A sequential Bayesian approach for inverting elastic seismic data in the frequency-ray parameter domain

Jinsong Chen, Lawrence Berkeley National Laboratory, Michael G. Hoversten, and Anusha Sekar, Chevron Energy Technology Company*

Summary

Accurately estimating low frequency trends in seismic attributes is very important for seismic imaging and full waveform inversion. In this study, we develop a sequential Bayesian model to invert elastic seismic data in the frequency-ray parameter domain for recovering the low frequency components of geophysical parameters. We first transform time-offset seismic data to the delay time and ray parameter domain through discrete Radom transformation and then to the frequency-ray parameter domain by temporal Fourier transformation. The frequency-p seismic data are divided into several frequency subgroups. We invert them sequentially starting from low to higher frequency subgroups using the results of previous inversion as priors for subsequent inversions. For each frequency subgroup, we can further divide the data into different ray parameter subgroups. We apply the method to a synthetic case that was developed based on real borehole logs. We invert the seismic data generated using frequencies from 3 to 15 Hz for acoustic and shear impedance and density. While simultaneously inverting seismic data at all the frequencies has trouble reaching a stationary distribution, the sequential approach is very effective for finding solutions. Comparison with the true model shows that the developed method can recover the low frequency (0-3 Hz) acoustic and shear impedance in the frequency band below that in the used seismic data (3-15 Hz).

Introduction

Transforming seismic data from the time-offset domain to the tau-p domain is a routine technique to analyze wide-aperture marine seismograms. The tau-p transform (or slant-stacking) is a discretized Radom transform that converts seismic profiles recorded as a function of time and offsets into the domain of vertical delay time, tau, and horizontal wave slowness or ray-parameter, p (Dietrich, 1988; Buland et al., 1996). There are many advantages of using seismic data in the tau-p domain. For example, both seismic reflections and refractions can be investigated simultaneously in the tau-p plane (Stoffa et al., 1981; Diebold and Stoffa, 1981). The relations between the earth model and the corresponding seismic response are far simpler in the tau-p domain (Dietrich, 1988).

However, inversion of seismic data in the time domain often suffers from the nonlinearity associated with high-frequency cycle-skipping artifacts (Brossier et al., 2009). Consequently, seismic inversion is more efficient when it is carried out in the frequency domain. This has been demonstrated by Pratt and Worthington (1990) and Sirgue and Pratt (2004) using Gauss-Newton based methods through successive inversion of single frequency data from low to higher frequencies. The deterministic sequential approach is suitable for estimating multi-resolution geophysical parameters by using selected frequencies (Brossier et al., 2009).

To take advantage of both tau-p transformation and frequency domain inversion, Tao and Sen (2013) applied the temporal Fourier transformation on the vertical delay time and developed a Gauss-Newton based method to invert seismic data in the frequency-ray parameter domain. This method allows inverting not only subgroups of frequencies but also selected ray-parameters and thus gains significant computational efficiency. This makes stochastic inversion of elastic seismic data for large size of problems feasible because stochastic inversion often needs to run forward models tens of thousands of times.

In this study, we use seismic data in the frequency-ray parameter domain as Tao and Sen (2013). Instead of using Gauss-Newton based deterministic methods for inversion, we develop a sequential stochastic approach to inverting seismic data. We divide seismic data into several subgroups and invert them sequentially. We test the developed Bayesian model using a synthetic data set based on logs from the north-west shelf of Australia.

Method

Parameterization

We extend the method developed by Chen and Hoversten (2012) for joint inversion of seismic AVA and CSEM data to invert elastic seismic data in the frequency-ray parameter domain, and adopt the same parameterization as Chen and Hoversten (2012) by dividing the subsurface into many layers. We group those layers into three zones and refer to them as the overburden, reservoir, and bedrock zones, respectively, starting from the air-earth interface. We consider P- and S-wave velocity ratio, P-wave impedance, density, lithotypes, and porosity in each layer of the reservoir zone as unknowns. Layered porosity is linked to P- and S-wave velocity, impedance, and density through a statistical rock physics model derived from geophysical logs at a nearby borehole, which depends on underlying lithotypes. The

spatial dependence of the parameters is carried out by lithotypes, which are assumed to form a Markov random field. Please see Chen and Hoversten (2012) for more details.

For the zones outside the reservoir, we estimate P- and S-wave impedance and density in each layer because we typically do not have data to develop rock physics in those zones. To account for the spatial correlation of those parameters, we add pairwise priors as used by Chen et al. (2012), which are equivalent to weighted averages of values around their neighborhoods. Since this work is built on many previous methods, we only focus on the new development after we replace the seismic convolution model with the frequency-p domain seismic forward simulator.

Elastic seismic data in the frequency-p domain

For a given set of common depth point (CMP) seismic data, we first do discrete Radon transformation (Brysk and McCowan, 1986; Buland et al., 1996) to get the tau-p data and then do temporal Fourier transformation on the tau-p data to get the frequency-p data.

We use the 1D algorithm developed by Chapman (2003) to calculate the elastodynamic response of a stack of plane layers to a plane spectral wave. The algorithm is based on the Kennett reflectivity method. The model propagates plane wave reflectivity through a stack of isotropic elastic layers. Q is added to both reflections and transmissions based on Kjartansson constant-Q model of attenuations, which models Q as a constant with frequency (Kjartansson, 1979). The native domain for computation is the frequency-ray parameter. Although the algorithm is often used to produce a point source response in the time-offset domain via the Hankel transform, we choose to compute elastodynamic responses in the native domain for computational efficiency. The modeling algorithm can also support VTI anisotropy, which we use to model Backus average induced anisotropy in this study. Future work could include inverting for these Thomsen parameters and joint inversion with CSEM data.

The noise characteristics of seismic data in the frequency-p domain play a key role in the inversion because they determine the likelihood function of seismic data. They are less investigated and less understood than those in the time-offset domain.

Sequential Bayesian approach

Suppose we have frequency-p seismic data at various frequencies. We divide the data into n subgroups based on their corresponding frequencies. For example, we let $\mathbf{d}_1$ be the seismic data at the lowest few frequencies, and let $\mathbf{d}_2$ includes the seismic data at the second lowest frequencies. Following the same criterion, we form data vectors $\mathbf{d}_3$, $\mathbf{d}_4$, $\cdots$, and $\mathbf{d}_n$. Let $\boldsymbol{\theta}$ represent all the unknown parameters. The Bayesian model for this inverse problem is given by

$$f(\boldsymbol{\theta} \mid \mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_n) \propto f(\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_n \mid \boldsymbol{\theta}) \times f(\boldsymbol{\theta}). \quad (1)$$

It is reasonable to assume errors associated with the frequency-p data at each frequency are conditionally independent given unknown parameters $\boldsymbol{\theta}$. As shown by Majumdar (1996) and Bonawitz et al. (2011), under the assumption of conditional independence, we have

$$f(\boldsymbol{\theta} \mid \mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_n) \propto \prod_{k=1}^{n} f(\mathbf{d}_k \mid \boldsymbol{\theta}) f(\boldsymbol{\theta}). \quad (2)$$

With the use of equation 2, it is straightforward to show:

$$f(\boldsymbol{\theta} \mid \mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_k) \propto f(\mathbf{d}_k \mid \boldsymbol{\theta}) \times f(\boldsymbol{\theta} \mid \mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_{k-1}). \quad (3)$$

Equation 3 implies that the posterior probability distribution of unknowns given data $\mathbf{d}_1$, $\mathbf{d}_2$, $\cdots$, $\mathbf{d}_k$ is equivalent to the posterior distribution given data $\mathbf{d}_k$ and with the prior being the posterior distribution given data $\mathbf{d}_1$, $\mathbf{d}_2$, $\cdots$, $\mathbf{d}_{k-1}$. This suggests a sequential approach to update unknown parameters, which is given below:

Step 1: Specify prior distribution $f(\boldsymbol{\theta})$ as the initial prior without conditioning to any seismic data. Let $k = 1$.
Step 2: Draw many samples using suitable Markov chain Monte Carlo (MCMC) methods from the posterior distribution defined by

$$f(\boldsymbol{\theta} \mid \mathbf{d}_k) \propto f(\mathbf{d}_k \mid \boldsymbol{\theta}) \times f(\boldsymbol{\theta}). \quad (4)$$

Step 3: Calculate statistics of unknowns using the latter half of the obtained samples, such as medians, and 95% quantiles.
Step 4: Form new prior distribution $f(\boldsymbol{\theta})$ from the above statistics. Suitable expansion of the previously obtained bounds may be needed.
Step 5: If $k > m$, the total number of frequency subgroups, stop; otherwise, let $k = k + 1$ and go to Step 2.

The above sequential procedure is similar to that given by Brossier et al. (2009), but their inner loops are Gauss-Newton based inversion. We use MCMC sampling methods to explore the joint posterior distribution (see Step 2).

Case Study

Synthetic model

We first apply the Bayesian method to a synthetic model generated from borehole logs. Figure 1 shows seismic P- and S-wave velocity and density as functions of normalized depth, where the black curves are original log data and the red line segments are their corresponding Backus averages over 5m thick intervals. The reservoir is in the depth range between the two blue lines. Although our ultimate goal is to estimate reservoir parameters, in the current study, we focus on estimating low frequency components of P- and S-impedance and density over the entire depth range needed for

full waveform inversion of any kind. Estimation of reservoir parameters is ongoing research.

Synthetic data and prior bounds

We generated synthetic data using the 1D algorithm by Chapman (2003) and the wavelet used by Chen and Hoversten (2012), which has the peak frequency of 29 Hz. We generated data by first calculating the true elastodynamic responses for frequencies 3-30 Hz with an increment of 1-Hz and ray-parameters from 0.0 to 0.7 s/km with an increment of 6.9e-3 s/km and then adding random noise that has the signal-to-noise ratio of 10. Figure 2 shows the real and imaginary part of the generated seismic data. We can see that the critical angle zone is in the ray-parameter range between 0.4 and 0.65 s/km.

We derived the prior bounds of unknown parameters from nearby borehole logs. We first fitted the seismic P- and S-wave impedance and density as linear functions of depth. We then subtracted three times of the fitting standard errors from the best fits to get lower bounds and added three times of the fitting standard errors to get upper bounds.

We adopt the following sequential procedure to invert the 5m-layer synthetic seismic data. First, we invert the 3-10 Hz seismic data for ray-parameters less than 0.3 s/km using 100m thick layers. Secondly, we invert the 11-15 Hz seismic data for ray-parameters less than 0.3 s/km using 50m thick layers. We set the prior bounds by expanding the previously obtained posterior bounds by 60%. Finally, we invert the 11-15 Hz seismic data for ray-parameters greater than 0.3 s/km. The prior bounds for the inversion are the posterior bounds from the previous inversion.

In this study, our parameterizations are based on 50m and 100m thick layers, but seismic data are generated from 5m thick layers. This mimics real situations where there are many unknown thin layers. However, this induces anisotropy. Without considering such anisotropy, there are significant discrepancies between the data generated from 5m thick and from 50m or 100m thick layers. In the current study, we obtained the anisotropic Thomsen parameters from the borehole logs and input them to the forward model. The ongoing study will consider them as unknowns.

Inversion results and discussion

We have run many inversions by choosing different combination of frequency and ray-parameter groups and report only the results of using 50m thick layers for frequencies from 3 to 15 Hz (see Figures 3 and 4). To show how well the inversion finds low frequency components of the true model, we compare the posterior estimates with the true 5m-layer model, both filtered using a smooth Tukey filter with a frequency window of 0-5 Hz.

Figure 3 compares the estimated results by using seismic data at frequencies from 3 to 15 Hz and ray-parameters less than 0.3 s/km. The blue and red dashed lines are the initial prior and 95% posterior bounds, respectively. The solid green, blue and red curves are the 5m-thick true model, the filtered true model, and the estimated modes after filtering. For P-wave impedance and density, the estimated modes follow the true model quite well. For S-wave impedance, although the estimated modes are different from the true values, the true values are all within the estimated 95% bounds.

Figure 4 shows the results after further conditioning to seismic data at ray-parameters greater than 0.3 s/km. The inversion significantly improves the estimates of S-wave impedance. It also reduces uncertainty in the estimates of P- and S-wave impedance considerably. Table 1 provides a quantitative comparison between the two inversion results. It is clear that the use of large ray-parameter seismic data is helpful for estimating low frequency components of S-wave impedance.

Conclusions

The case study results show the developed sequential Bayesian approach is effective for inverting seismic data in the frequency-p domain. The technique is able to accurately recover the 0-5 Hz components of the model when data with a low frequency of 3 Hz is used. This is particularly important for gradient-based full waveform inversion because starting from a model (i.e., a low frequency model) that is close enough to the underlying true model can avoid becoming trapped in local minimum.

Acknowledgments

The work is funded by Chevron Energy Technology Company (ETC). We thank Chevron for permission to publish this work. We also thank John Washbourne and Joseph Stefani from Chevron ETC for development of the fast 1D elastic code.

Table 1: Comparison between the estimated results using seismic data at ray-parameter less than 0.3 s/km and greater than 0.3 s/km.

		Ray-parameter <0.3 s/km	Ray-parameter > 0.3 s/km	Reduction (%)
RMS	Zs (MPa)	0.1880	0.1189	36.75
	Zp (MPa)	0.4184	0.3678	12.09
	Density (g/cc)	0.5816	0.6040	-3.85
Half width of 95% intervals	Zs (MPa)	0.9336	0.5871	37.11
	Zp (MPa)	1.9054	1.3297	30.21
	Density (g/cc)	0.3259	0.2293	29.64

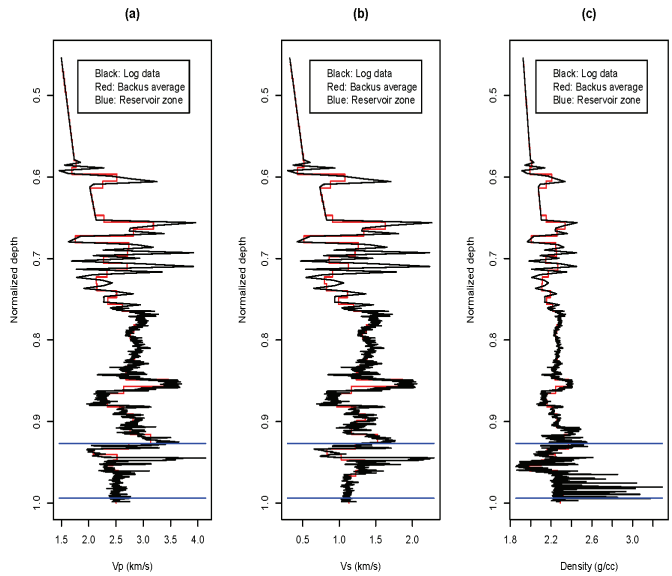

Figure 1: Borehole logs as functions of normalized depth, where the black curves are the original measurements and the red line segments are the Backus average over 5m thick layers.

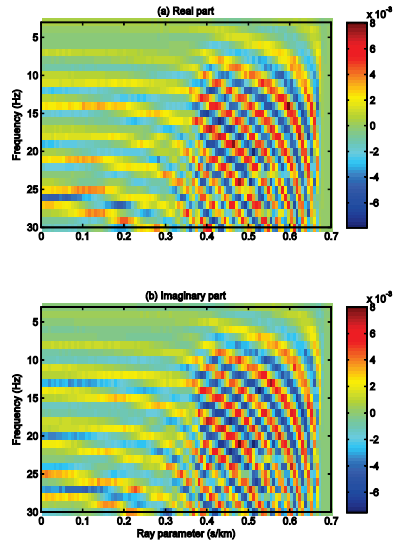

Figure 2: Seismic data in the frequency-ray parameter domain: (a) real part of data, (b) imaginary part of data amplitude.

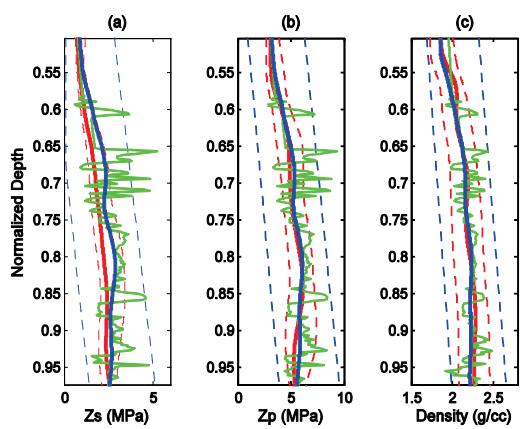

Figure 3: The estimated S-wave and P-wave impedance and density using the seismic data from 3 to 15 Hz and ray-parameters less than 0.3 s/km. The blue and red dashed lines are the initial prior bounds and 95% posterior predictive intervals. The green and blue curves are the 5m thick true models and the filtered true model using a smooth Tukey filter with a frequency window of 0-5 Hz. The solid red lines are the estimated modes after using the same filter. The estimated modes of P-wave impedance and density are close to their corresponding true values.

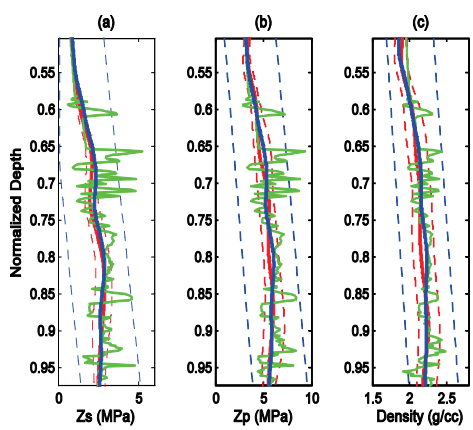

Figure 4: The estimated S-wave and P-wave impedance and density using the seismic data from 3 to 15 Hz and ray-parameters greater than 0.3 s/km. The blue and red dashed lines are the initial prior bounds and 95% posterior predictive intervals. The green and blue curves are the 5m thick true models and the filtered true model using a smooth Tukey filter with a frequency window of 0-5 Hz. The solid red lines are the estimated modes after using the same filter. The estimated modes of P-wave and S-wave impedance and density are close to their corresponding true values.