Re-establishment of gradient in frequency-domain elastic waveform inversion

Ju-Won Oh, Seoul National University, Seoul, Korea

Dong-Joo Min*, Seoul National University, Seoul, Korea, spoppy@snu.ac.kr

and

Felix J. Herrmann, The University of British Columbia, BC, Canada

GeoConvention 2012:Vision

Summary

To obtain solutions close to global minimum in waveform inversion, the gradients computed at each frequency need to be weighted to appropriately describe the residuals between modeled and field data. While the low-frequency components of the gradients should be weighted to recover the long-wavelength structures, the high-frequency components of the gradients need to be weighted when the short-wavelength structures are restored. However, the conventional elastic waveform inversion algorithms cannot properly weight the gradients computed at each frequency. When gradients are scaled using the pseudo-Hessian matrix inside the frequency loop, gradients obtained at high frequencies are over-emphasized. When the gradients are scaled outside the frequency loop, gradients are weighted by the source spectra. In this study, we propose applying weighting factors to the gradients obtained at each frequency so that gradients can properly reflect the differences between the true and assumed models satisfying the general inverse theory. The weighting factors are composed by the backpropagated residuals. Numerical examples for the simple rectangular-shaped model and the modified version of the Marmousi-2 model show that the weighting method enhances gradient images and inversion results compared to the conventional inversion algorithms.

Introduction

In the waveform inversion, model parameters can be updated using gradients to minimize the objective function built on the basis of residuals between modeled and field data. Considering that modeled and field data reflect the features of assumed and true models, the gradients should properly describe the differences between the assumed and true models. One of the reasons that the waveform inversion suffers from local minima problems may be that the gradient cannot appropriately describe such differences. In general, gradients computed at low frequencies contribute to reconstructing long-wavelength structures, whereas gradients calculated at high frequencies contribute to restoring short-wavelength structures. Only when they are properly weighted, we can obtain solutions close to global minimum.

In this study, we analyze the limitation of the conventional frequency-domain elastic waveform inversion using the pseudo-Hessian matrix as a pre-conditioner. Based on the analysis, we modify the conventional waveform inversion algorithm where the gradients are scaled with the pseudo-Hessian matrix inside the frequency loop. In our method, the gradients at each frequency are weighted by the back-propagated residuals to compute the model parameter update vector, through which the model parameter update vector can mainly be influenced by the differences between the assumed and true models. We investigate if the newly constructed gradients can properly reflect the differences between the assumed and true models for the rectangular-shaped model. We also examine if the weighting method yields reliable solutions for the modified version of the Marmousi-2 model when initial models are poorly estimated.

Theory and Method

The objective function using the l_2 norm of residuals between modelled and field data can be written by

$$E(\mathbf{p}) = \sum_{\omega}\sum_{s}\frac{1}{2}\left\|\mathbf{u}_s(\omega,\mathbf{p}) - \mathbf{d}_s(\omega)\right\|_2^2, \tag{1}$$

where $\mathbf{d}_s$ and $\mathbf{u}_s$ are the field and modelled data, respectively, ω and $\mathbf{p}$ indicate the angular frequency and the model parameter vector, respectively, and the subscript s denotes the source position. When we assume that the gradient method is applied to update the model parameter vector and the gradient is computed using the adjoint operator (Pratt et al. 1998), the model parameter update at the l-th iteration can be expressed as

$$\delta\,\mathbf{p}^{(l)} = \mathbf{p}^{(l+1)} - \mathbf{p}^{(l)} = -\alpha\nabla_{\mathbf{p}}E \tag{2}$$

with

$$\nabla_{\mathbf{p}}E = \sum_{\omega}\sum_{s}\mathrm{Re}\left\{\left(\mathbf{F}_s^v\right)^T\left(\mathbf{S}^{-1}\right)^T\left(\mathbf{u}_s - \mathbf{d}_s\right)^*\right\}, \tag{3}$$

where α is the step length, $\mathbf{F}^v$ and $\mathbf{S}$ indicate the virtual source and the complex impedance matrix, respectively, and the superscripts T and $*$ denote the transpose and the complex conjugate, respectively. The last two terms indicate the back-propagation of residuals (Pratt et al. 1998). In an effort to enhance the resolution of the deeper parts, we use the Hessian matrix as a pre-conditioner. To reduce computational burden, we use the diagonal of the pseudo-Hessian matrix (Shin et al. 2001) rather than the full or approximate Hessian matrices.

Jang et al. (2009) addressed that the gradient can be scaled by the diagonal of the pseudo-Hessian matrix either inside or outside the frequency loop, which can be given by

$$\delta\mathbf{p} = \sum_{\omega}\frac{\sum_{s}\mathrm{Re}\left\{\left(\mathbf{F}_s^v\right)^T\left(\mathbf{S}^{-1}\right)^T\left(\mathbf{u}_s - \mathbf{d}_s\right)^*\right\}}{\sum_{s}\mathrm{diag}\left\{\left(\mathbf{F}_s^v\right)^T\left(\mathbf{F}_s^v\right)^* + \beta\mathbf{I}\right\}} \tag{4}$$

and

$$\delta\mathbf{p} = \frac{\sum_{\omega}\sum_{s}\mathrm{Re}\left\{\left(\mathbf{F}_s^v\right)^T\left(\mathbf{S}^{-1}\right)^T\left(\mathbf{u}_s - \mathbf{d}_s\right)^*\right\}}{\sum_{\omega}\sum_{s}\mathrm{diag}\left\{\left(\mathbf{F}_s^v\right)^T\left(\mathbf{F}_s^v\right)^* + \beta\mathbf{I}\right\}}, \tag{5}$$

respectively, where β is a damping factor. Let us consider that the source spectra are included in the virtual source matrix and the residual vector. In eq. (4), the source spectra can be cancelled out in the numerator and the denominator, whereas in eq. (5), the source spectra cannot be cancelled out because the gradient and the pseudo-Hessian matrices are summed over frequencies. Consequently, in the second method, the model parameter update is influenced by the source spectra, which may indicate that the low- and high-end frequency components rarely contribute to the model parameter update. On the other hand, although the first method is not affected by the source spectra, it has another problem. In the first method, the model parameter update is not mainly influenced by the gradients themselves, because of the cancellation of the virtual sources in both the denominator and the numerator. This is not consistent with the general inverse theory. To make the first method consistent with the general inverse theory, we have normalized the model parameter updates computed at each frequency by their maximum values. However, this procedure equalizes the gradients computed at each frequency, which gives an effect of over-emphasizing the high-frequency components of gradients.

Considering that it is not easy to remove only the source wavelet conserving the general inverse theory, we modify the first method by applying weights to the gradients at each frequency. The weighting functions should play a role in re-establishing the original property of gradients at each frequency. For the weighting functions, we use the average of wavefields recorded at all the nodal point by backpropagating the residuals measured at the receivers when the source is located in the middle of the surface for computational convenience, which yields

$$\delta\mathbf{p} = \sum_{\omega} \gamma_{\omega} \frac{\delta\mathbf{p}_{\omega}}{\left|\delta\mathbf{p}_{\omega}\right|_{\max}} \quad (6)$$

with

$$\delta\mathbf{p}_{\omega} = \frac{\sum_{s} \mathrm{Re}\left\{\left(\mathbf{F}_{s}^{v}\right)^{T}\left(\mathbf{S}^{-1}\right)^{T}\left(\mathbf{u}_{s}-\mathbf{d}_{s}\right)^{*}\right\}}{\sum_{s} \mathrm{diag}\left\{\left(\mathbf{F}_{s}^{v}\right)^{T}\left(\mathbf{F}_{s}^{v}\right)^{*}+\beta\mathbf{I}\right\}} . \quad (7)$$

Examples

We first investigate if the two conventional methods and the new weighting method can properly describe the differences between the true and assumed models comparing gradients obtained at the first iteration using the three methods for the simple rectangular-shaped model as shown in Figure 1. For initial models, we assume the homogeneous models whose velocities are the same as those of the background media of Figure 1a. Figures 1b to 1d show the gradient images of λ, which indicate that the first conventional method over-emphasizes the high-frequency components and the second conventional method emphasizes the frequency components corresponding to the source spectra. On the other hand, the weighting method properly reflects the differences between the true and assumed models.

We apply the three methods to the modified version of the Marmousi-2 model, where density is constant at 2.0 g/cm^3 and S-wave velocities were adjusted so that Poisson's ratio can be constant at 0.25. To avoid the computational overburden, we use the central part of the model whose dimension is 9.2 km x 3.04 km. For the source wavelet, we use the first derivative of the Gaussian function with the maximum frequency of 10 Hz. For waveform inversion, we employ frequencies from 0.2 Hz to 10 Hz with an interval of 0.2 Hz. Figure 2 shows inversion results obtained by the two conventional methods and the new weighting method, when initial guesses are poorly estimated. In Figure 2, we can see that the weighting method yields better results than those of the two conventional methods.

Conclusions

In the conventional elastic waveform inversion using the pseudo-Hessian matrix, gradients can be scaled by the pseudo-Hessian matrices either inside or outside the frequency loop. Both methods have some limitations that the gradients computed at each frequency do not properly contribute to updating the model parameter vector. When we scale the gradient inside the frequency loop, the gradients computed at high frequencies are over-emphasized. On the other hand, when the gradient is scaled outside the frequency loop, the gradients are weighted by the source spectra. To obtain solutions close to global minimum, the gradients computed at each frequency should be properly weighted according to the differences between the assumed and true models. This can be achieved by introducing appropriate weights to the gradients. In this study, we proposed the weighting method based on the inversion algorithm that scales gradients with the pseudo-Hessian matrix inside the frequency loop. The weighting factors are obtained using the backpropagated residuals. Numerical examples showed that the gradients obtained by the weighting method are appropriately weighted to describe the differences between the assumed and true models. Inversion results for the modified version of the Marmousi-2 model indicate that the weighting method is not sensitive to initial guesses. In this study, we only provided examples for ideal cases, fixing the density and Poisson's ratio and using the low-frequency components. Further study is needed to investigate the feasibility of the weighting method for real field data.

Acknowledgements

This work was financially supported by the Brain Korea 21 project of Energy System Engineering, the Basic Science Research Program funded by the Ministry of MEST, the Energy Efficiency and Resources funded by the Ministry of MKE, and the "Development of Technology for CO_2 Marine Geological storage" program funded by the Ministry of MLTM in Korea.

References

Jang, U., Min, D. J., and Shin, C., 2009, Comparison of scaling methods for waveform inversion: Geophysical Prospecting, **57**, 49-59.

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

Shin, C., Jang, S., and Min, D. J., 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, 592-606.

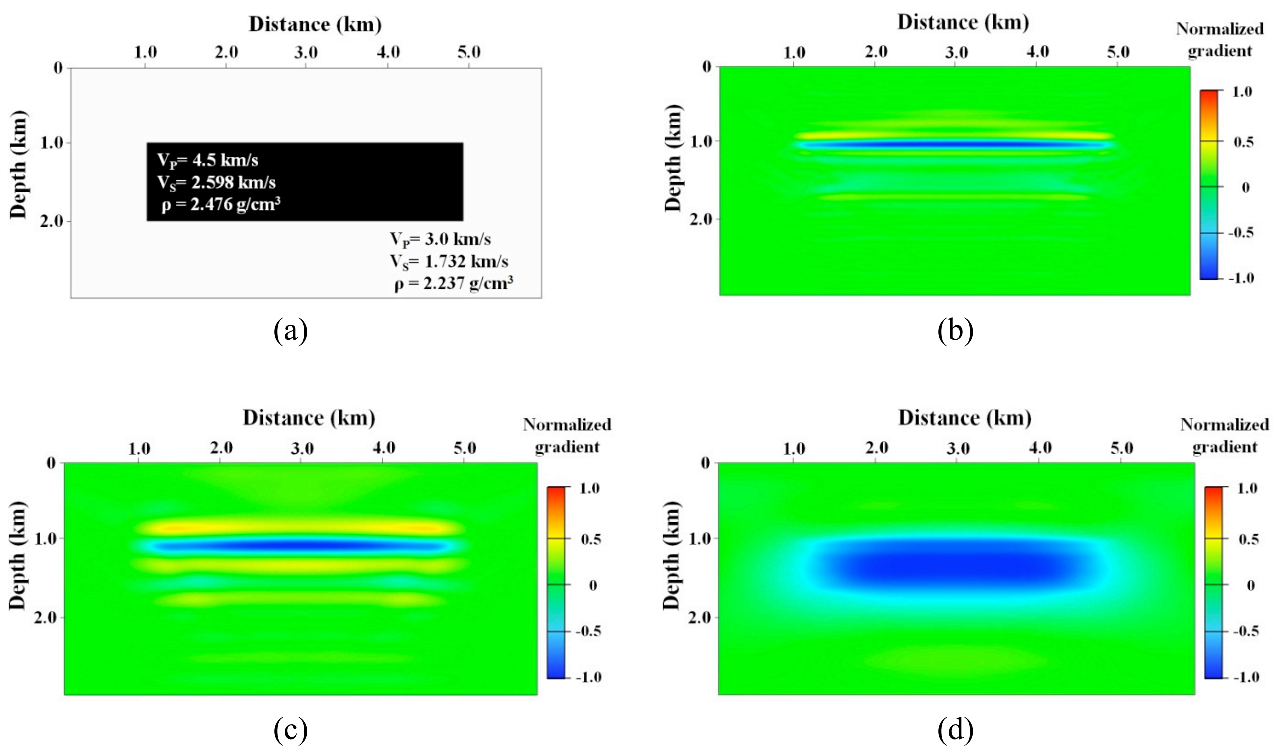
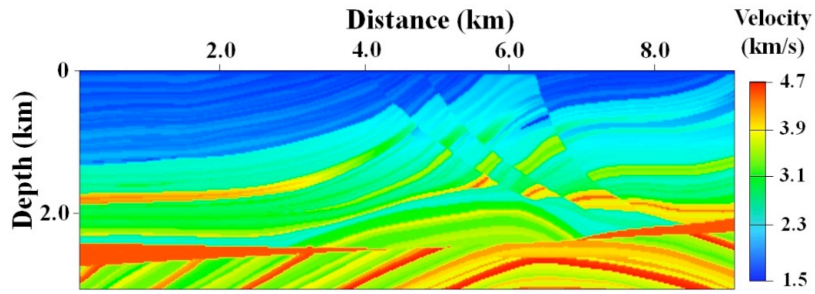
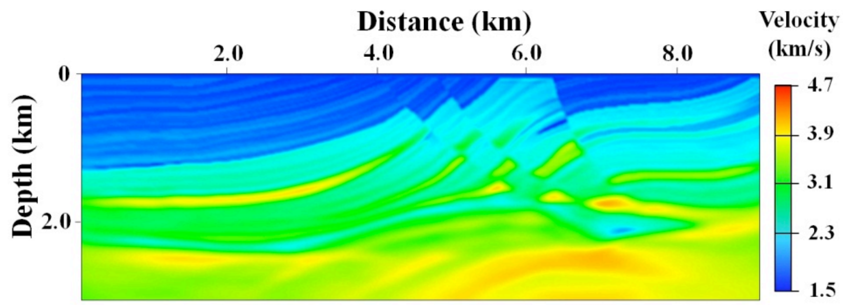


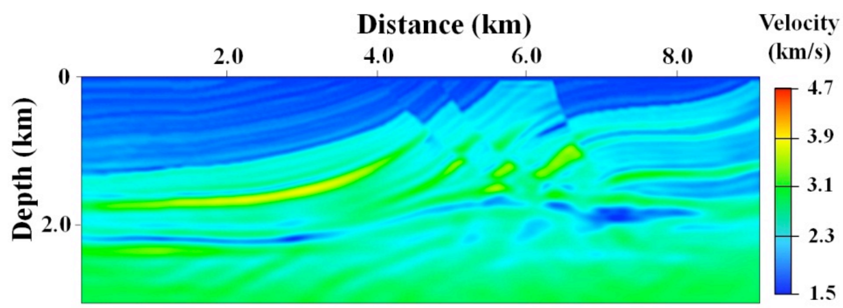
Figure 1: (a) The rectangular-shaped model and its gradient images of λ obtained at the first iteration by the (b) first and (c) second conventional methods and (d) the weighting method. Initial guesses for parameters are the same as those of the background media.



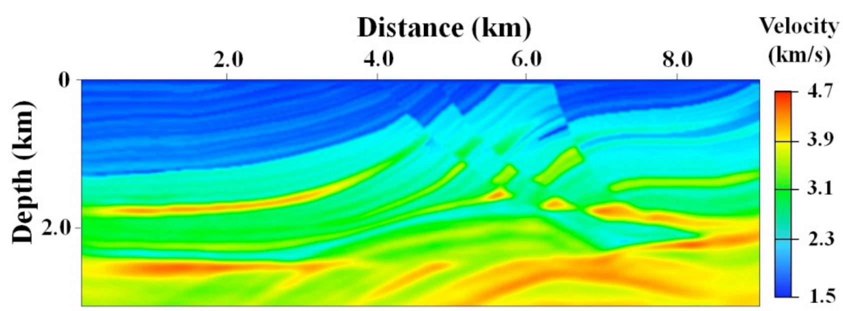
(a)



(b)



(c)



(d)

Figure 2: (a) The true P-wave velocity model of the modified version of the Marmousi-2 model and its inversion results obtained by the (b) first and (c) second conventional methods and (d) the weighting method. The initial guesses for P-wave velocity are poorly estimated, changing from 1.5 to 3.02 km/s, and those for S-wave velocity are determined from the P-wave initial guesses and Poisson's ratio (0.25).