Introduction

The least-squares inverse strategy has been considered as an important method in estimating the subsurface parameters (e.g., full waveform inversion (FWI)) and building the reflectivity profile (e.g., least-squares reverse time migration (LSRTM)) through an iterative process. While the gradient based method for FWI or LSRTM is regarded as a crude scaling strategy and it suffers from slow convergence rate. The gradient can be considerably enhanced by multiplying the inverse Hessian, which is extremely expensive to be calculated directly.

The approximate Hessian in Gauss-Newton method, which ignores the nonlinear term in full Hessian matrix (Pratt et al., 1998), is still too expensive to be calculated directly. Shin et al. (2001) proposed to use the pseudo-Hessian matrix as a substitution for the approximate Hessian. While the diagonal pseudo-Hessian is limited to balance the amplitudes, for ignoring the receiver-side Green’s functions. When considering the reciprocity theory, the receiver-side Green’s functions can be replaced by the source-side Green’s functions under the assumption that the sources and receivers are collocated (Plessix and Mulder, 2006), which forms the double illumination strategy. The receiver-side Green’s functions can also be constructed using the phase encoding method (Tang, 2009) which can reduce the computational burden significantly. In this research, we illustrated a numerical example based on a 50 × 50 homogeneous model to compare different Hessian approximations and presented the gradients scaled by different Hessian approximations based on the Marmousi model.

General principle of the least-squares inverse problem

As a least-squares local optimization, FWI seeks to minimize the difference between the synthetic data and observed data (Lailly, 1983; Tarantola, 1984) and update the model iteratively. The misfit function $\phi$ is given in a least-squares norm. The model perturbation in each iteration consists of the gradient and update the model iteratively. The misfit function $\phi$ makes the inverse problem approximately linear. Thus the full Hessian matrix can be substituted by the approximate Hessian (the second term in equation (2)), used in the Gauss-Newton method:

$$ H^{(n)}(\mathbf{r}, \mathbf{r}) = H_1^{(n)} + H_2^{(n)}, $$

And the first term of full Hessian matrix $H_1^{(n)}$ is always neglected for computation convenience, which makes the inverse problem approximately linear. Thus the full Hessian matrix can be substituted by the approximate Hessian (the second term in equation (2)), used in the Gauss-Newton method:

$$ H_2^{(n)} = \sum_{\mathbf{r}_g, \mathbf{r}_s} \int d\omega \mathcal{R} \{ \omega^4 G(\mathbf{r}', \mathbf{r}_g, \omega) G^*(\mathbf{r}', \mathbf{r}_g, \omega) G(\mathbf{r}', \mathbf{r}_g, \omega) G^*(\mathbf{r}', \mathbf{r}_g, \omega) \}, $$

Under the assumption of the infinite receiver coverage, the receiver-side Green’s functions can be approximated as $d^2(\mathbf{r}_g, \mathbf{r}')$, where $d(\mathbf{r}_g, \mathbf{r}')$ means the distance from position $\mathbf{r}'$ to the receiver position $\mathbf{r}_g$. When the depth in vertical is quite smaller than the distance in horizontal, $d^2(\mathbf{r}_g, \mathbf{r}')$ can be approximated as a constant scalar $L$. This approximation to Hessian matrix is equivalent to the pseudo-Hessian proposed by Shin et al. (2001). The pseudo-Hessian is constructed using the forward modeling wavefields,
used as the virtual sources in reverse rime migration.

\[ H_{pseudo}^{(n)} = \sum_{r_s} \int d\omega \omega^4 \Re \{ G(r'', r_s, \omega) G^*(r', r_s, \omega) \}, \quad (4) \]

The diagonal part of the pseudo-Hessian can be formed when \( r' = r'' \).

For finite receiver coverage, the pseudo-Hessian is limited to compensate the geometrical spreading effects and balance the biased amplitudes, for missing the receiver-side Green’s functions. By introducing the reciprocity theory, the receiver-side Green’s functions can be constructed by the reciprocal wavefields from the receiver. In this condition, if the receivers and the sources are collocated, the reciprocal wavefields can be substituted by the forward modeling wavefields, which gives the approximation of the diagonal Hessian as (Plessix and Mulder, 2006):

\[ H_{double} = \sum_{r_s} \int d\omega \Re \{ \omega^4 \| G(r, r_s, \omega) \|^2 \| G^*(r, r_s, \omega) \|^2 \}, \quad (5) \]

So, we can notice that preconditioning the gradient using \( H_{double} \) is equivalent to applying a double illumination compensation to the gradient.

The phase encoding technique was firstly introduced for imaging (Romero et al., 2000) and then applied for calculating the gradient in FWI. We can also employ phase encoding technique to construct the Hessian matrix, which can reduce the computational cost considerably but unfortunately involves strong crosstalk artifacts. The crosstalk artifacts can be reduced effectively by slant stacking over sufficient ray parameters. In this research, we used a chirp phase encoding strategy, which is a combination of the linear phase encoding and random phase encoding methods, to calculate the receiver-side Green’s functions. Hence, the receiver-side phase encoded Hessian can be expressed as:

\[ H_{chirp} = \sum_{r_s, p} \int d\omega \Re \{ \omega^4 G(r', r_s, \omega) G^*(r', r_s, \omega) G(r'', r_g, \omega) G^*(r'', r_g, \omega) e^{i\omega(p_r + \epsilon \Delta p)(x_g' - x_g)} \}, \quad (6) \]

**Numerical examples**

In this section, two numerical examples are illustrated for analysis and comparison. The first numerical example is a homogeneous model with a constant background velocity. Different Hessian approximations are presented for comparison. Then we practice the scaling methods on a 2D Marmousi model to examine the effects of the preconditioners.

Figure 1 shows the homogeneous model with a constant velocity of 2500m/s used to calculate the Hessian matrix. The model consists of 50x50 grid cells with 5m horizontal and vertical grid intervals. 9 sources are arranged from 25m to 225m with a spacing of 125m and 50 receivers are distributed from...
Figure 2 The Hessian approximations constructed with 9 sources and 50 receivers.

0m to 250m with spacing of 25m. Figure 2 shows the exact approximate Hessian and its approximations for this homogeneous model. Figure 2a is the exact approximate Hessian. As we can see that the Hessian is dominated by the diagonal band. Figure 2b is the pseudo-Hessian, the energy of which is less concentrated on the diagonal band parts. Figure 2c is formed based on the double illumination method. Figure 2e and f are the linear and chirp phase encoded Hessian respectively and the ray parameters range from -0.3s/km to 0.3s/km with a step of 0.1s/km. We can see that these three approximations are very close to the exact approximate Hessian. But the double illumination method relies more on the source and receiver distribution. Figure 2d is the Hessian approximation contaminated by crosstalk artifacts and we can notice that the energy is not distributed regularly caused by the crosstalk artifacts. The second numerical example is based on a modified Marmousi model. Figure 3a shows the exact P-wave velocity model and Figure 3b shows the smoothed P-wave velocity. Figure 3c is the true reflectivity profile obtained by following Plessix and Mulder (2006). Comparing with the true reflectivity, the gradient shown in Figure 3d is poorly scaled. The amplitudes for the deep reflectors are very weak, while too much energy is concentrated on the shallow part of the image or gradient. Figure 4a shows the diagonal pseudo-Hessian which overestimates the illumination energy. Figure 4b is the diagonal Hessian approximation by the double illumination method, which is a very good approximation to the exact diagonal Hessian. While for the areas that the sources don’t cover, the energy is underestimated, as indicated by the white arrows. Figure 4c and d are obtained using the linear phase encoding method with 7 and 31 simulations respectively. Figure 4e and f are obtained using the chirp phase encoding method with 7 and 31 simulations respectively. It can be seen that with increasing the number of simulations, the crosstalk artifacts can be suppressed effectively. And for the same number of simulations, the chirp phase encoding method can obtain a better approximation. Figure 5 show the gradients in the first iteration preconditioned by the corresponding Hessian approximations shown in Figure 4. Figure 5a is the gradient scaled by the diagonal pseudo-Hessian. We can see that the amplitude of the deep reflectors have been recovered obviously. By introducing the receiver-side Green’s functions, the multiple scattering effects are suppressed, as denoted by the black arrows and the resolution is improved.
The Hessian matrix works like a sharpening or focusing filtering to the gradient. Furthermore, the double illumination method based diagonal Hessian and the chirp phase encoded diagonal Hessian can balance the amplitudes better.

**Conclusion**

The Hessian matrix in the least-squares inverse problem can serve as a nonstationary deconvolution operator to compensate the geometrical spreading effects and improve the resolution of the gradient. The diagonal part of the pseudo-Hessian is limited to balance the biased amplitude in the gradient. Three other strategies can balance the amplitude further and sharp or focus the gradient. The chirp phase encoding method can approach the exact approximate Hessian better compared to the linear phase encoding method with the same computational cost.

**ACKNOWLEDGMENTS**

This research was supported by the Consortium for Research in Consortium of Elastic Wave Exploration Seismology (CREWES).

**References**


