GeoConvention
Calgary • Canada • March 7-11 2016
OPTIMIZING RESOURCES

Source-independent Hessian-free Gauss-Newton Full-waveform Inversion

Wenyong Pan and Kris Innanen
CREWES Project, Department of Geoscience, University of Calgary

Summary

The estimation of soruce wavelet is very important for implementing full-waveform inversion (FWI) successfully. Many FWI algorithms estimate the soruce signature iteratively in the inversion process. In this paper, a source-independent method is adpoted by normalizing the wavefield. Furthermore, the gradient-based methods for FWI suffer from slow local convergence rate. A Hessian-free (HF) Gauss-Newton method is implemented in this research by solving the Newton system with a conjugate-gradient (CG) method. The CG method only needs Hessian-vector products instead of constructing the Hessian matrix explicitly. In this paper, the Hessian operator in HF Gauss-Newton method is corrected combining with the source-independent method. We demonstrate with numerical examples that the proposed strategies can improve the convergence rate and reduce the computational burden.

Introduction

In recent years, full-waveform inversion (FWI) is becoming increasing popular in estimating the subsurface parameters by iteratively minimizing a l-2 norm misfit function, which measures the difference between the modelled data and observed data. FWI is exposed to many difficulties, one of which being the lack of information of the source wavelet. Small perturbations of the source wavelet may result in large disturbances in inverted models. Hence, a robust and efficient FWI algorithm needs the information of source wavelet, which is generally unknown in seismic exploration.

In traditional seismic data processing, the source wavelet is estimated from seismic data trace. Many FWI algorithms iteratively update the source wavelet estimation in the inversion process (Zhou and Greenhalgh, 2003; Sun et al., 2014). Zhou and Greenhalgh (2003) consider the source weight as unknown parameter and update it iteratively with unknown velocity. Sun et al., (2014) developed a shallow-response based variable projection type of strategy to estimate the source wavelet alongside the model parameters during FWI. Another class of approach for handling the source effects is to make the inverse problem independent of source wavelet (Lee and Kim, 2003). Lee and Kim (2003) developed a source-independent FWI strategy by normalizing the wavefield in the frequency domain with respect to the frequency response of a reference trace, in which the source spectrum is eliminated. In this paper, a variable projection method is used for data calibration in frequency-domain FWI (Rickett, 2013; Li et al., 2013).

The traditional optimization methods for FWI in exploration geophysics are gradient-based methods (e.g., non-linear conjugate-gradient (NCG) method). In NCG method, the search direction is just the linear combination of the current gradient and previous search direction. Within the adjoint-state method, the gradient of the misfit function can be calculated efficiently by applying a zero-lag cross-correlation between the forward modelled wavefield and back-propagated data residual wavefield (Pratt et al., 1998; Pan et al., 2014c,a, 2015, 2014b***)***. Thus, the gradient-based methods are computationally attractive for large-scale inverse problem but also suffer from slow convergence rate.

In this paper, the NCG, L-BFGS and HF Gauss-Newton methods (Pan et al., 2016) are implemented with the source-independent strategy. In particular, the Gauss-Newton Hessian is corrected with the variable projection strategy, which improves the convergence rate and reduces the computation burden for the Hessian-free Gauss-Newton method.

Review of non-linear least-squares inverse problem

As a non-linear least-squares optimization problem, FWI seeks to estimate the subsurface parameters by iteratively minimizing the difference between the synthetic data and observed data. The misfit function is formulated in a least-squares form:

$$\Phi(\mathbf{m}) = \frac{1}{2}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\sum_{\omega}\left\|\mathbf{d}_{obs}(\mathbf{x}_r,\mathbf{x}_s,\omega)\text{-}\mathbf{s}(\mathbf{x}_s,\,\omega)\mathbf{d}_{syn}(\mathbf{x}_r,\mathbf{x}_s,\omega)\right\|^2, \tag{1}$$

where $\mathbf{x}_r$ and $\mathbf{x}_s$ indicate the locations of receivers and sources, ω is the angular frequency and $\mathbf{s}(\mathbf{x}_s,\,\omega)$ indicates the source weight vector:

$$\mathbf{s}(\mathbf{x}_s,\,\omega)=\frac{\sum_{\mathbf{x}_s}\mathbf{d}_{obs}(\mathbf{x}_r,\mathbf{x}_s,\omega)\mathbf{d}^*_{syn}(\mathbf{x}_r,\mathbf{x}_s,\omega)}{\sum_{\mathbf{x}_r}\mathbf{d}_{syn}(\mathbf{x}_r,\mathbf{x}_s,\omega)\mathbf{d}^*_{syn}(\mathbf{x}_r,\mathbf{x}_s,\omega)}, \tag{2}$$

To minimize the quadratic approximation of the misfit function, the updated model at k+1 iteration can be written as the sum of the model at the kth iteration and the search direction:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \mu_k\Delta\mathbf{m}_k, \tag{3}$$

where μ_k is the step length, a scalar constant calculated through a line search method satisfying the weak Wolfe condition. The gradient can be constructed by the correlation of partial derivative wavefield and complex conjugate of the data residuals. The Jacobian matrix can be expressed as:

$$\mathbf{J}=\frac{\partial\mathbf{d}_{syn}}{\partial\mathbf{m}} = P\mathbf{A}(\mathbf{m},\,\omega)^{-1}\mathbf{A}_{\mathbf{m}}(\mathbf{m},\,\omega)\mathbf{u}\,, \tag{4}$$

where P is the detection operator, $\mathbf{A}(\mathbf{m},\,\omega)$ is the impedance matrix and $\mathbf{A}_{\mathbf{m}}(\mathbf{m},\,\omega)=\frac{\partial\mathbf{A}(\mathbf{m},\,\omega)}{\partial\mathbf{m}}$. The gradient for minimizing the objective function is obtained as:

$$\mathbf{g}=\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\sum_{\omega}\mathbf{u}^*(\mathbf{x}_r,\mathbf{x}_s,\omega)\mathbf{A}^*_{\mathbf{m}}(\mathbf{m},\,\omega)\mathbf{A}^*(\mathbf{m},\,\omega)^{-1}P^*\Delta\mathbf{d}(\mathbf{x}_r,\mathbf{x}_s,\omega), \tag{5}$$

The approximate Hessian used in Gauss-Newton method is expressible as:

$$\mathbf{H}_a=\mathbf{u}^*(\mathbf{x}_r,\mathbf{x}_s,\omega)\mathbf{A}^*_{\mathbf{m}}(\mathbf{m},\,\omega)\mathbf{A}^*(\mathbf{m},\,\omega)^{-1}P^*P\mathbf{A}(\mathbf{m},\,\omega)^{-1}\mathbf{A}^*_{\mathbf{m}}(\mathbf{m},\,\omega)\mathbf{u}, \tag{6}$$

When considering the objective function illustrated in equation (1), the Jacobian matrix becomes (Li et al., 2013):

$$\tilde{\mathbf{J}}=\mathbf{s}(\mathbf{x}_s,\omega)\mathbf{J}(\mathbf{x}_r,\mathbf{x}_s,\omega)+\frac{\mathbf{d}_{syn}(\mathbf{x}_r,\mathbf{x}_s,\omega)}{\sum_{\mathbf{x}_r}\left|\mathbf{d}_{syn}(\mathbf{x}_r',\mathbf{x}_s,\omega)\right|^2}\times\left(\sum_{\mathbf{x}_r}\mathbf{d}_{obs}(\mathbf{x}_r',\mathbf{x}_s,\omega)\frac{\partial\mathbf{d}^*_{syn}(\mathbf{x}_r',\mathbf{x}_s,\omega)}{\partial\mathbf{m}}-2\mathbf{s}(\mathbf{x}_s,\omega)\sum_{\mathbf{x}_r}\mathbf{d}_{syn}(\mathbf{x}_r',\mathbf{x}_s,\omega)\mathbf{J}(\mathbf{x}_r',\mathbf{x}_s,\omega)\right). \tag{7}$$

And the Gauss-Newton Hessian becomes:

$$\tilde{\mathbf{H}}_a = \tilde{\mathbf{J}}^*\tilde{\mathbf{J}}. \tag{8}$$

Examples

In this section, we first apply the source-independent Hessian-free Gauss-Newton (SI-HF-GN) FWI on a Gaussian-anomaly model in comparison with non-linear conjugate gradient (NCG), L-BFGS and HF-GN

methods. The inversion results verify the effectiveness of the variable projection strategy in improving the convergence rate and reducing the computation cost.

The Gaussian Anomaly Model

In this numerical example, the NCG, L-BFGS and HF Gauss-Newton methods are applied on a Gaussian-anomaly model. We examine the quadratic convergence rate of the HF Gauss-Newton FWI in reconstructing the velocity model compared to NCG and L-BFGS methods. We also demonstrate that with the corrected Gauss-Newton Hessian (equation (8)), the convergence rate of the Hessian-free Gauss-Newton method is improved and the computation burden is reduced.

The Gaussian-anomaly model consists of 50 X 100 grid cells with a grid interval of 10 m in both horizontal and vertical directions. A total of 49 sources are deployed from 20 m to 980 m with a source interval of 20 m and a depth of 20 m. A total of 100 receivers are distributed from 10 m to 1000 m with a receiver interval of 10 m and a depth of 20 m. A Ricker wavelet with a 30 Hz dominant frequency is used as the source function. Figure 1 shows the true Gaussian-anomaly P-wave velocity model. The initial velocity model is a homogeneous model with a constant velocity of 2 km/s.

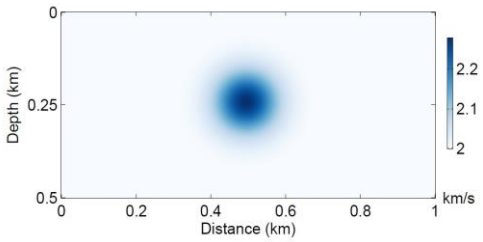

Figure 1. The true Gaussian anomaly model.

For comparison of different methods, we set the stopping criteria for inversion as: the maximum outer iteration $k_{\max} = 21$ and the minimum normalized misfit $\phi_{\min} = 2e-5$. To illustrate the performances of different preconditioning strategies, the stopping criteria for the inner CG algorithm is also defined: the maximum inner iteration $\tilde{k}_{\max} = 50$ and the minimum relative residual $\gamma_{\min} = 1e-2$ .Figures 2a, 2b, 2c and 2d show the inversion results by SD, L-BFGS, HF-GN and SI-HF-GN methods respectively. Figures 3a and 3b show the normalized misfit vs. iteration and RLSE vs. iteration as the iteration proceeds.

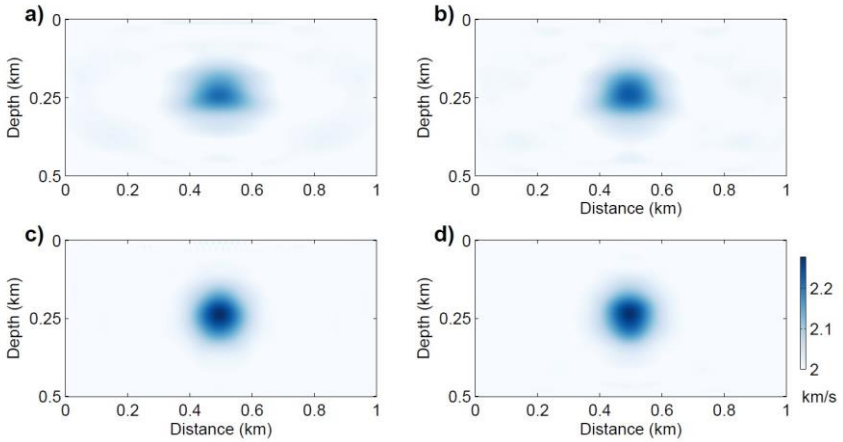

Figure 2. (a) SD method; (b) L-BFGS method; (c) HF-GN method; (d) SI-HF-GN method.

As we can see, compared to SD and L-BFGS methods, the HF Gauss-Newton methods can approach the true model much better. Furthermore, we notice that we the corrected Gauss-Newton Hessian, the

source-independent HF Gauss-Newton method can converge faster. We use both of the misfit and RLSE to evaluate the quality of the inversion result.

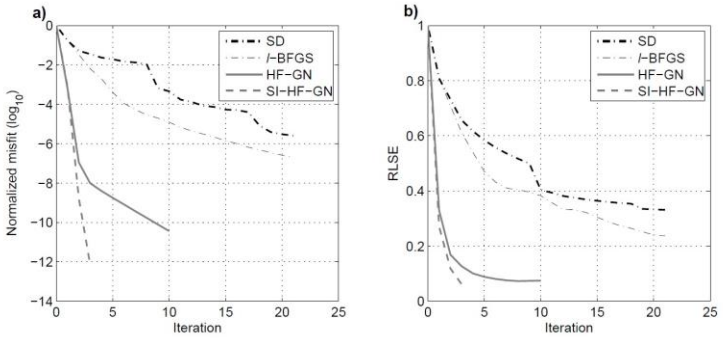

Figure 3. Normalized misfit vs. Iteration (a) and RLSE vs. Iteration (b).

We then apply the source-independent Hessian-free Gauss-Newton FWI on a modified Marmousi model. Figure 4a shows the true P-wave velocity model. The initial velocity model shown in Figure 4b is obtained by smoothing the true P-wave velocity model with a Gaussian function.

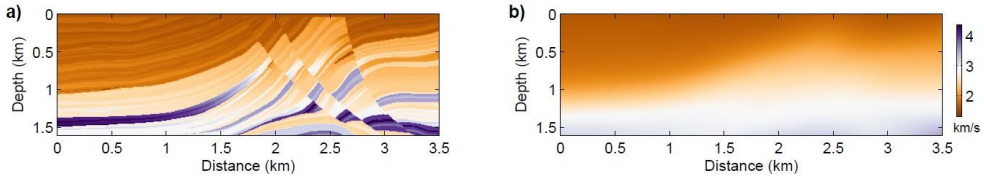

Figure 4. (a) True P-wave velocity model; (b) Initial P-wave velocity model.

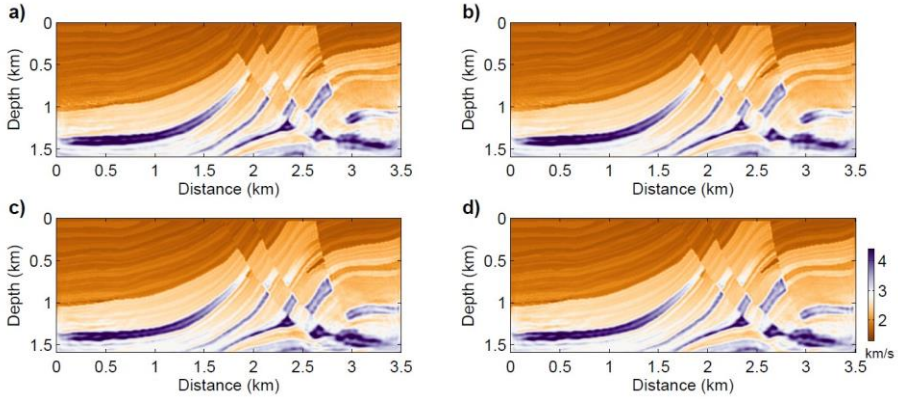

Figure 5. (a) NCG method; (b) L-BFGS method; (c) HF-GN method; (d) SI-HF-GN method.

The inversion results by NCG, L-BFGS and HF Gauss-Newton and source-independent HF Gauss-Newton methods are shown in Figures 5a, 5b, 5c and 5d. We observe that the Gauss-Newton Hessian correction, a better inversion result is achieved.

Conclusions

In this paper, we develop a source-independent Hessian-free Gauss-Newton FWI, in which the Gauss-Newton Hessian is corrected with the source-independent strategy. The numerical examples show that the proposed method can provides faster convergence rate and better inversion result with reducing the computational burden.

Acknowledgements

The authors thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. Author 1 was also supported by a SEG/Chevron scholarship.

References

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration: Conference on Inverse Scattering, Theory and Applications, SIAM, Expanded Abstracts, 206-220.

Lee, K. H., and Kim, H. J., 2003, Source-independent full-waveform inversion of seismic data: Geophysics, 68, 45–59.

Li, M., Rickett, J., and Abubakar, A., 2013, Application of the variable projection scheme for frequency domain full-waveform inversion: Geophysics, 78, R249–R257.

Nash, S. G., 1985, Preconditioning of truncated-newton methods: SIAM J. Sci. Statist. Comput, 6, 599-616.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, 133, 341-362.

Pan, W., K. A. Innanen, and G. F. Margrave, 2014a, A comparison of different scaling methods for least-squares migration/inversion: EAGE Expanded Abstracts, We G103 14.

Pan, W., K. A. Innanen, G. F. Margrave, and D. Cao, 2015a, Efficient pseudo-Gauss-Newton full-waveform inversion in the t-p domain: Geophysics, 80, no. 5, R225-R14.

Pan, W., K. A. Innanen, G. F. Margrave, M. C. Fehler, X. Fang, and J. Li, 2015b, Estimation of elastic constants in HTI media using Gauss-Newton and Full-Newton multi-parameter full waveform inversion: SEG Technical Program Expanded Abstracts, 1177-1182.

Pan, W., G. F. Margrave, and K. A. Innanen, 2014b, Iterative modeling migration and inversion (IMMI): Combining full waveform inversion with standard inversion methodology: SEG Technical Program Expanded Abstracts, 938-943.

Pan, W., K. A. Innanen, and W. Liao, 2016, Preconditioning for the Hessian-free Gauss-Newton full-waveform inversion: Geophysics, in revision.

Rickett, J., 2013, The variable projection method for waveform inversion with an unknown source function: Geophysical Prospecting, 61, 874–881.

Sun, D., Jiao, K., Vigh, D., and Coates, R., 2014, Source wavelet estimation in full waveform inversion: EAGE Expanded Abstracts, Tu E106 09.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, 49, 1259-1266.

Virieux, A. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, no. 6, WCC1-WCC26.

Zhou, B., and Greenhalgh, S. A., 2003, Crosshole seismic inversion with normalized full-waveform amplitude data: Geophysics, 68, 1320–1330.