ASEG-PESA 2013
'THE EUREKA MOMENT'
11-14 AUGUST 2013 • MELBOURNE, AUSTRALIA

Robust scaling strategy for frequency-domain acoustic full waveform inversion

Ju-Won Oh
Seoul National University
1, Gwanak-ro, Gwanak-gu, Seoul
juwon24@snu.ac.kr

Dong-Joo Min
Seoul National University
1, Gwanak-ro, Gwanak-gu, Seoul
spoppy@snu.ac.kr

SUMMARY

One of the limitations in seismic waveform inversion is noise, which plays a role in impeding that the FWI converges on the global minimum solution. We first analyse noise behaviours in the inverse problem and limitations of the conventional FWI strategy for noisy data. For a more robust inversion algorithm, we propose applying a frequency-depth scaling strategy. The frequency-depth scaling strategy is designed using the denoise function for both the spectral scaling and the flexible damping factor of the Levenberg-Marquardt method for the depth scaling so that the influence of noise-dominated gradient directions can be weakened during the inversion. Numerical example for synthetic data with low-frequency-dominant random noise for the Marmousi-2 model shows that our new scaling strategy can improve the robustness of the frequency-domain FWI with little human intervention and without any prior information.

Key words: full waveform inversion, inverse problem, denoise function, Levenberg-Marquardt method

INTRODUCTION

Seismic full waveform inversion (FWI), which is a technique to identify subsurface physical properties, has been extensively studied. Nevertheless, FWI still has some problems to be resolved such as local minima, the effect of noise and lack of reliable low-frequency data. Because the noise can hinder that the FWI converges on the global minimum solution, there have been attempts to increase the robustness of FWI for noisy data by introducing robust objective functions. Brossier et al. (2010) compared the robustness of various objective functions for noisy data and concluded that the l_1-norm objective function is the most robust. However, most robust objective functions are not robust for the data containing random noise, whose S/N ratio is spectrally different. Oh and Min (2012) addressed that, for more robust FWI, we have to consider not only the robustness of the objective function but also the spectral distribution of noise to prevent the noisy single-frequency gradient from contributing the inverse process. However, to increase the robustness of the FWI for noisy data, we also need to consider the proper depth scaling strategy. For these reasons, to minimise the influence of random noise in frequency-domain acoustic FWI, we propose applying a frequency-depth scaling strategy that combines the spectral scaling strategy using the denoise function suggested by Oh and Min (2012) and the depth scaling strategy using the constraint of the Levenberg-Marquardt method (Lines and Treitel, 1984).

INVERSE THEORY FOR NOISY DATA

In frequency-domain waveform inversion using the Levenberg-Marquardt method, the model parameter change vector (so-called gradient direction), which minimises the differences between the observed ($\mathbf{d}_{s,\omega}$) and modelled data ($\mathbf{u}_{s,\omega}$), can be expressed as follows:

$$\delta\mathbf{p} = \left[\sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^{T}\mathbf{J}_{s,\omega} + \beta\mathbf{I}\right)\right]^{-1}\left[\sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^{T}(\mathbf{u}_{s,\omega} - \mathbf{d}_{s,\omega})\right)\right] \quad (1)$$

where $\delta\mathbf{p}$ and β are the model parameter change vector and the damping factor (i.e., Lagrange multiplier), respectively, and the terms ω and s denote the angular frequency and source, respectively. The superscripts T and * indicate the transpose and complex conjugate, respectively. The term $\mathbf{J}_{s,\omega}$ is the Jacobian matrix consisting of partial derivative wavefields.

In eq. (1), if the observed data are noise-free, model parameters can approach the global minimum solution, although there are numerous local minima. However, when the observed data includes noise components, the global minimum of objective function can deviate from those of the actual subsurface structures. In this case, because the observed data can be decomposed by

$$\mathbf{d} = \mathbf{d}^{event} + \mathbf{d}^{noise}, \quad (2)$$

the gradient direction in eq. (1) can be also decomposed as follows

$$\nabla_{\mathbf{p}}^{\text{total}} E = \nabla_{\mathbf{p}}^{\text{event}} E + \nabla_{\mathbf{p}}^{\text{noise}} E, \quad (3)$$

where

$$\nabla_{\mathbf{p}}^{\text{event}} E = \left[\sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^{T}\mathbf{J}_{s,\omega} + \beta\mathbf{I}\right)\right]^{-1}\left[\sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^{T}(\mathbf{u}_{s,\omega} - \mathbf{d}_{s,\omega}^{\text{event}})\right)\right] \quad (4)$$

and

$$\nabla_{\mathbf{p}}^{\text{noise}} E = \left[\sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^{T}\mathbf{J}_{s,\omega} + \beta\mathbf{I}\right)\right]^{-1}\left[\sum_{\omega}\sum_{s}\left(-\mathbf{J}_{s,\omega}^{T}\mathbf{d}_{s}^{\text{noise}}(\omega)\right)\right], \quad (5)$$

which are called 'pure seismic gradient direction' and 'noise direction', respectively, throughout the paper to distinguish

from the total gradient direction ($\nabla_{\mathbf{p}}^{\text{total}} E$). The noise direction, $\nabla_{\mathbf{p}}^{\text{noise}} E$, corresponds to the reverse time migration image obtained by back-propagating only the noise components. As described by eq. (3), the total gradient direction in the noise-included inverse problem is affected by the noise direction, which degrades the inversion results by noise-induced artefacts during the inversion.

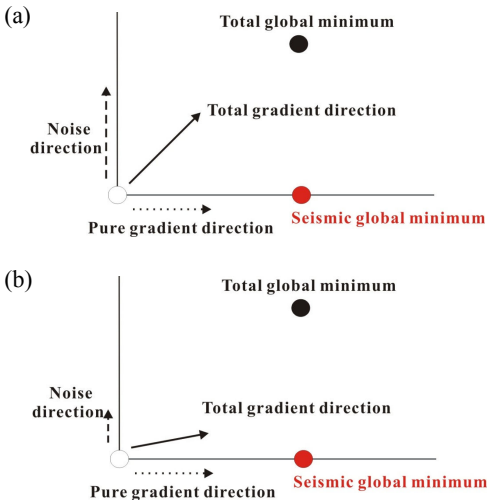

Figure 1. Schematic diagrams illustrating the parameter search directions in the noise-included FWI algorithm (a) without and (b) with constraints refraining from the data-fitting process along the noise direction

Figure 1 shows the schematic diagrams, which illustrate the relationship among the pure gradient, noise and total gradient directions. When observed data are noise-free, the inversion process concentrates on finding the seismic global minimum solution, which is the goal of seismic inversion. However, when noises are included in observed data (Figure 1a), the inversion process converges to not the seismic global minimum but the total global minimum while updating the combination of the pure seismic-gradient direction and the noise direction. Consequently, the optimized solution, which minimises the noise-included objective function, can deviate from the seismic global minimum solutions due to the parameter updates along the noise direction. To reduce this negative influence of noises on the inversion process, we must constrain the data-fitting process along the noise direction to make the model parameter update converge to the seismic global minimum solution, as shown in Figure 1b.

In this study, for computational convenience, we calculate the descent direction suggested by Ha et al. (2009) as follows:

$$\delta \mathbf{p}_{\mathrm{d}} = \sum_{\omega} \left[\mathrm{NRM} \left[\mathrm{diag} \left\{ \mathbf{H}_n(\omega) \right\} + \beta \mathbf{I} \right]^{-1} \nabla_{\mathbf{p}}^{\text{single}} E(\omega) \right], \quad (6)$$

with

$$\mathbf{H}_n(\omega) = \sum_{s} \left[\left(\mathbf{F}_{s,\omega}^{v} \right)^{T} \mathbf{A} \left(\mathbf{F}_{s,\omega}^{v} \right)^{*} \right], \quad (7)$$

$$\mathrm{diag}(\mathbf{A}) = \mathrm{Re} \left\{ \sum_{i=1}^{ns} \left| g_{i,1} \right| \quad \sum_{i=1}^{ns} \left| g_{i,2} \right| \quad \cdots \quad \sum_{i=1}^{ns} \left| g_{i,np} \right| \right\} \quad (8)$$

and

$$\nabla_{\mathbf{p}}^{\text{single}} E(\omega) = \sum_{s} \mathrm{Re} \left\{ \left[\mathbf{F}_{s,\omega}^{v} \right]^{T} \left[\mathbf{S}_{\omega}^{-1} \right]^{T} \left[\mathbf{u}_{s,\omega} - \mathbf{d}_{s,\omega} \right]^{*} \right\}, \quad (9)$$

where $\mathbf{H}_n(\omega)$, g_i and $\nabla_{\mathbf{p}}^{\text{single}} E(\omega)$ are the new pseudo-Hessian matrix (Choi et al., 2008), the impulse response for the i$^{\text{th}}$ source and the single-frequency gradient, respectively. The terms *ns* and *np* denote the number of sources and the number of nodal points, respectively. The terms $\mathbf{F}_{s,\omega}^{v}$ and $\mathbf{S}_{\omega}^{-1}$ are the virtual source matrix and the inverse of the modelling operator, respectively. The term NRM represents the normalising operator, which divides each single-frequency descent direction vector by its maximum absolute value.

ROBUST SCALING STRATEGY

To overcome the aforementioned limitation of the conventional FWI strategy for noisy data, we suggest a frequency-depth scaling strategy, which includes the spectral scaling strategy using the denoise function and the depth scaling strategy using the flexible damping factor. Because background noises (so-called ambient ground motions) are induced by natural and cultural sources, such as the wind, ocean and human activities, the spectrum of noises is incoherent with that of seismic signals. This means that certain frequency bands can be dominantly contaminated depending on the signal-to-noise ratio (S/N ratio) at each frequency. For this reason, to refrain the parameter update along the noise direction in the spectral aspect, we need to suppress the contribution of these noise-dominated frequency components to the inverse process.

Using the denoise function (Oh and Min, 2012), which acts as a frequency-filter, the descent direction can be rewritten by

$$\delta \mathbf{p}_{d} = \sum_{\omega} g(\omega) \left[\mathrm{NRM} \left[\mathrm{diag} \left\{ \mathbf{H}_n(\omega) \right\} + \beta \mathbf{I} \right]^{-1} \nabla_{\mathbf{p}}^{\text{single}} E(\omega) \right], \quad (10)$$

where

$$g(\omega) = \frac{\sum_{r} \left| \sum_{s} u_{s,r}(\omega) \right|}{\sum_{r} \left| \sum_{s} d_{s,r}(\omega) \right|}. \quad (11)$$

The terms *s* and *r* denote the source and receiver numbers, respectively. According to Oh and Min (2012), by summing the monochromatic observed and modelled data over all shot gathers, the denoise function tends to be proportional to the S/N ratio at each frequency, particularly for the random noise-included data. Finally, the denoise function has a similar shape to a frequency-filter and prevents the noisy single-frequency descent direction from contributing to the total descent direction.

In eq. (1), the S/N ratio of the residual wavefields is determined by not only the observed data but also modelled data, which means that the S/N ratio decreases as the iteration proceeds because modelled seismic signals get closer to observed seismic signals. In other words, the S/N ratio of the residual wavefields can spatially vary depending on the model parameter restoration. For example, if the upper structures are well recovered, the residual wavefield for early arrival reflected waves gets dominated by noise. For this reason, to restrict the parameter along the noise direction in the spatial aspect, we need to apply depth scaling strategy, which reduces the parameter update for the well recovered depth interval.

As the depth scaling strategy, we employ the flexible damping factor, β , which is expressed by

$$\beta = \frac{k}{100} \times \max \left| \mathbf{H} \right|, \quad (12)$$

where $\max \left| \mathbf{H} \right|$ and k indicate the maximum value of the Hessian matrix and its percentage, respectively. As an initial damping factor, we use sufficiently large value (10 % of the maximum value of the new pseudo-Hessian matrix in this

study). When the inversion process diverges (i.e., the trend for the RMS error increases), the damping factor is divided by 2 so that it can decrease to 50 % of its former value. This approach allows the inversion process to focus on the shallower structures in the early stage and the deeper structures in the late stage like the layer-stripping FWI (Wang and Rao, 2009).

NUMERICAL EXAMPLE

To verify the robustness of the frequency-depth scaling strategy, we perform the frequency-domain acoustic FWI for the Marmousi-2 model. Figure 2 shows the P-wave velocity structure of the modified Marmousi-2 model. For an initial P-wave velocity model, we use a gradually increasing model ranging from 1.5 to 4.54 km/s. The parameters used for the inversion are listed in Table 1. We apply the frequency-depth scaling strategy for data containing random noise dominant at low-frequency components. Figure 3 shows synthetic seismograms without and with the low-frequency-dominated random noise.

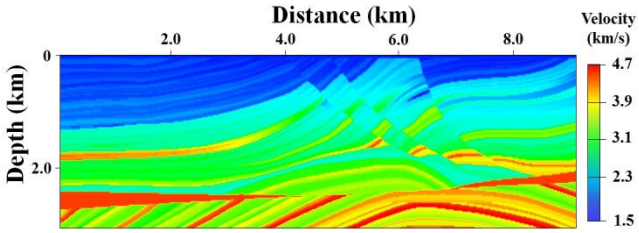

Figure 2. The P-wave velocity structure of the Marmousi-2 model

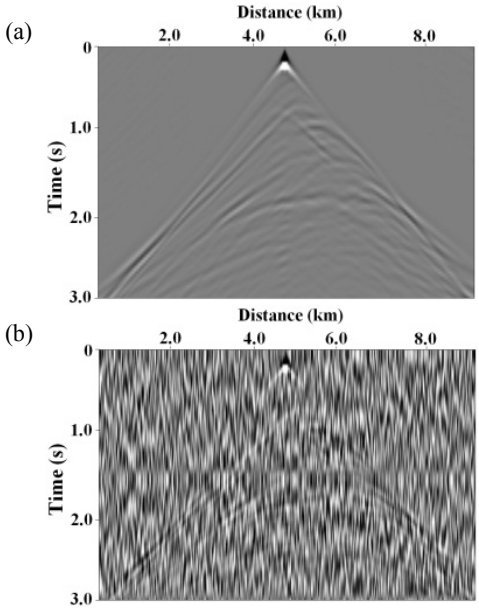

Figure 3. Synthetic seismograms for pressure fields (a) without and (b) with the low-frequency random noise

Figure 4 shows the recovered velocity structures. When acoustic FWI is used without the frequency-depth scaling strategy (Figure 4a), the recovered velocities have some distortions induced by the low-frequency random noise. The velocity structure recovered using the depth scaling strategy (Figure 4b) shows great improvements compared with that obtained using the fixed damping strategy (Figure 4a). However the inversion result still shows some low-frequency noise-induced artefacts. When we apply the spectral scaling strategy (Figure 4c), these low-frequency noise-induced artefacts are reduced due to the high-pass filter-like behaviour of the denoise function although deeper structures are still not recovered well because of the improper damping condition of the conventional FWI. Finally, when we apply the depth scaling strategy along with the spectral scaling strategy (Figure 4d), the inversion result is more compatible with the true model than the results obtained with either the depth or the spectral scaling strategy, even though we use noisy low-frequency data.

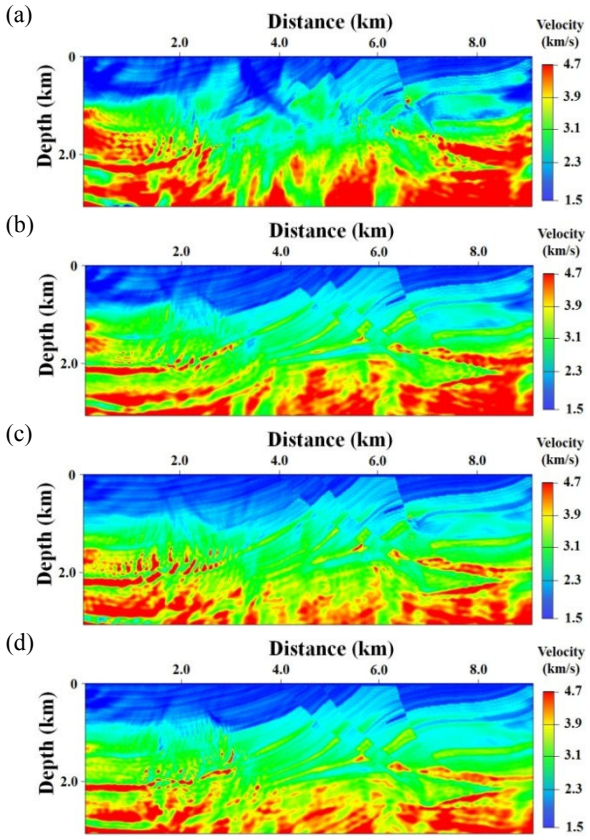

Figure 4. P-wave velocity structures recovered at the 200th iteration for the data with low-frequency random noise by (a) the conventional FWI, (b) only the depth scaling strategy using the flexible damping factors, (c) only the spectral scaling strategy using the denoise function and (d) the frequency-depth scaling strategy using both the denoise function and flexible damping factor

CONCLUSIONS

We suggested the robust scaling strategy for frequency-domain acoustic FWI for noisy observed data, particularly data containing random noise. For the spectral scaling strategy, we employ a denoise function defined as the ratio of the modelled data to real data summed over sources and receivers. For the depth scaling strategy, we employ the flexible damping factor from the Levenberg-Marquardt method. The inversion results for synthetic data containing low-frequency-dominated random noise for the Marmousi-2 model show that the denoise function effectively filters out the noisy low-frequency descent direction. In addition, the flexible damping factor acts like a depth filter, which means that the damping factor controls the energy concentration in the gradient during the inversion. In the early iterations, the energy of the gradient

is concentrated in the shallow parts, whereas in the later iterations, the energy concentration moves to the deeper parts. The denoise function and the damping factor are determined with little human intervention and without any prior information about the noise spectrum or subsurface structures during the inversion and improve the inversion results for data containing random noise.

ACKNOWLEDGMENTS

This work was financially supported by the Human Resources Development program (No. 20124010203200) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Knowledge Economy, the "Development of Technology for CO_2 Marine Geological Storage" program funded by the Ministry of Land, Transport and Maritime Affairs (MLTM) of Korea and the Korea CCS R&D Centre (KCRC) grant funded by the Korea government (Ministry of Education, Science and Technology) (No. 2012-0008926).

REFERENCES

Brossier, R., Operto, S. and Virieux, J., 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion?. *Geophysics*, 75, R37-R46.

Choi, Y., Min, D.J. and Shin, C., 2008, Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Experience of elastic Marmousi-2 synthetic data. *Bulletin of the Seismological Society of America*, 98, 2402 -2415.

Ha, T., Chung, W. and Shin, C., 2009, Waveform inversion using a back-propagation algorithm and a Huber function norm. *Geophysics*, 74, R15-R24.

Lines, L.R. and Treitel, S., 1984, Tutorial: A review of least-squares inversion and its application to geophysical problems. *Geophysical Prospecting*, 32, 159-186.

Oh, J.W. and Min, D.J., 2012, Frequency-domain elastic full waveform inversion using a denoise function. *Expanded Abstracts of the 82nd Annual International Meeting*, SEG, Las Vegas, 1-5.

Wang, Y. and Rao, Y., 2009, Reflection seismic waveform tomography. *Journal of Geophysical Research*, 114, B03304.

Table 1. Parameters used for inversion

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
9.2 km × 3.04 km	230	461	0.04 km	0.02 km	3 sec	15 Hz	1/3 Hz