We P1 16

Full Waveform Inversion Using An Energy-Based Objective Function with Efficient Calculation of the Gradient

Y. Choi* (King Abdullah University of Science & Technology), T. Alkhalifah (King Abdullah University of Science & Technology)

Summary

Full waveform inversion (FWI) using an energy-based objective function has the potential to provide long wavelength model information even without low frequency in the data. However, without the back-propagation method (adjoint-state method), its implementation is impractical for the model size of general seismic survey. We derive the gradient of the energy-based objective function using the back-propagation method to make its FWI feasible. We also raise the energy signal to the power of a small positive number to properly handle the energy signal imbalance as a function of offset. Examples demonstrate that the proposed FWI algorithm provides a convergent long wavelength structure model even without low-frequency information, which can be used as a good starting model for the subsequent conventional FWI.

Introduction

Full waveform inversion (FWI) has been widely studied and developed to delineate accurate description of the subsurface content. Nevertheless, FWI still suffers from the local minima problem mainly resulting from the cycle-skipping between the observed and modelled data. The available frequency low enough in the data or a starting model close enough to the true one can help FWI converge to an accurate solution. In general, however, the frequency band of seismic data is usually not low enough and obtaining a good starting model is not a trivial task as it requires using elaborate approaches such as travel-time tomography, migration velocity analysis, and etc.

Recently, many studies of FWI have been devoted to solving the cycle-skipping problem in the case of a lack of low frequency information. Choi and Alkhalifah (2015) unwrapped the phases at relatively high frequency in the frequency-domain and inverted the unwrapped phase to avoid the cycle-skipping problem. Wu et al. (2014) suggested FWI based on the envelope objective function, where the envelope of seismic trace yields artificial low frequency components which help FWI construct a long wavelength model.

On the other hand, Jeon et al. (2014) proposed the energy objective function for FWI to construct a macro-velocity model, which can be used as a good starting model for a subsequent conventional FWI. The energy (or energy signal) of a seismic trace is defined as an integration of squared wavefield with a shifting time window. They showed that the energy objective function has a convex shape free of local minima and their FWI approach generates a convergent long wavelength velocity model. However, to calculate the gradient of the energy objective function, they explicitly calculate the partial derivative of the energy signal using a finite-difference method, which makes their FWI approach impractical especially for the model size of a general exploration seismic survey.

In this abstract, we derive the gradient of the energy objective function using the back-propagation method (adjoint-state method) to make the proposed FWI practical for the model size of a general seismic survey. We also provide theoretical insights into the feasibility of the energy objective function. On the other hand, the energy signal has singularly strong amplitude (energy) near zero-offset because of the integration of the squared wavefield, therefore, the conventional least-square misfit function does not properly handle the energy signal. Jeon et al. (2014) took the logarithm of the energy signal to address the features of the energy signal. We propose raising the energy signal to the power of a very small positive number (e.g., 10^{-5}) to properly handle it in FWI. In the examples, we apply the proposed FWI with the back-propagation method to the low-cut filtered synthetic data generated in a salt-dome model. Examples demonstrate the feasibility and potential of FWI using the energy objective function especially in case of a lack of low frequency information.

Energy signal

The energy of a seismic trace is originally defined as

$$e \equiv \int_0^T [u(\tau)]^2 d\tau, \quad (1)$$

where $u(\tau)$ is a seismic trace with time variable and T is the recording time. On the other hand, we estimate the energy signal (Jeon et al. 2014) as

$$u_e(t) = \int_0^t [u(\tau)]^2 d\tau. \quad (2)$$

The energy signal in equation 2 is an integration of the squared seismic trace with shifting time window. If the shifting time window is replaced with the Heaviside step function, the energy signal is expressed as a convolution of the Heaviside step function and squared trace:

$$u_e(t) = \int_{-\infty}^{\infty} H(t-\tau) [u(\tau)]^2 d\tau, \quad (3)$$

where $u(\tau)$ is a causal signal. The Heaviside step function in equation 3 plays the role of a filter, which amplifies very low frequency components. On the other hand, the squared wavefield in

equation 3 is expressed as an auto-convolution in the frequency-domain, which yields artificial low frequency components. Therefore, artificial very low frequency components dominate in the energy signal in equations 2 and 3. Figure 1 shows the original low-cut filtered trace, squared trace, Heaviside step function, energy signal, and their corresponding frequency-domain spectra. Although the original trace is low-cut filtered below 5 Hz (Figures 1a and 1e), the squared trace includes artificial low frequency components (Figures 1b and 1f). The spectrum of Heaviside step function exponentially decreases as frequency increases (Figure 1g), thus the energy signal includes mainly artificial very low-frequency components (Figures 1d and 1h).

Figure 2 shows the low-cut filtered seismogram and its corresponding energy signal seismogram with different clipping ranges. The energy signal has singularly very strong amplitude near zero offset (Figures 2b and 2c), thus the least square misfit function cannot properly handle the energy signal seismogram through FWI.

Figure 1 The first row shows the time-domain traces and the second row is their corresponding frequency-domain spectra: the (a, e) original low-cut filtered trace, (b, f) its squared trace, (c, g) Heaviside step function, and (d, h) energy signal.

Figure 2 The (a) filtered seismogram and (b, c) corresponding energy signal seismogram with different clipping ranges.

Objective functions using energy signal and gradient calculation

As an alternative to the conventional misfit function, Jeon et al. (2014) took the logarithm of the energy signal to address the features of the energy signal seismogram. In addition to the logarithm, we suggest raising the energy signal to the power of very small positive number. The objective functions using logarithm and power for the energy signal are written as

$$E_{\log} = \sum_{s,r} \left[\frac{1}{2} \int_0^T \left[\ln \left\{ \frac{u_e(t)}{d_e(t)} \right\} \right]^2 dt \right] \quad \text{and} \quad E_{\text{power}} = \sum_{s,r} \left[\frac{1}{2} \int_0^T \left[(u_e(t))^\alpha - (d_e(t))^\alpha \right]^2 dt \right], \quad (4)$$

respectively, where $u_e(t)$ and $d_e(t)$ are the energy signals of the modelled and observed wavefield, respectively, and α is a power of energy signal. The logarithm and the power with a very small number in equation 4 can mitigate the effect of the singularly very strong amplitude near zero offset in Figure 2b. We set the value of α as 10^{-5} in the examples. The gradients of the objective functions in equation 4 can be obtained by taking the derivative of the objective functions with respect to the k th model parameter:

$$\frac{\partial E_{\log}}{\partial p_k} = \sum_{s,r} \left[\int_0^T \left[\frac{\partial u_e(t)}{\partial p_k} \frac{1}{u_e(t)} \ln \left\{ \frac{u_e(t)}{d_e(t)} \right\} \right] dt \right] \quad \text{and} \quad \frac{\partial E_{\text{power}}}{\partial p_k} = \sum_{s,r} \left[\int_0^T \left[\frac{\partial u_e(t)}{\partial p_k} \alpha (u_e(t))^{\alpha-1} \left\{ (u_e(t))^\alpha - (d_e(t))^\alpha \right\} \right] dt \right]. \quad (5)$$

Jeon et al. (2014) explicitly calculated the partial derivative of energy signal ($\partial u_e(t)/\partial p_k$) using the finite-difference method, which is very expensive to calculate and makes FWI impractical for the model size of general seismic survey. In this abstract, we derive the gradient of the above objective functions using the back-propagation method in a similar way to reverse-time migration. The partial derivative of energy signal in equation 5 is expressed in a vector form from equation 2 as

$$\left(\frac{\partial \mathbf{u}_e}{\partial p_k} \right)^T = 2 \left(\sum_j^1 \left\{ u_j \frac{\partial u_j}{\partial p_k} \right\} \quad \sum_j^2 \left\{ u_j \frac{\partial u_j}{\partial p_k} \right\} \quad \cdots \quad \sum_j^n \left\{ u_j \frac{\partial u_j}{\partial p_k} \right\} \right), \quad (6)$$

where the subscript ‘ j ’ stands for the time sample element and n is the number of total time samples of vector $\mathbf{u}_e$. By expanding each element in equation 6, the partial derivative of energy signal can be rewritten as

$$\left(\frac{\partial \mathbf{u}_e}{\partial p_k} \right)^T = 2 \left(\frac{\partial \mathbf{u}}{\partial p_k} \right)^T \mathbf{A}, \quad \text{where} \quad \mathbf{A} = \begin{bmatrix} u_1 & u_1 & u_1 & \cdots & u_1 \\ 0 & u_2 & u_2 & \cdots & u_2 \\ 0 & 0 & u_3 & \cdots & u_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_n \end{bmatrix}, \quad (7)$$

where $\mathbf{u}$ is a vector form of the original modelled wavefield. Finally, the gradients in equation 5 are expressed in vector/matrix form as

$$\frac{\partial E_{\log}}{\partial p_k} = \sum_{s,r} \left[\left(\frac{\partial \mathbf{u}}{\partial p} \right)^T (\mathbf{A} \mathbf{r}_{\log}) \right], \quad \text{where} \quad r_{\log}(t) = \frac{1}{u_e(t)} \ln \left\{ \frac{u_e(t)}{d_e(t)} \right\}, \quad (8)$$

$$\frac{\partial E_{\text{power}}}{\partial p_k} = \sum_{s,r} \left[\left(\frac{\partial \mathbf{u}}{\partial p} \right)^T (\mathbf{A} \mathbf{r}_{\text{power}}) \right], \quad \text{where} \quad r_{\text{power}}(t) = (u_e(t))^{\alpha-1} \left\{ (u_e(t))^\alpha - (d_e(t))^\alpha \right\}. \quad (9)$$

From equations 8 and 9, we can efficiently calculate the gradients using the back-propagation method (or adjoint-state method), which is almost the same as the conventional FWI (or RTM) except for adjoint sources given by the new residual seismograms $\mathbf{A} \mathbf{r}_{\log}$ and $\mathbf{A} \mathbf{r}_{\text{power}}$.

Examples

We generate synthetic data from the SEG/EAGE salt model shown in Figure 3a and filter out the data below 3 Hz. The maximum frequency of data is 8 Hz. We place 370 shots at a depth of 20 m and 770 receivers at the same depth. One of seismograms is shown in Figure 2a. The starting velocity model is linearly increasing with depth (Figure 3b). The logarithmic objective function provides similar results to that of the power objective function, thus we display only results of the power objective function to save space in this abstract. The proposed FWI algorithms generate long wavelength structures in the inverted model and the gradient (Figures 4a and 4b) even though the data lacks low frequency information. The subsequent conventional FWI starting from the previous inverted model in Figure 4a generates a quite good convergent velocity model delineating well both the inside and the shape of salt (Figure 4c), whereas the one starting from the initial model in Figure 3b provides poor results (Figure 4d). The examples demonstrate that the proposed FWI algorithm generates a convergent long wavelength velocity model, which can be used as a good starting model for the subsequent FWI.

Figure 3 The (a) SEG/EAGE salt model and (b) starting model for FWI.

Figure 4 The (a) inverted model obtained from FWI using the energy-based objective function, (b) gradient at the first iteration, and subsequent conventional FWI results starting from the (c) previous inverted model and (d) initial model in Figure 3b.

Conclusions

FWI using the energy objective function has a potential to provide long wavelength structure information even without low frequency in the data, since the energy signal mainly includes artificial low frequency components. However, without the back-propagation method (or adjoint-state method), FWI using the energy objective function is impractical for the model size of a general seismic survey. We derive the gradient of objective function using the back-propagation method to make the proposed FWI practical. The energy signal has singularly very strong amplitudes near zero offset, in which the conventional misfit function does not work. We propose raising the energy signal to the power of a small positive number in addition to taking the logarithm to properly deal with the energy signal imbalance. Numerical examples show that the proposed FWI algorithm with the back-propagation method provides a convergent long wavelength structure model even without low frequency information, which can be used as a good starting model for the subsequent conventional FWI.

Acknowledgements

We are grateful to King Abdullah University of Science and Technology for financial supports. For computer time, this research used the resources of the Super Computing Laboratory at KAUST in Thuwal, Saudi Arabia.

References

- Choi, Y. and Alkhalifah, T. [2015] Unwrapped phase inversion with an exponential damping. *Geophysics*, **80**, R251-R264.
- Jeon, S., Ha, W., Kwon, J. and Shin, C. [2014] Full waveform inversion using the energy objective function. *76th EAGE Conference & Exhibition, Expanded Abstracts, Amsterdam*, Tu E106 02.
- Wu, R.S., Luo, J. and Wu, B. [2014] Seismic envelope inversion and modulation signal model. *Geophysics*, **79**, WA13-WA24.