Preconditioned transmission + reflection joint traveltime tomography with adjoint-state method for subsurface velocity model building

Introduction

Inverting the subsurface velocity structure is one of the core issues in seismology. Traveltime tomography, which has high reliability and strong stability, has been widely used in velocity model building. The adjoint-state traveltime tomography (AST) (Sei and Symes, 1994; Leung and Qian, 2006; Taillard et al., 2009) is an important technique among the traveltime inversion methods. This method uses the traveltime field calculated by eikonal equation and the adjoint field calculated by the adjoint equation to generate the gradient of the objective function, avoiding the ray tracing and the calculation of Fréchet derivative matrix. Simultaneously, it has extremely low computational memory requirement and is particularly suitable for parallel computing (Noble et al., 2010; Benaichouche et al., 2015). Due to those advantages, it has gained a lot of attention in the field of velocity model building recently (Huang and Bellefleur, 2012; Li et al., 2014; Waheed et al., 2016; Li et al., 2017; Han and Dong, 2019).

In the early stages, AST refers to the adjoint-state transmitted first arrival traveltime tomography. In the last few years, it has been extended to the adjoint-state reflected traveltime tomography and adjoint-state joint transmitted and reflected traveltime tomography (ASJT) (Huang and Bellefleur, 2012; Li et al., 2014). However, there will be singular values at the source and the receiver positions in the gradient when using the conventional AST/ASJT method directly, which will slow down the inversion convergence. Reasonable utilization of Hessian information can accelerate convergence in AST (Breutadeau et al., 2014; Benaichouche et al., 2015). How to precondition the ASJT is worth studying.

In this paper, we utilize the diagonal elements of approximate Hessian as preconditioners for ASJT to increase the inversion accuracy and convergence rate. The corresponding preconditioned ASJT method is referred to as PASJT. Two numerical experiments show that the precondition method eliminates the singular values in the original gradient and ameliorates the characteristics of gradient, thereby improving the accuracy of the inversion result and accelerating the convergence rate.

Theory and Method

The eikonal equation can be expressed as:

\[ \begin{align*}
\left| \nabla t_r(x) \right|^2 &= \frac{1}{v(x)} \quad x \in \Omega / \Gamma_s, \\
\nabla t_r(x_s) &= \nabla t_r(x), \\
\end{align*} \]

where \( \Omega \) denotes the computational domain, \( \Gamma_s \) denotes the seismic source locations, which is a subset of \( \Omega \). \( v(x) \) denotes the seismic wave velocity at the spatial position \( x \). \( t_r(x_s) \) denotes the priori traveltime located at the source point \( x_s \), generally, it is equal to \( 0 \). \( t_r \) denotes the traveltime field of the transmitted wave. \( \nabla \) denotes gradient operator and \( \| \| \) denotes modulo length operation. If the reflected traveltime field also needs to be calculated, an additional equation is needed as below

\[ \begin{align*}
\left| \nabla t_p(x) \right|^2 &= \frac{1}{v(x)} \quad x \in \Omega / \Gamma_r, \\
\nabla t_p(x_s) &= \nabla t_p(x), \\
\end{align*} \]

where, \( t_p \) denotes the traveltime field of the reflected wave. \( \Gamma_r \) denotes the area where the reflectors (secondary seismic sources) are located, and \( x_s \) is the position of the reflection point (scattering point) in this area. According to Huygens principle, as well as the traveltime field on the reflectors is recorded as the secondary source, the traveltime field of the reflected wave can be calculated using equation (2).

According to AST method with independence of surface normal vectors (Han and Dong, 2019), the ASJT objective function can be defined as

\[ J(v) = J_T(v) + J_R(v) = \frac{1}{2} \int_\Omega \omega_1 \left( t_r(x) - t^{obs}_r(x) \right)^2 \delta_{x,s} + \frac{1}{2} \int_\Omega \omega_2 \left( t_p(x) - t^{obs}_p(x) \right)^2 \delta_{x,s}, \]

where \( t^{obs}_r \) and \( t^{obs}_r \) denote the observed transmitted and reflected traveltime, respectively. \( \omega_1 \) and \( \omega_2 \).
denote the weighting factors of the observed transmitted and reflected traveltime respectively, and in this paper we set both \(\omega_1\) and \(\omega_2\) being equal to 0.5. The restriction operator \(\delta_{x,x}\) controls the objective function at the position of receivers \(g\). \(J_f\) and \(J_h\) are the transmitted and reflected part of the joint objective function \(J\) with respect to velocity. \(J\) is iteratively minimized using the following relation

\[
v_{n+1} = v_n - \alpha \left( \omega_1 \cdot \text{diag}(H_f)^{-1} \nabla J_f + \omega_2 \cdot \text{diag}(H_h)^{-1} \nabla J_h \right),
\]

where \(\alpha\) is the step length, \(\text{diag}(H_f)\) and \(\text{diag}(H_h)\) are the diagonals of the approximate Hessian. 
\[
\nabla J_f = \partial J_f / \partial v(x) = - \lambda_f(x) / v^*(x) \quad \text{and} \quad \nabla J_h = \partial J_h / \partial v(x) = - \lambda_h(x) / v^*(x)
\]

are the gradients of the transmitted and reflected objective function, which can be calculated by the adjoint-state method, with the corresponding adjoint equation

\[
\begin{align*}
\nabla \cdot (\lambda_f(x) \nabla t_f) &= (t_f(x) - t_{\text{true}}(x)) \delta_{x,g} \\
\nabla \cdot (\lambda_h(x) \nabla t_h) &= (t_h(x) - t_{\text{true}}(x)) \delta_{x,g} \\
\nabla \cdot (\lambda_{h}^s(x) \nabla t_1) &= \lambda_{h}^s(x) \delta_{x,x}
\end{align*}
\]

where \(\lambda_f\) and \(\lambda_h\) denote the transmitted and reflected adjoint-state variables with \(\lambda_h = \lambda_h^s + \lambda_h^r\) (\(\lambda_h^s\) and \(\lambda_h^r\) are the receiver-side and the shot-side reflected adjoint-state variables), and \(\delta_{x,x}\) controls the equation at the position of reflector \(x\).

Finally, referring to Benaichouche (2015), we divide the original diagonal of the Hessian \(\text{diag}(H)\) into \(\text{diag}(H_f)\) and \(\text{diag}(H_h)\) as preconditioners for transmitted and reflected gradients. These two preconditioners can be estimated by sequentially calculating the following additional adjoint equation (6) that is similar to equation (5).

\[
\begin{align*}
\nabla \cdot (\lambda_f(x) \nabla t_f) &= t_{\text{true}}(x) \delta_{x,g} \\
\nabla \cdot (\lambda_h(x) \nabla t_h) &= t_{\text{true}}(x) \delta_{x,g} \\
\nabla \cdot (\lambda_{h}^s(x) \nabla t_1) &= \lambda_{h}^s(x) \delta_{x,x}
\end{align*}
\]

where \(t_{\text{true}}\) is a constant value, which is equal to 1 in this paper. The preconditioners can also be explained in the framework of the traveltime tomography with scattering integral method (Li et al., 2017). Physically, these preconditioners remove the influence of ray density and singularities, hence accelerate the convergence rate.

**Numerical Experiments**

Two numerical examples are used to demonstrate the effectiveness of our PASJT method. The first example is an inclusion model (Fig.1a) with a constant velocity gradient model as the background velocity model (Fig.1b). The grids number of the model is 401*101, and the spacing of horizontal and vertical directions is 20 m. In this test, the reflector is artificially set at the bottom of the model.

*Figure 1 (a) True inclusion velocity model and (b) initial velocity model.*

Figure 2 shows the gradient (Fig.2a) and the preconditioned gradient (Fig.2b) during the inversion using the single-shot-receiver pair geometry system (The shot point and the receiver point are located at 2 km and 6 km on the surface, respectively). It can be clearly seen that in the vicinity of the shot and the
receiver points, the gradient value is significantly high (Fig.2a The red box), we call it gradient singularity. After preconditioning, the singular values in the gradient are effectively eliminated (Fig.2b).

**Figure 2** The first-iteration gradient of single-shot-receiver pair inversion for (a) ASJT (b) PASJT.

Figure 3 show the normalized gradients of ASJT (Fig.3a) and PASJT (Fig.3b) after multi-shot stacking. In this experimental process, 101 shots are evenly distributed along the model surface with an interval of 80m, and 401 receivers for each shot are distributed at each grid point along the top surface. It can be seen that the gradient of ASJT (Fig.3a) is stretched from the position of the inclusion velocity anomaly (X=4 km, Z=1 km) to top surface, even some large values formed at 1-2 km and 6-7 km near the surface and at 3-5 km on the bottom position of the reflector.

By imposing preconditioners, the large gradient values of PASJT (Fig.3b) are mainly concentrated in the vicinity of the inclusion velocity anomaly, which indicates that the preconditioners can modify the original gradient effectively.

**Figure 3** The first-iteration gradients of multi-shot-receiver pair joint inversion for (a) ASJT and (b) PASJT.

To further illustrate the effectiveness of the preconditioners, we employ ASJT and PASJT to the complex foothill model (Fig.4a), with the constant velocity gradient model as the initial model (Fig.4b). The grid number of the model is 1255*300, and the spacing of horizontal and vertical directions is 10 m. The reflector is set at the bottom of the model. 126 shots and 1255 receivers for each shot are located on the surface. For the convenience of comparison, we only show the shallow part of the model.

**Figure 4** Foothill model experiments. (a) true model, (b) constant gradient initial model, inversion result of (c) ASJT (d) PASJT.

Figure 4c and Figure 4d show the results of ASJT and PASJT, respectively. It should be noted that although there will be singular values in the first-iteration gradient, ASJT can still get a general result (Fig.4c) with enough iterations in which the gradient values were corrected iteratively. This is why the traditional ASJT succeed but is not efficient enough.
Figure 5 shows the velocity profiles of ASJT (Fig.4c) and PASJT (Fig.4d) results at 200m below the surface. It can be seen that the PASJT result (Fig.5, The blue line ) is significantly closer to the true model (Fig.5, The black dotted line ). Moreover, the convergence speed of the PASJT (Fig.6, The solid line) is faster than that of the ASJT (Fig.6, The dotted line). Evidently, compared to ASJT, PASJT not only has higher inversion accuracy, but also has higher convergence.

**Figure 5** The velocity profiles of ASJT (the red line) and PASJT (the blue line) at 200m below the surface.

**Figure 6** Decrease of logarithmic objective function with iterations for ASJT (the dotted line) and PASJT (the solid line).

### Conclusions

Compared with the traditional ASJT method, the proposed PASJT is easy to apply by solving an additional adjoint equation. In addition, the preconditioned ASJT method have the following three advantages: (1) The gradient can be preconditioned at the cost of a little extra computational memory. (2) The PASJT method makes the characteristics of gradient better in inversion, thereby improving the accuracy of inversion results; (3) PASJT method accelerate the convergence rate.

### Acknowledgements

This research is financially supported by the National Natural Science Foundation of China (grant nos. 41874127) and the National Key R&D Program of China (grant no. 2019YFC0312004).

### References


