denotes the expectation with respect to the distribution $q(\Theta)$. The above functions lead to the following iterative algorithm to find $q(\Theta)$. Given the initial estimates $q(r)$, $q(s')$, $q(\lambda)$, $q(\alpha_1)$, $q(\alpha_2)$, $\tilde{q}(\theta)$, $\tilde{q}(s')$, $\tilde{q}(\lambda)$, $\tilde{q}(\alpha_1)$, $\tilde{q}(\alpha_2)$, for $k=1,2$...until a stopping criterion is met in the following algorithm. Lastly repeat the algorithm for $l = 1, 2, \cdots$

Find

$$q^k(r) = \arg\min_{q(r)}$$

$$C_{KL} \left\{ q^k(r) \frac{p(r)}{q^k(r)} \right\}$$

Find

$$q^{k+1}(s') = \arg\min_{q'(s')}$$

$$C_{KL} \left\{ q^{k+1}(s') \frac{p(s')}{q^{k+1}(s')} \right\}$$

Find

$$q^{k+1}(\lambda) = \arg\min_{q(\lambda)}$$

$$C_{KL} \left\{ q^{k+1}(\lambda) \frac{p(\lambda)}{q^{k+1}(\lambda)} \right\}$$

Find

$$q^{k+1}(\alpha_1) = \arg\min_{q(\alpha_1)}$$

$$C_{KL} \left\{ q^{k+1}(\alpha_1) \frac{p(\alpha_1)}{q^{k+1}(\alpha_1)} \right\}$$

Find

$$q^{k+1}(\alpha_2) = \arg\min_{q(\alpha_2)}$$

$$C_{KL} \left\{ q^{k+1}(\alpha_2) \frac{p(\alpha_2)}{q^{k+1}(\alpha_2)} \right\}$$

where the criterion can be assumed to be $\|E(r)q^k(r) - E(r)q^{k+1}(r)\|^2 / \|E(r)q^{k+1}(r)\|^2 \leq \epsilon$, where $\epsilon$ is a stated bound to terminate algorithm. Furthermore, the above equations lead to the following estimations of the observation model and hyperparameters as

$$E(r) = \left(M^k(r)\right)^{\alpha_2} E(s')^r y,$$
\[ \text{cov}^{+1+1}(r) = \begin{bmatrix} M^k(R) \end{bmatrix}^{-1} \] (19)

With
\[ M^k(r) = [E \left( \alpha_s^\alpha \right)^2 + (1 - \lambda) \left( \alpha_s^\alpha \right)^2] \]
\[ + \alpha_s^k E \left( S^k \right)^2 E \left( S^k \right)^T \]
\[ + \alpha_s^k \text{cov}^{+1}(r) \] (20)

\[ E^{+1+1}(s') = \left[ M^k \left( S^k \right) \right]^{-1} \alpha_s^k E^k \left( R \right)^T \]
\[ \text{cov}^{+1+1}(s') = \left[ M^k \left( S^k \right) \right]^{-1} \] (21)

We defined the signal to noise ratio (SNR) for channel \( l \) as:
\[ \text{SNR}^l(0) = 10 \log \left( \frac{\lambda \alpha_s^{-1} E^l(s)}{\alpha_s^{-1}} \right) \] (28)

where \( l = 1, \cdots, d \) is the index of channel \( l \), \( E^l(s) \) is the wavelet energy and \( N \) is the wavelet length: \( E^l(s) = N^{-1} \left( \sum_{i=1}^{N} s_i^2 \right) \).

We generated a 1D reflectivity sequence of 100 sample, shown in Fig. 1, using the BG model with the following parameters ( \( \alpha^{-1} = 1 \), \( \alpha^{-1} = 2 \times 10^{-4}, \lambda = 0.1 \). We then convolved it with four different Ricker wavelet respectively and added Gaussian noise, with SNR of 20 dB. We assume that \( d=4 \) channels are available. The corresponding noisy seismic data is presented in Fig. 2. The estimated reflection coefficients by the proposed variational Bayesian algorithm and the ML algorithms are shown in Fig. 3 and Fig. 4 respectively. Fig. 5 shows the estimated reflection coefficients by single channel variational Bayesian algorithm.

In order to study improvements brought by the different deconvolution method, we consider the following performance indices:
\[ \text{MSE} = \left\| r - \hat{r} \right\|^2, \] (26)

\[ \text{Fig. 1. Synthetic reflectivity.} \]

\[ \text{Fig. 2. Seismic data (SNR=20dB).} \]
where $MSE$ represents the error energy of the estimated wavelet. The $MSE$ for ML algorithm, and variational Bayesian are shown in Fig. 6 with SNR=20, while $MSE$ for single channel and multichannel performance indices are shown in Fig. 7.

We can easily see the results on simulated data that the better recovery for the reflectivity coefficients using the variational Bayesian algorithm compared to ML algorithm. Moreover, we can see the significant advantage for multichannel blind seismic deconvolution than signal channel seismic deconvolution. And the accuracy of the reflectivity recovery will improve with the increase of the channel.

5. Conclusion

In this paper, we propose a new approach for multichannel blind deconvolution using variational Bayesian algorithm related. The poster distribution of the observation model and hyperparameters are given by Kullback–Leibler divergence. An application on simulated data shows the improvement of the proposed algorithm compared to ML algorithm. Furthermore, we can see the recovered reflectivity by multichannel algorithm is better than the single channel one.

Acknowledgement

The authors gratefully acknowledge Teachers Scientific Research Fund of China Earthquake Administration (20130110) for supporting this work.

References


[2]. Ram Idan, Cohen Israel, Raz Shalom, Multichannel Deconvolution of Seismic Signals Using Statistical


