Simultaneous inversion for velocity and attenuation by waveform tomography

Fengxia Gao a,b, Yanghua Wang b,*

a School of Geophysics and Information Technology, China University of Geosciences (Beijing), 100081, China
b Centre for Reservoir Geophysics, Department of Earth Science and Engineering, Imperial College London, SW7 2BP, UK

A R T I C L E   I N F O

Article history:
Received 9 October 2015
Received in revised form 22 May 2016
Accepted 24 May 2016
Available online 27 May 2016

Keywords:
Attenuation
Waveform tomography
Tuning factor
Full waveform inversion
Simultaneous inversion

A B S T R A C T

Seismic waveform tomography can invert for the velocity and attenuation (Q^-1) variations simultaneously. For this simultaneous inversion, we propose two strategies for waveform tomography. First, we analyze the contributions of the real part and the imaginary part of the gradients, associated with the velocity and attenuation parameters respectively, and determine that the combination of the real part of the gradient subvector for the velocity parameter and the imaginary part of the gradient subvector for the attenuation parameter would produce an optimal inversion result. Second, we attempt to balance the sensitivities of the objective function to the velocity and the attenuation parameters. Considering the magnitude differences between these two-type parameters in the simultaneous inversion, we apply preliminarily a normalization to both the velocity model and the attenuation model. However, for balancing their sensitivities, we further adjust the corresponding model updates using a tuning factor. We determine this tuning parameter adaptively, based on the sensitivities of these two parameters, at each iteration. Numerical tests demonstrate the feasibility and reliability of these two strategies in full waveform inversion.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Seismic waveform tomography is an efficient way to obtain subsurface models, defined by various physical parameters including velocity, density and attenuation, etc. Among these parameters, velocity is the most commonly inverted parameter by waveform tomography (Gauthier et al., 1986; Brossier et al., 2009; Rao et al., 2006; Sourbier et al., 2009; Wang and Rao, 2006, 2009; Kim et al., 2011; Rao et al., 2016). However, other parameters are also important, such as the attenuation parameter Q^-1, which can be used as a lithology and fluid indicator. Therefore, simultaneous inversion for both the velocity and attenuation parameters is necessary for reservoir geophysics.

Waveform tomography for the attenuation Q^-1, or Q directly, can be implemented either in the time or frequency domain. Hicks and Pratt (2001) extracted Q parameter using frequency-domain waveform tomography. Wang (2008, Chapter 12) and Rao and Wang (2009) proposed a strategy to invert for the velocity and the attenuation Q^-1 sequentially in waveform tomography, based on their different sensitivities. Quan and Harris (1997) and Cavalca and Fletcher (2008) derived the attenuation models using ray based tomography. Cheng et al. (2015) estimated both velocity and Q model through viscoacoustic waveform inversion in the time domain. When using waveform tomography to invert for attenuation, there is a variety of parameterizations, such as the imaginary part of the complex-valued slowness or velocity (Pratt et al., 2004; Wang, 2008; Rao and Wang, 2009; Kamei and Pratt, 2013; Rao and Wang, 2015), the inverse square-root of the complex-valued velocity (Hak and Mulder, 2011), and Q^-1 or Q directly (Malinowski et al., 2011). In this study, we invert for Q^-1 in the tomographic inversion.

We implement seismic waveform tomography for the velocity and attenuation parameters in the frequency domain. In frequency-domain waveform tomography, only several distinct frequencies are inverted (Wang and Rao, 2009), and the wave equation can be easily modified to include the attenuation by using a complex-valued velocity. However, when inverting for the velocity and the Q^-1 models simultaneously, there are two fundamental questions we should pay attention to.

First, should the real part or the imaginary part of the gradient subvector be applied to calculate the model perturbation? Although the velocity and the Q^-1 parameters are real valued, the gradient of the objective function with respect to these two parameters are complex-valued. In the velocity-only inversion, only the real part is used to update the velocity model (Wang and Rao, 2009). In two-parameter inversion, the real part of the gradient subvector with respect to the velocity is useful for the velocity model, just like in the case of the velocity-only inversion, but the imaginary part of the gradient subvector with respect to the attenuation parameter can give a better inverted

* Corresponding author.
E-mail address: yanghua.wang@imperial.ac.uk (Y. Wang).

http://dx.doi.org/10.1016/j.jappgeo.2016.05.012
0926-9851/© 2016 Elsevier B.V. All rights reserved.
Q\(^{-1}\) model. This is one of the strategies which will be discussed in this paper in detail.

Secondly, we know that the velocity and Q\(^{-1}\) parameters have different physical units and magnitudes, but how to effectively balance the differences and improve both parameters in a simultaneous inversion? To take care of the magnitude difference in the inversion, it is necessary to have a normalization. But a more critical issue is the difference of sensitivity to these parameters. In this paper, we propose a tuning parameter to balance the updates in the velocity model and the Q\(^{-1}\) model, and show how to properly choose such a tuning parameter in an iterative inversion.

The paper is arranged in the following sections. Section 2 summarizes briefly the frequency-domain waveform inversion theory and provides explicit expressions of gradient subvectors we concerned. Section 3 analyzes the sensitivities of the objective function with respect to the velocity and attenuation parameters. Section 4 focuses on up-scaling the updates for the Q\(^{-1}\) model. Finally, Section 5 presents the numerical experiments and a discussion on the results.

2. Frequency-domain waveform tomography

The 2D viscoacoustic wave equation in the frequency domain can be expressed as (Malinowski et al., 2011)

\[
\left(\frac{\omega^2}{K(x,z)} - \frac{\partial}{\partial x}\left(\frac{1}{\rho(x,z)} \frac{\partial}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{1}{\rho(x,z)} \frac{\partial}{\partial z}\right)\right)P(x,z,\omega) = F(x,z,\omega),
\]

where \(P(x,z,\omega)\) is the frequency domain seismic wavefield, \(\rho(x,z)\) is density, \(F(x,z,\omega)\) is the frequency domain source, \(K(x,z)\) is the complex bulk modulus,

\[
K(x,z) = \rho(x,z)\ddot{v}(x,z),
\]

where \(\ddot{v}(x,z)\) is a complex velocity, expressed as (Blanch et al., 1995; Wang, 2008)

\[
\ddot{v}(x,z) = \nu(x,z) \left(1 - \frac{i}{2Q(x,z)}\right),
\]

and \(Q\) is the quality factor, which is assumed to be frequency independent (Wang and Guo, 2004).

\[
\nu(x,z) = \nu(x,z) \exp(-\alpha x) \exp\left[i(k_b - \omega t)\right], \quad \alpha = \frac{\omega}{2vQ}.
\]

The frequency-domain wave Eq. (1) can be represented in a matrix-vector form as

\[
AP = F,
\]

where \(A\) is the complex-valued matrix which is a function of frequencies and model properties, \(\rho(x,z)\) and \(\nu(x,z)\) (which includes \(Q^{-1}(x,z)\)), \(P\) is the frequency domain wavefield vector and \(F\) is the frequency domain source vector.

For waveform tomography, the objective function is set as

\[
J = \frac{1}{2} A^\dagger \phi(m) A \phi(m),
\]

where \(\phi(m)\) is the misfit function, and can be expressed as

\[
J = \frac{\partial P(m)}{\partial m} \partial P(m).
\]

\[
\frac{\partial P(m)}{\partial m} = \frac{\partial P(m)}{\partial a} \frac{\partial a}{\partial m} + \frac{\partial P(m)}{\partial b} \frac{\partial b}{\partial m},
\]

where \(\phi(m)\) is the misfit function, and can be expressed as

\[
\phi(m) = \frac{1}{2} A^\dagger \phi(m) A \phi(m),
\]

where \(J\) is the Fréchet derivative, and can be expressed as

\[
J = \frac{\partial P(m)}{\partial m} A^\dagger \phi(m). \quad \text{(7)}
\]

This Fréchet matrix can be worked out as the following (Pratt et al., 1998; Ravaut et al., 2004). Taking the first-order derivative to Eq. (4), with respect to a parameter \(m_i\),

\[
A \frac{\partial A}{\partial m_i} + \frac{\partial A}{\partial m_i} P = 0, \quad \text{we obtain}
\]

\[
\frac{\partial P}{\partial m_i} = -A^{-1} \frac{\partial A}{\partial m_i} P. \quad \text{(9)}
\]

Then the element of the gradient can be expressed as

\[
\frac{\partial \phi}{\partial m_i} = P^\dagger A^\dagger \phi \left(\frac{\partial A}{\partial m_i}\right) A P \phi(m).
\]

3. Sensitivity analysis

To see the sensitivity difference between the velocity and the Q\(^{-1}\) parameter, let us first examine the variation of the misfit function with respect to the variation of these two parameters.
The models used in the test are shown in Fig. 1. In Fig. 1(a), the background velocity is 5000 m/s and the velocity of the square anomaly is 4000 m/s. The counterparts for the Q\(^{-1}\) model are 0.02 and 0.025, respectively, in Fig. 1(b). In the test, only the value of the anomaly is changed (Malinowski et al., 2011).

The model size in Fig. 1 is 40 × 40 with the grid interval of 25 m. Only one shot is applied in the test, which is located in the middle of the model and at the depth of 25 m. There are 40 receivers set at the surface with the space interval of 25 m.

The relative variation range for velocity and Q\(^{-1}\) is -50% to 100%. The relative variation is calculated by

\[
\text{var} = \frac{m - m_{\text{true}}}{m_{\text{true}}} \times 100\%
\]

where \(m_{\text{true}}\) is the true value, and \(m\) is the perturbed value.

As shown in Fig. 2(a), the value of the misfit function will be bigger when the difference between the true velocity model and the varied velocity model increases. Fig. 2(b) is the misfit function versus Q\(^{-1}\). The misfit function varies slowly when Q\(^{-1}\) is smaller than the true value, but varies greatly when Q\(^{-1}\) is bigger than the true value.

The magnitude of the misfit function with velocity variation is bigger than the one with Q\(^{-1}\) variation by around 3 magnitude orders. For example, the magnitude difference is about 3 orders at variation of 50%. Therefore, in the simultaneous inversion for the velocity and the Q\(^{-1}\) models by minimizing the objective function in Eq. (5), the velocity parameter is much more sensitive than the Q\(^{-1}\) parameter. That is, the velocity model can be easily updated, in comparison to the Q\(^{-1}\) reconstruction.

4. Upscaling the update for the Q\(^{-1}\) model

In order to balance the difference between the two parameters, a straightforward treatment is normalization. This is because the velocity parameter and the Q\(^{-1}\) parameter have different magnitudes. We normalize both the velocity and Q\(^{-1}\) parameters by

\[
\mathbf{m} = \left[ \frac{v}{Q^{-1}} \right] = \left[ \begin{array}{c} \frac{v - v_{\text{min}}}{v_{\text{max}} - v_{\text{min}}} \\ \frac{Q^{-1} - Q_{\text{min}}}{Q_{\text{max}} - Q_{\text{min}}} \end{array} \right].
\]

where \(\mathbf{m}\) is the normalized model vector. In the following tests, we set \(v_{\text{min}} = 1000\) m/s, \(v_{\text{max}} = 6000\) m/s, \(Q_{\text{min}}^{-1} = 0\), and \(Q_{\text{max}}^{-1} = 0.2\). Now we conduct analysis on the gradients of normalized models.

We use an overthrust model (Fig. 3) for the following tests. Fig. 3(a) is the velocity model, and Fig. 3(b) is a Q\(^{-1}\) model, where the Q model values are empirically calculated from the velocity model:

\[
Q = 21.86 \times \exp \left( -40.12 \left( \ln(V) \right) -49.89. \right)
\]

This unpublished relationship is derived by authors from field data examples. Note that no relationship between these two models is assumed during the inversion.

The model size is 300 × 75 with the grid interval of 25 m. There are 151 shots located at a depth of 75 m with the space interval of 50 m. 151 receivers are set at the surface at a spatial interval of 50 m. A Ricker wavelet with the peak frequency of 15 Hz is used as the source wavelet.

For iterative inversion, we set the initial models by simply smoothing the true models, by low-pass filtering, as shown in Fig. 3(c–d). Fig. 4 shows the gradients of the objective function with respect to velocity and Q\(^{-1}\) parameters at the first iteration when the frequency used is 3.72 Hz.

Fig. 4a shows the real part of the gradient subvector of the velocity model. Fig. 4b is the imaginary part of the gradient subvector of the Q\(^{-1}\) model. Both are calculated based on the initial models. Recalled that the velocity and attenuation models are normalized. Fig. 4 reveals that the magnitude difference between the two gradient subvectors is about 3 orders. If using the same step length for these two parameters, the model update might be too small to have any moderate change in the Q\(^{-1}\) model, and thus the inverted Q\(^{-1}\) model will be nearly the same as the initial model.

In order to balance the updates to the normalized models, we can simply amplify the gradient subvector of the Q\(^{-1}\) model by a positive tuning factor:

\[
\frac{\partial \phi}{\partial \mathbf{Q}_{\text{tune}}} = \left[ \begin{array}{c} 1 & 0 \\ 0 & \lambda \end{array} \right] \frac{\partial \phi}{\partial \mathbf{Q}}
\]

where \(\partial \phi/\partial \mathbf{V}\) is the gradient for normalized velocity parameter, \(\partial \phi/\partial \mathbf{Q}^{-1}\) and \(\partial \phi/\partial \mathbf{Q}_{\text{tune}}^{-1}\) are the gradients for normalized Q\(^{-1}\) parameter before and after using the tuning factor \(\lambda\). This tuning parameter
upscales the gradient subvector for the $Q^{-1}$ model and is estimated by

$$
\lambda = \left| \frac{\partial \phi}{\partial v} \right|^{-1} \left| \frac{\partial \phi}{\partial Q^{-1}} \right|^{-1},
$$

where $\left| \frac{\partial \phi}{\partial v} \right|$ is the L2-norm of the gradient subvector for normalized velocity parameter, and $\left| \frac{\partial \phi}{\partial Q^{-1}} \right|$ is the L2-norm of the gradient subvector for normalized $Q^{-1}$ parameter.

This quantitative tuning factor is similar to a weighting matrix related to different model parameters suggested by Wang (1999), in which a weighting matrix was defined by the ratio of Frechet sub-matrices, with respect to various classes of model parameters. Since the $\lambda$ value in Eq. (18) is estimated in each iteration, the proposed method is an adaptive selection method.

The model updates are obtained by using the memoryless quasi-Newton method, in which only the gradient and the updated model from the previous iteration are used to construct the approximate Hessian matrix theoretically. However, it does not calculate the inverse Hessian matrix, since the model update is obtained directly from the previously updated model and the gradients of the previous and current iterations (Liu et al., 2015). With the introduction of the tuning factor $\lambda$, the model update is calculated by using the tuned gradient for the normalized $Q^{-1}$ parameter. And then the models are updated by

$$
\begin{bmatrix}
\tilde{v}^{(k+1)} \\
\tilde{Q}^{(k+1)}
\end{bmatrix}
= \begin{bmatrix}
\tilde{v}^{(k)} \\
\tilde{Q}^{(k)}
\end{bmatrix}
+ \alpha_k \begin{bmatrix}
\Delta \tilde{v} \\
\Delta \tilde{Q}^{-1}
\end{bmatrix},
$$

where $\Delta \tilde{v}$ and $\Delta \tilde{Q}^{-1}$ are the model updates for normalized velocity and normalized $Q^{-1}$, respectively, $\alpha_k$ is the step length in the $k$th iteration for both parameters.

As for the step length $\alpha_k$, it is calculated by a parabolic line-search method (Ravaut et al., 2004; Sourbier et al., 2009). In this method, three step lengths, 0, $\alpha_1$, and $\alpha_2$, are selected first. The corresponding misfit functions to these three step lengths are required to satisfy $\phi(0) > \phi(\alpha_1) > \phi(\alpha_2)$. Secondly, a parabolic curve can be constructed with the three step length and misfit function pairs. Then we can find the smallest misfit function value on the parabolic curve, and the corresponding step length is taken as the best step length in current iteration.
In each iteration, the updated physical models are

\[ v^{(k+1)} = \bar{v}^{(k+1)} (v_{\text{max}} - v_{\text{min}}) + v_{\text{min}}, \]  

\[ Q^{-1(k+1)} = \bar{Q}^{-1(k+1)} (Q_{\text{max}}^{-1} - Q_{\text{min}}^{-1}) + Q_{\text{min}}^{-1}. \]  

5. Model tests

Now let us test the real part and the imaginary part of the gradient subvector for \( Q^{-1} \) used under the \( \lambda \) adaptively selecting scheme.

According to Eq. (14), for the velocity update, we use the real part of the gradient, associated with the velocity parameter, and for the \( Q^{-1} \) update, we use the imaginary part of the gradient, associated with the \( Q^{-1} \) parameter. The gradient vector of the objective function, with respect to the normalized parameters, is

\[ \frac{\partial \phi}{\partial m} = \begin{bmatrix} \text{Re} \left\{ -\left( \frac{\partial P}{\partial v} \right)^{H} \delta P (v_{\text{max}} - v_{\text{min}}) \right\} \\ \text{Im} \left\{ -\left( \frac{\partial P}{\partial Q^{-1}} \right)^{H} \delta P (Q_{\text{max}}^{-1} - Q_{\text{min}}^{-1}) \right\} \end{bmatrix}. \]  

To verify the scheme above, we also test the case when we use the real part of both sub-vectors of the gradient in model updating.

In the test, the inversion uses seismic data sequentially from 11 frequencies (3.72, 5.68, 7.65, 9.61, 11.57, 13.53, 15.49, 17.45, 19.41, 21.37, 23.33 Hz). The frequency selection strategy followed Wang and Rao (2009)'s method. For a single frequency, 20 iterations are carried out.

To evaluate the quality of the inverted tomograms, we estimate the relative errors between the inverted models and the true ones by

\[ \text{error} = \frac{1}{\sqrt{N}} \frac{m_{\text{inv}} - m_{\text{true}}}{m_{\text{true}}}. \]  

Table 1

<table>
<thead>
<tr>
<th>Gradient of ( Q^{-1} )</th>
<th>Imaginary</th>
<th>Real</th>
</tr>
</thead>
<tbody>
<tr>
<td>The errors in ( v )</td>
<td>0.0669</td>
<td>0.0669</td>
</tr>
<tr>
<td>The errors in ( Q^{-1} )</td>
<td>0.0842</td>
<td>0.2654</td>
</tr>
</tbody>
</table>
where $\mathbf{m}_{\text{inv}}$ represents the inverted model, $\mathbf{m}_{\text{true}}$ is the true model and $N$ is the total number of model grid points.

Fig. 5(a–b) display the inverted velocity model and the inverted $Q^{-1}$ model, respectively, with the imaginary part of the gradient for normalized $Q^{-1}$ used. As shown in Table 1, the relative errors between the inverted model and the true model for velocity and $Q^{-1}$ parameters are 0.0669 and 0.0842, respectively. Also from Fig. 5a–b, it can be seen that both the inverted velocity $Q^{-1}$ models are in good accordance with the true models not only in structure but also in model values.

Fig. 6 plots the values of each iteration, and the average of $\lambda$ values over these iterations is 158.29. We also test the case when $\lambda$ is set equal to 158.29 at all iterations, and the inversion results are close to those using the adaptive $\lambda$ selection method. This implies that the $\lambda$ adaptive selection method is stable.

Fig. 7(a–b) are the inversion result, using the real part of the gradient for the normalized $Q^{-1}$ parameter and the adaptive $\lambda$ values. The relative errors for velocity and $Q^{-1}$ parameters are 0.0649 and 0.2654, respectively, as listed in Table 1. From Fig. 7(a–b), it’s clear that the inverted $Q^{-1}$ model is smeared severely at the shallow part.

Fig. 8 is the $\lambda$ curve for all iterations. The average of the $\lambda$ values is 161.13, which is close to the average $\lambda$ value in the imaginary part of the gradient for $Q^{-1}$ used case. This is because in both cases, the complex gradients for velocity and $Q^{-1}$ parameters are used to calculate $\lambda$.

The relative errors listed in Table 1 indicate that the result using the imaginary part of $Q^{-1}$ gradient subvector is better than the other.

6. Conclusions

In the simultaneous inversion for the velocity and $Q^{-1}$ models by waveform tomography, we have proposed two strategies. First, the combination of the real part of the gradient subvector for the velocity and the imaginary part of the gradient subvector for $Q^{-1}$ will produce a better inversion result. Second, considering the magnitude and sensitivity differences between the two-type parameters, after preliminary normalization, the tuning factor will further balance the sensitivity difference between the velocity and the $Q^{-1}$ parameter. The tuning factor is selected adaptively in each iteration. Jointly using these strategies, we are able to improve the inversion of the velocity model and the $Q^{-1}$ model, in the simultaneous inversion.

Acknowledgments

The authors are grateful to the sponsors of the Centre for Reservoir Geophysics, Imperial College London, for supporting this research.

References


Cavalcá, M., Fletcher, R., 2008. Deriving 3D Q models from surface seismic data using attenuated traveltine tomography. 70th EAGE conference and Exhibition, paper H003.


