geoconvention
Calgary • Canada • May 15-19 2017

Full waveform inversion based on nonlinear gradient

Yu Geng and Kris Innanen
CREWES project, Department of Geoscience, University of Calgary

Summary

Full waveform inversion (FWI), which is a highly nonlinear inverse problem, use full waveform information to build high resolution subsurface structures. However, FWI is usually solved iteratively as a local optimization problem, and during each iteration, the gradient is ususally obtained from the Born approximation, which indicates a linear relationship between the scattering points and scattered wavefields. Mitigating linearization error, or to say including nonlinearity, during each updates can be important to improve both FWI results and convergence rates. In this paper, we study a nonlinear gradient by including higher order Born series into the conventional gradient. Numerical examples show the new nonlinear gradient can help building a better model especially in the deeper regions and help FWI converge faster.

Introduction

By minimizing the misfit between recorded data and modeled data, full waveform inversion (FWI) (Lailly, 1983, Tarantola, 1984) has become an attractive tool to build high resolution velocity model. Although waveform changes are highly nonlinear with respect to the model parameter changes, the FWI problem is usually treated as a local optimization problem, and solved iteratively by a linearized version of the nonlinear problem at each step. During each iteration, a local gradient is calculated to solve the local optimization problem and this gradient is usually studied in the framework of Fréchet derivative, or sensitivity kernel, generally derived under the Born approximation, which indicates a linear relationship between the model perturbations and the change of the waveform.

In FWI, several iterations are needed for the inversion to converge toward the minimum of the misfit function, and nonlinearity is mainly mitigated through iterations and multi-scale strategies. Linearization error could be significant in many cases (Wu and Zheng, 2014), especially when large-scale perturbations are involved. As shown in recently applications of FWI (Virieux and Operto, 2009), FWI can successfully build high resolution models for shallow regions, which are sampled mostly by diving waves and post-critical reflections. However, in the deeper regions, which are usually sampled by small-angle backscattered data, such as pre-critical reflections, FWI behaves more like a migration, which can only build the high wavenumber components of the subsurface models. After Xu et al. (2012) revisited the FWI theory to develop a new gradient to better handle the inversion of low wavenumber components of the model sampled by pre-critical reflections, several work has been done to better handle the linearization errors respond to small-angle backscattered data by developing different kinds of gradient (Wu and Alkhalifah, 2015, Zhou et al., 2015).

In this paper, following Innanen's work (2014, 2015), to include nonlinearity into each iteration, we study a nonlinear gradient by including higher order Born series into the conventional gradient. At each iteration, the gradient is constructed not only from the updated model itself but also a perturbation as well. By adding this perturbation, a range of scattering processes can be included in the gradient, each involving one interaction with to-be-determined perturbation and multiple interaction with the added perturbation. This added perturbation is highly related to the data residual, and can be determined using a linear inversion method. Numerical examples based on this nonlinear gradient are shown to verify the nonlinearity are added correctly, with faster convergence rate and better inversion results.

Theory and Method

In this paper, the space-frequency domain isotropic acoustic wave equation with constant density is used to describe the wave motion with the squared slowness parameter $s(\mathbf{r})$. At each iteration, FWI is to minimize the misfit function,

$$\phi(s)=\frac{1}{2}\sum_{\mathbf{r}_s}\sum_{\mathbf{r}_g}\sum_{\omega}\left\|\delta P\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_n\right)\right\|^2, \tag{1}$$

which involves the data residual $\delta P\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_n\right)$ between recorded data $P\left(\mathbf{r}_g,\mathbf{r}_s,\omega\right)$ and modeled data $G\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_n\right)$, by updating the given initial model $s_n(\mathbf{r})$ with a perturbation $\delta s_n(\mathbf{r})$ to approach the real model $s(\mathbf{r})$. The gradient of the misfit function, used to calculate the perturbation, can be calculated with the sensitivity $\partial G/\partial s$ and the complex conjugate of the data residual

$$g_n(\mathbf{r})=-\sum_{\mathbf{r}_s}\sum_{\mathbf{r}_g}\sum_{\omega}\mathrm{Re}\left(\omega^2\frac{\partial G\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_n\right)}{\partial s(\mathbf{r})}\delta P^*\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_n\right)\right). \tag{2}$$

FWI is a highly nonlinear problem, since the relationship between the wavefield and model parameter is nonlinear. One direct way to help mitigating this nonlinearity is to have broadband seismic data with wide-offset range and very good initial model before starting the FWI procedure, however, even with different types of misfit definition proposed (van Leeuwen and Mulder, 2008, Brossier, Operto, and Virieux, 2015), the gradient used to update the model is still obtained under the Born approximation. The other way to mitigate the nonlinearity is taking multiple forward scattering into account. Suppose that the at n+1th iteration, the updated model is determined by $s_{n+1}(\mathbf{r})=s_n(\mathbf{r})+\delta s_n(\mathbf{r})$, when adding a small variation $\delta s(\mathbf{r})$ localized at $\mathbf{r}$ to s_{n+1}, the perturbation wavefield caused by this added variation can be written as

$$\begin{aligned}\delta G\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_{n+1},\delta s(\mathbf{r})\right)=&\omega^2 G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n\right)G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n\right)\delta s(\mathbf{r})\\&+\omega^4\int d\mathbf{r}'\,G\left(\mathbf{r}_g,\mathbf{r}',\omega\,|\,s_n\right)G\left(\mathbf{r}',\mathbf{r},\omega\,|\,s_n\right)G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n\right)\delta s_n(\mathbf{r}')\delta s(\mathbf{r})\\&+\omega^4\int d\mathbf{r}'\,G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n\right)G\left(\mathbf{r},\mathbf{r}',\omega\,|\,s_n\right)G\left(\mathbf{r}',\mathbf{r}_s,\omega\,|\,s_n\right)\delta s_n(\mathbf{r}')\delta s(\mathbf{r})+\ldots\end{aligned}, \tag{3}$$

where the first term can be interpreted as scattering process involving one interaction the added variation $\delta s(\mathbf{r})$ only, and the higher order terms can be interpreted as multiple interactions with $\delta s_n(\mathbf{r})$ and one interaction with $\delta s(\mathbf{r})$. The sensitivity of δG respect to $\delta s(\mathbf{r})$ can then be written as

$$\begin{aligned}\frac{\partial G\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_{n+1}\right)}{\partial s(\mathbf{r})}=\omega^2\Big(&G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n\right)G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n\right)\\&+\delta G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n,\delta s_n\right)G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n\right)+G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n\right)\delta G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n,\delta s_n\right)\Big)\end{aligned}. \tag{4}$$

Substituting sensitivities (4) back to the gradient (2), we can get a nonlinear gradient which contains higher order Born series as

$$\begin{aligned}g_{n+1}(\mathbf{r})=-\sum_{\mathbf{r}_s}\sum_{\mathbf{r}_g}\sum_{\omega}\mathrm{Re}\Big(\omega^2\delta P^*\left(\mathbf{r}_g,\mathbf{r}_s,\omega\,|\,s_{n+1}\right)\Big(&G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n\right)G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n\right)\\&+\delta G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n,\delta s_n\right)G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n\right)+G\left(\mathbf{r}_g,\mathbf{r},\omega\,|\,s_n\right)\delta G\left(\mathbf{r},\mathbf{r}_s,\omega\,|\,s_n,\delta s_n\right)\Big)\Big)\end{aligned}. \tag{5}$$

It is obvious that when $\delta s_n(\mathbf{r})$ is zero, the nonlinear gradient (5) is equivalent to the conventional gradient at nth iteration. Perturbation $\delta s_n(\mathbf{r})$ is highly related to the data residual, so instead of calculating $\delta s_n(\mathbf{r})$ during FWI, we use a linear inversion to calculate $\delta s_n(\mathbf{r})$ first in each iteration before calculating gradient (5).

Examples

We first show the single frequency sensitivity in the case of a homogeneous model with one reflector as in Figure 1. The conventional kernel (Figure 1a) is calculated in the homogeneous model without the reflector, and the nonlinear kernel (Figure 1b) is calculated in the homogeneous model with the reflector as the perturbation $\delta s_n(\mathbf{r})$. It can be seen that compared to the conventional kernel, which has the forward scattering between source and receiver along the first Fresnel zone at the shallow depth and the back scattering along the secondary Fresnel zones as the outer fringes in the deeper regions, the added two terms in the nonlinear kernel build two more first Fresnel zones between both source and receiver to the reflector, which can further provide the ability to update the low wavenumber components of deeper part of the model.

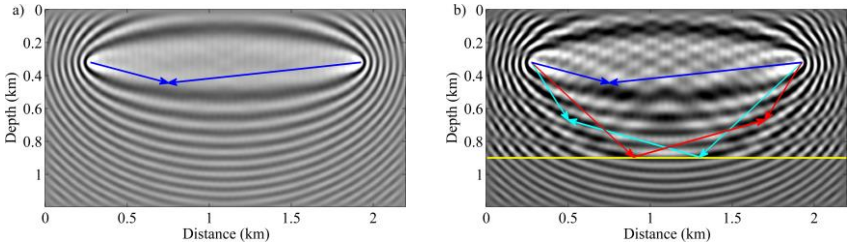

Figure 1: Single frequency sensitivity kernel in a homogeneous model with one reflector. a) Conventional sensitivity; b) nonlinear sensitivity. Yellow line indicates the reflector, blue, cyan and red lines stands for related wave paths.

The Marmousi model is used to perform the synthetic test of FWI with the nonlinear gradient. The true model is shown in Figure 2a, we add a 500m water layer on the top of the model. We use a linearly increasing model as the initial model, which is exact at the added water layer, but changes linearly along depth direction from 1.5km/s to 4km/s, as shown in Figure 2b. We use 461 fixed receivers and 46 sources along the surface to generate data for 5 frequencies (2Hz, 3.3Hz, 5.5Hz, 9Hz and 14.9Hz) for the inversion. Figure 2c shows the inversion result from FWI with conventional gradient with 10 iterations for each frequency, and Figure 2d shows the inversion result from FWI with nonlinear gradient with same iterations. Velocity profiles along x = 4km, 6km are shown in Figure 3. Since the initial model is far away from the true velocity model, we update the velocity at the water layer in this test. However, even with frequency as low as 2Hz, conventional FWI can only reconstruct rough structures of the model, but nonlinear gradient FWI gives accurate high resolution result even at the deeper regions. Figure 4a shows the norm of data residual vector vs. number of iterations, and Figure 4b shows the relative model least-squares error vs. number of iterations. It is observed that FWI with nonlinear gradient converges faster than FWI with conventional gradient.

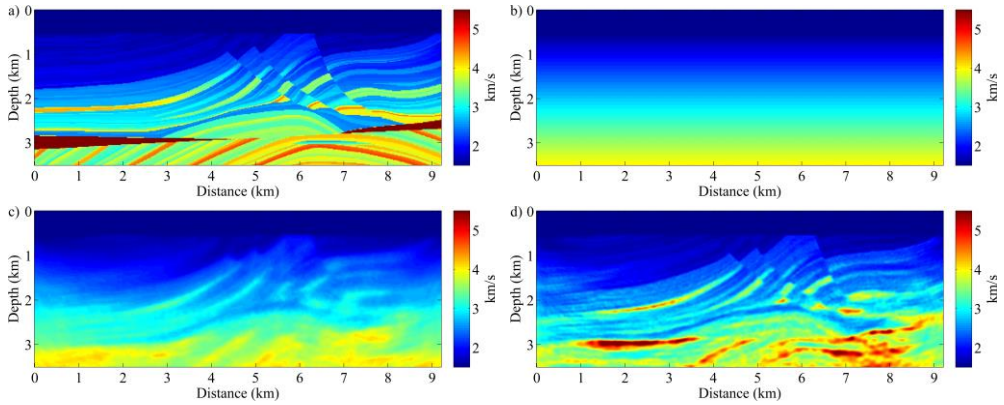

Figure 2: a) True Marmousi velocity model; b) initial velocity model; Inversion result with conventional gradient and b) nonlinear gradient d) with 10 iterations for 5 frequencies starting from 2Hz to 15Hz.

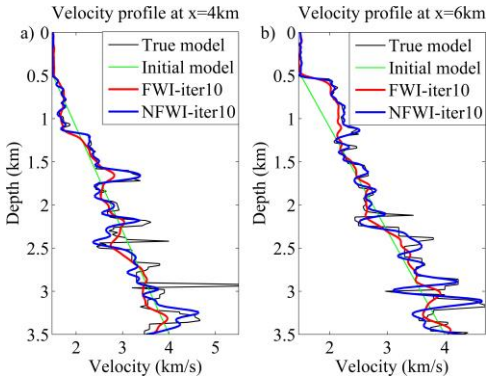

Figure 3: Velocity profiles along a) x = 4km and b) x = 6km.

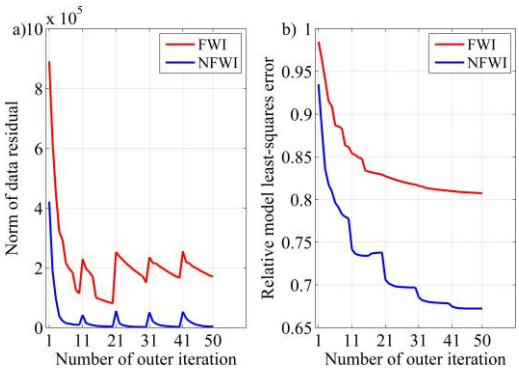

Figure 4: a) Norm of data residual vector vs. number of iterations; b) relative model least-squares error vs. number of iterations.

Next, we use uniformly sampled 12 frequencies between 4Hz to 15Hz to perform FWI with both conventional and nonlinear gradient with the same initial model as in Figure 2b. Figure 5a shows the inversion result from conventional gradient with 10 iterations for each frequency, and Figure 5b shows the inversion result from nonlinear gradient with same iterations. Velocity profiles along x = 4km, 6km are shown in Figure 6. With the conventional gradient, velocity is updated incorrectly, especially at the shallow area around z = 1km to 1.5km along vertical profile x = 4km and deeper area along the vertical profile x = 6km, where inversion with nonlinear gradient provides a fairly good result.

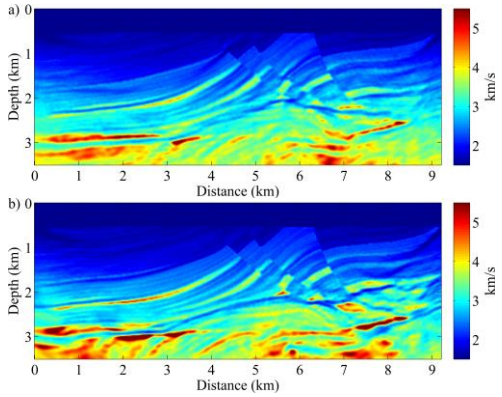

Figure 5: Inversion results using a) conventional gradient, b) nonlinear gradient with 10 iterations for 12 frequencies starting from 4Hz to 15Hz.

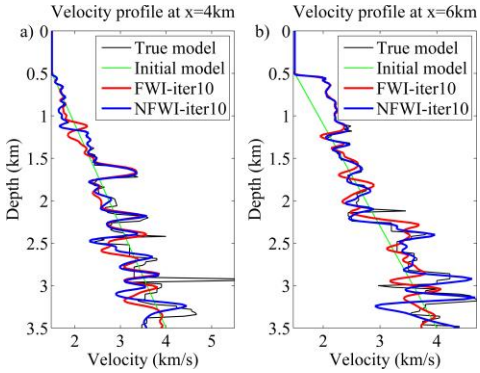

Figure 6: Velocity profiles along a) x = 4km and b) x = 6km.

Conclusions

In this study, by including both zero order and higher order terms in the calculation of sensitivities, we have constructed a nonlinear gradient to better handle nonlinearity in FWI, which can provide a better update to the model and converges faster than FWI with the conventional gradient. The inversion results of the Marmousi model with different frequency bands illustrate the nonlinear gradient FWI can build better velocity model in deeper regions even in the case of lack low frequency information.

Acknowledgements

We thank the sponsors of CREWES for support. This work was funded by CREWES and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 379744-08.

References

Brossier, R., S. Operto, and J. Virieux, 2015, Velocity model building from seismic reflection data by full-waveform inversion: Geophysical Prospecting, 63, 354-367.

Innanen, K. A., 2014, Seismic full waveform inversion with nonlinear sensitivities: CREWES Annual Report, 26.

Innanen, K. A., 2015, Residual dependent FWI sensitivities based on direct nonlinear inverse scattering: CREWES Annual Report, 27.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society of Industrial and Applied Mathematics, Expanded Abstracts. 206-220.

Tarantola, A., 1984, Inversion of Seismic-Reflection Data in the Acoustic Approximation: Geophysics, 49, 1259-1266.

van Leeuwen, T., and W. A. Mulder, 2008, Velocity analysis based on data correlation: Geophysical Prospecting, 56, 791-803.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, Wcc1-Wcc26.

Wu, R. S., and Y. C. Zheng, 2014, Non-linear partial derivative and its De Wolf approximation for non-linear seismic inversion: Geophysical Journal International, 196, 1827-1843.

Wu, Z. D., and T. Alkhalifah, 2015, Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion: Geophysics, 80, R317-R329.

Xu, S., D. Wang, F. Chen, G. Lambaré, and Y. Zhang, 2012, Inversion on reflected seismic wave: SEG Technical Program Expanded Abstracts. 1-7.

Zhou, W., R. Brossier, S. Operto, and J. Virieux, 2015, Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation: Geophysical Journal International, 202, 1535-1554.