Waveform inversion with exponential damping using a deconvolution-based objective function

Yunseok Choi* and Tariq Alkhalifah, King Abdullah University of Science and Technology

Summary

The lack of low frequency components in seismic data usually leads full waveform inversion into the local minima of its objective function. An exponential damping of the data, on the other hand, generates artificial low frequencies, which can be used to admit long wavelength updates for waveform inversion. Another feature of exponential damping is that the energy of each trace also exponentially decreases with source-receiver offset, where the least-square misfit function does not work well. Thus, we propose a deconvolution-based objective function for waveform inversion with an exponential damping. Since the deconvolution filter includes a division process, it can properly address the unbalanced energy levels of the individual traces of the damped wavefield. Numerical examples demonstrate that our proposed FWI based on the deconvolution filter can generate a convergent long wavelength structure from the artificial low frequency components coming from an exponential damping.

Introduction

Full waveform inversion (FWI) has been widely studied because theoretically it can provide detailed subsurface velocity structures. FWI, however, still faces a big problem in the lack of low frequencies in the observed data. Low frequency components are essential in FWI to construct a long wavelength structure and the lack of low frequency components usually leads FWI to local minima. A starting velocity model close enough to the true model solves the local minima problem resulting from the lack of low frequency components, but obtaining a good starting velocity model is not a trivial task.

Many studies have been devoted to solving the local minima problem resulting from the lack of low frequency components. Warner and Guasch (2014 a,b) suggested the adaptive waveform inversion where the weighted Wiener filter is minimized to avoid the local minima. Wu et al. (2014) constructed the objective function using the envelope of the whole trace which includes the artificial low frequency components and helps FWI update long wavelength structures. On the other hand, Xu et al. (2012) and Wu and Alkhalifah (2015) simulated the reflected waves from the migrated images and inverted the reflected waveforms to obtain a long wavelength gradient.

In this abstract, we propose a time-domain FWI with exponential damping applied to the data. In this case, we need to mute energy (possibly noise) prior to the first arrival before applying the exponential damping. Such

damping along the time-axis amplifies events near the first arrival and, thus, generates artificial low frequencies. We invert these low frequencies in our proposed FWI to obtain a long wavelength structure. The exponential damping also decreases the energy (or amplitude) of traces exponentially with offset. After applying an exponential damping, most of the energy is focused at the near offset and early time in the seismogram. Therefore, the least-square misfit-based objective function cannot address the feature of the exponentially damped wavefield.

Exponential damping has been used widely for Laplace-domain FWI (Shin and Cha, 2008, 2009) and the unwrapped phase inversion in the frequency-domain (Choi and Alkhalifah, 2015) to mitigate the cycle-skipping problem. In the Laplace- and frequency-domain FWI, the exponentially damped wavefields can be properly handled using the complex logarithm or phase value. In the time-domain, however, it is not easy to deal with the exponentially damped wavefield for FWI, since the energy of traces exponentially decreases with offset. We suggest a deconvolution-based objective function to properly deal with the exponentially damped wavefield in the time-domain FWI. Since the deconvolution filter includes the division process between the modeled and observed data, it can properly address the unbalanced energy levels of each traces in the damped wavefield.

Warner and Guasch (2014 a,b) calculated the deconvolution (Wiener) filter in the time-domain for adaptive waveform inversion. We estimate the deconvolution filter in the frequency-domain using only low frequency band and take the inverse Fourier transform of it. We apply a weighting function to the deconvolution filter and minimize the L2 norm of the weighted deconvolution filter to update velocity model.

In the numerical examples, we first generated the synthetic data from the Marmousi velocity model and filtered out low frequencies of the data below 5 Hz. We applied our proposed FWI algorithm to the low-cut filtered synthetic data. Examples demonstrate that our proposed FWI algorithm can generate a good convergent long wavelength structure without low frequency components.

Theory

The observed and modeled data with an exponential damping are expressed as

$$\hat{d}(t) = d(t)\exp(-\alpha t) \text{ and } \hat{u}(t) = u(t)\exp(-\alpha t), \quad (1)$$

where $d(t)$ and $u(t)$ are the original observed and modeled data, $\hat{d}(t)$ and $\hat{u}(t)$ are the exponentially damped wavefields, and α is a damping factor.

Figure 1: (a) An original observed wavefield filtered out below 5 Hz and (b) its exponentially damped version with a damping factor of 4.

Figure 2: The frequency spectrum within 0 ~ 5 Hz for each traces of (a) the original observed wavefield in Figure 1a and (b) the damped wavefield in Figure 1b. The frequency spectrum of each trace is normalized by its maximum value for clear comparison.

Figure 1 shows one of shot gathers with a low frequency cut of 5 Hz and its exponentially damped version with a damping factor of 4. In the damped seismogram (Figure 1b), the events near the first arrival are relatively large compared to the later arrivals, and this feature of damping can generate artificially low frequencies. We can also note that the energy (or amplitude) of traces in the damped seismogram (Figure 1b) is exponentially decreasing as a function of offset. Figure 2 shows the frequency spectrums of the original and damped seismograms in Figure 1 in the frequency-space domain. In Figure 2, we normalized each spectrum of the trace with its maximum value for clear comparison. Since we pick the first arrival time and mute noises prior to the first arrival, some weak low frequency components are shown below 5 Hz (Figure 2a). On the other hand, the exponentially damped wavefields have much stronger amplitude in low frequency components compared to the undamped wavefields (Figure 2b). These

artificial low frequencies can be used to construct long wavelength structures through FWI.

The conventional least-square misfit function for FWI, however, does not properly address the exponentially decreasing energy of traces along the offset-axis in the damped seismogram in Figure 1b. Because most of the energy is focused at very near offsets and early time in the damped seismogram, FWI based on the least-square misfit updates only the shallow part of the model.

As an alternative to the least-square misfit, we employ the deconvolution filter for the objective function. Warner and Guasch (2014 a,b) calculated the deconvolution filter in the time-domain for the adaptive waveform inversion, where an elaborate optimization is required to solve the matrix problem. On the other hand, we calculate the deconvolution filter in the frequency-domain, where no optimization is required, and also selectively choose a frequency band to calculate an optimal deconvolution filter. We first take the Fourier transform of the damped wavefield:

$$\tilde{d}(\omega) = FFT[\hat{d}(t)] \text{ and } \tilde{u}(\omega) = FFT[\hat{u}(t)], \quad (2)$$

where FFT stands for the Fourier transform operation. The normalization of the modeled data by the observed data for the deconvolution is usually stabilized as

$$\tilde{f}(\omega) = \frac{\tilde{u}(\omega) \cdot \text{conj}(\tilde{d}(\omega))}{\tilde{d}(\omega) \cdot \text{conj}(\tilde{d}(\omega)) + \varepsilon}, \quad (3)$$

where ‘conj’ indicates the complex conjugate and ε is the stabilization term. This normalization method, however, is not suitable for FWI, because the deconvolution filter does not go to the Dirac delta function even when the modeled data is very close to the observed data. We propose another normalization method:

$$\tilde{f}(\omega) = \begin{cases} \frac{\tilde{u}(\omega)}{\tilde{d}(\omega)}, & \text{when } \text{abs}(\tilde{d}(\omega)) \geq \varepsilon \\ 1, & \text{when } \text{abs}(\tilde{d}(\omega)) < \varepsilon \end{cases}, \quad (4)$$

where ‘abs’ stands for the absolute value of the complex number. In this method, the deconvolution filter in the time-domain becomes close to the Dirac delta function when the modeled data converges to the observed data. We take the inverse Fourier transform of $\tilde{f}(\omega)$ in equation 4 to get the time-domain deconvolution filter:

$$f(t) = FFT^{-1}[\tilde{f}(\omega)], \quad (5)$$

where $f(t)$ is a deconvolution filter in the time-domain and FFT^{-1} means the inverse Fourier transform.

We construct the objective function for FWI using the deconvolution filter:

$$E = \sum_{\text{shot, receiver}} \|f(t)w(t)\|_2^2, \quad (6)$$

where $w(t)$ is a weighting function. Since $f(t)$ should be close to the Dirac delta function as the modeled data converges to the observed data, we set $w(t)$ as a linearly increasing function in t away from $t=0$ and minimize the objective function to update velocity model.

Since the deconvolution filter in equations 4 and 5 includes a normalization of the modeled data by the observed data, the objective function in equation 6 can properly handle the exponentially decreasing energy level of traces along the offset-axis, shown in the damped seismogram in Figure 1b.

We obtain the gradient expression by taking the derivative of the objective function with respect to the k th model parameter p_k :

$$\frac{\partial E}{\partial p_k} = \sum_{\text{shot, receiver}} \left\{ \frac{\partial u(t)}{\partial p_k} r(t) \right\}, \quad (7)$$

where

$$r(t) = FFT^{-1} \left[\frac{\tilde{d}(\omega) \cdot FFT[w^2(t)f(t)]}{\tilde{d}(\omega) \cdot \text{conj}(\tilde{d}(\omega))} \right] \cdot \exp(-\alpha t). \quad (8)$$

The only difference between the proposed FWI and conventional FWI is the back-propagated residual expressed in equation 8. To calculate the gradient, we back-propagate the new residual in equation 8 and then estimate the zero-lag correlation of the back-propagated wavefield and weighted forward-propagated wavefield.

Examples

We test the proposed FWI algorithm on the Marmousi synthetic data. We generate the synthetic (observed) data with a maximum frequency of 16 Hz (peak frequency is 8 Hz) and filter out frequencies below 5 Hz. Representative shot-gather and frequency spectrum are shown in Figures 1 and 2. The starting model for FWI is a linearly increasing velocity model (1.5 ~ 4 km/s). We apply an exponential damping with a damping factor of 4 to the synthetic data and invert the artificial low frequency components. We use the steepest descent direction to update velocity model.

Figure 3 shows the inverted models for the damped wavefield. The frequency range used for FWI is 0 ~ 3 Hz. We compare the result of our proposed FWI with that of conventional FWI based on the least-square misfit function. The conventional FWI updates only the shallow part of the model (Figure 3a), since the least-square misfit function cannot address the exponentially decreasing energy of traces along the offset-axis in the damped wavefield shown in Figure 1b. On the other hand, our proposed FWI provides a good convergent long wavelength structure model (Figure 3a), which demonstrates that the

deconvolution-based objective function works well for the exponentially damped wavefield and its artificial low frequency components.

Figure 3: The inverted models for a frequency band of 0 ~ 3 Hz using the (a) conventional least-square misfit objective function and (b) deconvolution-based objective function.

We estimate and display the deconvolution filter for each trace at the first and last iteration of our proposed FWI (Figure 4). The deconvolution filter is displayed in time-distance domain for the same shot gather in Figure 1. At the first iteration, the deconvolution filter has some high values at non-zero time-lag especially at far offsets (Figure 4a), whereas it becomes close to the Dirac delta function at the last iteration (Figure 4b), which shows the convergence of our proposed FWI.

For a subsequent FWI, we reduce the damping factor to 2 and extend the frequency band to 0 ~ 5 Hz for the proposed FWI. The starting model is the previous result shown in Figure 3b. The inverted model includes higher wavenumber structures (Figure 5a). Finally, we apply the conventional FWI to the observed data without damping starting from the velocity model in Figure 5a. The final inverted model shows highly detailed and compatible structures with the Marmousi model (Figure 5b). For comparison, we apply the conventional FWI to the observed data without damping starting from a linearly increasing velocity model and the inverted model converges to a wrong velocity structure (Figure 5c) resulting from the lack of low frequency components and the starting model far from the true model.

Numerical examples demonstrate that our proposed FWI based on the deconvolution filter can properly deal with the

exponentially damped wavefield and generate a good convergent long wavelength structure using the artificial low frequencies coming from an exponential damping, whereas the least-square misfit function does not work for the exponentially damped wavefield.

Figure 4: The deconvolution filters estimated in the propose FWI at the (a) first and (b) last iteration.

Figure 5: The subsequent FWI results (a) starting from Figure 3b with a reduced damping factor of 2 and a extended frequency band of 0 ~ 5 Hz, (b) starting from Figure 5a without damping by the conventiona FWI, and (c) starting from the linearly increasing velocity model without damping by the conventiona FWI.

Discussions

Exponential damping applied to the observed data generates artificial low frequencies, but also amplifies noise prior to the first arrival, thus we must pick the first arrival traveltimes and mute noises prior than the first arrival before applying an exponential damping.

A strong exponential damping can make amplitude of far offset trace smaller than the computer precision in some cases. In this case, the division process included in the deconvolution filter could give us erroneous results. Therefore, a strong exponential damping is not a good choice for our proposed FWI. We choose a damping factor of 4 for the inversion examples, which is not strong but moderate. Choosing an appropriate damping factor should be careful considering the computer precision.

Conclusions

An exponential damping applied to seismic data generates artificial low frequencies. We proposed a FWI algorithm to invert such artificial low frequencies. However, the least-square misfit function does not work well for damped wavefields due to the bias in amplitudes towards near offsets as a resulting of the damping. We suggest the deconvolution filter as an alternative objective function to invert the exponentially damped wavefields. Since the deconvolution filter includes a division between the modeled and observed data, it can properly handle the exponentially decreasing energy levels along the offset axis in the damped wavefield. We apply a weighting function to the deconvolution to enhance the non-zero-time energy and form an objective function based on the weighted deconvolution filter to update the velocity model. The only difference between conventional FWI and this deconvolution-based FWI is the residual seismogram, which is back-propagated to calculate the backward wavefield. Numerical examples show that the proposed FWI method based on the deconvolution filter successfully generates long wavelength structures from the artificially low frequency components and the subsequent FWI provides a highly detailed and convergent velocity model, whereas the least-square misfit function does not.

Acknowledgments

We are grateful to King Abdullah University of Science and Technology for financial support.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2016 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Choi, Y., and T. Alkhalifah, 2015, Unwrapped phase inversion with an exponential damping: *Geophysics*, **80**, no. 5, R251–R264, <http://dx.doi.org/10.1190/geo2014-0498.1>.
- Shin, C., and Y. Ho Cha, 2008, Waveform inversion in the Laplace domain: *Geophysical Journal International*, **173**, 922–931, <http://dx.doi.org/10.1111/j.1365-246X.2008.03768.x>.
- Shin, C., and Y. Ho Cha, 2009, Waveform inversion in the Laplace-Fourier domain: *Geophysical Journal International*, **177**, 1067–1079, <http://dx.doi.org/10.1111/j.1365-246X.2009.04102.x>.
- Warner, M., and L. Guasch, 2014, Adaptive waveform inversion — FWI without cycle skipping — Theory: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, We-E106-13.
- Warner, M., and L. Guasch, 2014, Adaptive waveform inversion: Theory: 84th Annual International Meeting, SEG, Expanded Abstracts, 1089–1093, <http://dx.doi.org/10.1190/segam2014-0371.1>.
- Wu, R. S., J. Luo, and B. Wu, 2014, Seismic envelope inversion and modulation signal model: *Geophysics*, **79**, no. 3, WA13–WA24, <http://dx.doi.org/10.1190/geo2013-0294.1>.
- Wu, Z., and T. Alkhalifah, 2015, Simultaneous inversion of the background velocity and the perturbation in full waveform inversion: 77th Annual International Conference and Exhibition, EAGE, Extended Abstracts, Th-N104-14.
- Xu, S., D. Wang, F. Chen, G. Lambare, and Y. Zhang, 2012, Inversion on reflected seismic wave: 82nd Annual International Meeting, SEG, Expanded Abstracts, <http://dx.doi.org/10.1190/segam2012-1473.1>.