

PSEUDO-HESSIAN MATRIX FOR THE LOGARITHMIC OBJECTIVE FUNCTION IN FULL WAVEFORM INVERSION

WANSOO HA¹, WOOKEEN CHUNG² and CHANGSOO SHIN¹

¹ *Department of Energy Systems Engineering, Seoul National University, Seoul 151-742, South Korea.*

² *Department of Energy and Resources Engineering, Korea Maritime University, Busan 606-791, South Korea. wkchung@hhu.ac.kr*

(Received December 27, 2011; revised version accepted March 29, 2012)

ABSTRACT

Ha, W.S., Chung, W.K. and Shin, C.S., 2012. Pseudo-Hessian matrix for the logarithmic objective function in full waveform inversion. *Journal of Seismic Exploration*, 21: 201-214.

The pseudo-Hessian matrix of the l_2 objective function has been used in previous inversion studies because it was observed to yield satisfactory results in both the frequency and Laplace domains. However, there are fundamental differences between the pseudo-Hessian of the logarithmic objective function and that of the l_2 objective function. In this study, we first derive the pseudo-Hessian matrix for full waveform inversion using the logarithmic objective function. We then compare the two pseudo-Hessian matrices in both Laplace- and frequency-domain waveform inversions using the logarithmic objective function. A Laplace-domain inversion using the pseudo-Hessian of the logarithmic objective function yields a better result than that obtained using the pseudo-Hessian of the l_2 objective function. On the other hand, frequency domain inversion results using the two pseudo-Hessians yield similar results. The differences originate from the different behaviour of wavefields in the Laplace and frequency domains.

KEY WORDS: pseudo-Hessian matrix, logarithmic objective function, frequency domain, waveform inversion.

INTRODUCTION

Full waveform inversion is a method used to build more accurate subsurface models required for seismic imaging. Since Tarantola's pioneering work (Tarantola, 1984), many authors have improved the inversion technique (Gauthier et al., 1986; Mora, 1987; Amundsen, 1991; Pratt et al., 1998; Operto et al., 2004; Brenders and Pratt, 2007; Ben-Hadj-Ali et al., 2008).

The objective function is one of the most important factors in a successful seismic inversion. Many researchers have suggested various objective functions for full waveform inversion. Claerbout and Muir (1973) suggested the l_1 objective function, which is more robust to outlier noise. Amundsen (1991) suggested the Cauchy objective function and the hyperbolic secant objective function, both of which are also robust to noise. Bube and Langan (1997) suggested a hybrid l_1/l_2 objective function. Symes and Carazzone (1991) advanced the differential semblance objective function, which exploits coherency in the seismic wavefield.

Shin and Min (2006) proposed a logarithmic objective function for full waveform inversion in the frequency domain. The logarithmic objective function enables one to invert amplitude and phase separately (Shin et al., 2007). This has been widely used, especially in the Laplace domain (Shin and Cha, 2008; Shin and Cha, 2009; Bae et al., 2010; Chung et al., 2010; Ha et al., 2010) due to its scaling effect between near-and far-offset wavefields.

To reduce the computational burden, Shin and Min (2006) used the pseudo-Hessian (Shin et al., 2001) instead of the full or approximate Hessian (Pratt et al., 1998). Because a pseudo-Hessian is an approximation of the approximate Hessian, inversion using the pseudo-Hessian typically requires a greater number of iterations before convergence. Nevertheless, the pseudo-Hessian is an efficient alternative to the approximate Hessian due to the vastly lower computational cost required to calculate the former. The pseudo-Hessian that Shin and Min (2006) used was derived from the l_2 objective function. Although the pseudo-Hessian yielded satisfactory results, we anticipate that better results might be obtained when using the logarithmic objective function derived from the pseudo-Hessian.

In this study, we derive a pseudo-Hessian of the logarithmic objective function. First, we review the full waveform inversion using the logarithmic objective function and the pseudo-Hessian of the l_2 objective function. Then, we derive the pseudo-Hessian for the logarithmic objective function. We then compare the logarithmic inversion results obtained using the pseudo-Hessian of the l_2 objective function with those of the logarithmic objective function using the Marmousi velocity model (Versteeg, 1994) in both the Laplace and frequency domains.

REVIEW OF THE WAVEFORM INVERSION USING A LOGARITHMIC OBJECTIVE FUNCTION

In the frequency domain, we can define the logarithmic objective function (Shin and Min, 2006; Shin et al., 2007) at each frequency as

$$E = \frac{1}{2} \sum_i^{N_s} \sum_j^{N_r} |\ln(\tilde{u}_{ij}/\tilde{d}_{ij})|^2, \quad (1)$$

where \tilde{u}_{ij} is the modeled wavefield, \tilde{d}_{ij} is the observed wavefield, N_s is the number of shots, and N_r is the number of receivers. By differentiating the objective function with respect to a model parameter m_k , we obtain the gradient direction as

$$\partial E/\partial m_k = \Re\left\{ \sum_i^{N_s} \sum_j^{N_r} [(1/\tilde{u}_{ij})(\partial\tilde{u}_{ij}/\partial m_k)[\ln(\tilde{u}_{ij}/\tilde{d}_{ij})]^*] \right\}, \quad (2)$$

where \Re means the real part of a complex number, $*$ denotes the complex conjugate, and $k = 1, 2, \dots, M$, where M is the number of model parameters. To calculate the gradient direction, it is essential to compute the partial derivative wavefield, $\partial\tilde{u}_{ij}/\partial m_k$. However, the direct calculation requires considerable computational resources. As an indirect method, the partial derivative wavefield in the gradient direction can be obtained by the product between the inverse of the impedance matrix and the virtual source vector (Pratt et al., 1998) as

$$\partial\tilde{u}_i/\partial m_k = \mathbf{S}^{-1}[-(\partial\mathbf{S}/\partial m_k)\tilde{u}_i] = \mathbf{S}^{-1}\mathbf{v}_{ik}, \quad (3)$$

where

$$\mathbf{S} = -\omega^2\mathbf{M} + \mathbf{K}, \quad (4)$$

$$\mathbf{v}_{ik} = -(\partial\mathbf{S}/\partial m_k)\tilde{u}_i.$$

ω is the angular frequency, \mathbf{M} is the mass matrix, \mathbf{K} is the stiffness matrix, \mathbf{S} is the complex impedance matrix, and \mathbf{v}_{ik} is the virtual source vector. To apply the adjoint method (Pratt et al., 1998), eq. (2) can be augmented by adding zero elements. Thus, the gradient direction can be rewritten as a matrix equation:

$$\partial E/\partial m_k = \Re\left\{ \sum_i^{N_s} [\mathbf{v}_{ik}^T(\mathbf{S}^{-1})^T\mathbf{r}_i] \right\}, \quad (5)$$

where

$$\mathbf{r}_i = \left\{ (1/\tilde{u}_{i,1}) [\ln(\tilde{u}_{i,1}/\tilde{d}_{i,1})]^* \dots (1/\tilde{u}_{i,N_r}) [\ln(\tilde{u}_{i,N_r}/\tilde{d}_{i,N_r})]^* 0 \dots 0 \right\}^T . \quad (6)$$

In the Gauss-Newton method, the approximated Hessian is required to normalize the gradient direction (Pratt et al., 1998). Instead, Shin and Min (2006) used the pseudo-Hessian (Shin et al., 2001) derived from the approximate Hessian of the l_2 objective function. The approximate Hessian matrix of the l_2 objective function can be expressed as (Pratt et al., 1998)

$$\mathbf{H}_{kl}^a = \Re \left\{ \sum_i^{N_s} \sum_j^{N_r} [(\tilde{u}_{ij}/\partial m_k)(\partial \tilde{u}_{ij}/\partial m_l)^*] \right\} . \quad (7)$$

By substituting (3) into (7), following Pratt et al. (1998), we can obtain the approximate Hessian matrix as a multiplication of matrices and vectors:

$$\mathbf{H}_{kl}^a = \Re \left\{ \sum_i^{N_s} [(-\partial \mathbf{S}/\partial m_k) \tilde{\mathbf{u}}_i]^T (\mathbf{S}^{-1})^T (\mathbf{S}^{-1})^* [(-\partial \mathbf{S}/\partial m_l) \tilde{\mathbf{u}}_i]^* \right\} . \quad (8)$$

The term $(\mathbf{S}^{-1})^T (\mathbf{S}^{-1})^*$, which is composed of the matrix multiplication of the inverse of the complex impedance matrix and the conjugated inverse of the complex matrix, is a diagonal dominant matrix (Shin et al., 2001). Therefore, the pseudo-Hessian of the l_2 objective function can be obtained as follows:

$$\mathbf{H}_{kk}^p = \Re \left\{ \sum_i^{N_s} [(-\partial \mathbf{S}/\partial m_k) \tilde{\mathbf{u}}_i]^T [(-\partial \mathbf{S}/\partial m_k) \tilde{\mathbf{u}}_i]^* \right\} . \quad (9)$$

We add a damping constant λ to the diagonal elements of the pseudo-Hessian for stabilization (Levenberg, 1994; Marquardt, 1963). Accordingly, the k -th model parameter at $(n+1)$ -th iteration can be calculated as

$$m_k^{n+1} = m_k^n - \alpha_n [\text{diag}(\mathbf{H}_{kk}^p) + \lambda]^{-1} (\partial E/\partial m_k) , \quad (10)$$

using the diagonal elements of the pseudo-Hessian matrix, where α_n is a step length (Shin et al., 2007).

PSEUDO-HESSIAN OF THE LOGARITHMIC OBJECTIVE FUNCTION

In the previous section, we reviewed the gradient direction of the logarithmic objective function and the pseudo-Hessian of the l_2 objective function. In this section, we will derive the pseudo-Hessian of the logarithmic

objective function.

The Hessian of the logarithmic objective function can be obtained by differentiating the gradient direction in (2) with respect to another model parameter m_l , ($l = 1, 2, \dots, M$) as follows:

$$\partial^2 E / \partial m_l \partial m_k \approx \Re \left\{ \sum_i^{N_s} \sum_j^{N_r} \left[\begin{aligned} & -(1/\tilde{u}_{ij}^2) (\partial \tilde{u}_{ij} / \partial m_l) (\partial \tilde{u}_{ij} / \partial m_k) [\ln(\tilde{u}_{ij} / \tilde{d}_{ij})]^* \\ & + (1/\tilde{u}_{ij}) (\partial^2 \tilde{u}_{ij} / \partial m_l \partial m_k) [\ln(\tilde{u}_{ij} / \tilde{d}_{ij})]^* \\ & + (1/\tilde{u}_{ij}) (\partial \tilde{u}_{ij} / \partial m_k) [(1/\tilde{u}_{ij}) (\partial \tilde{u}_{ij} / \partial m_l)]^* \end{aligned} \right] \right\}. \quad (11)$$

We obtain the approximate Hessian by ignoring the second derivative of the modeled wavefield (Pratt et al., 1998) as

$$\mathbf{H}_{kl}^a \approx \Re \left\{ \sum_i^{N_s} \sum_j^{N_r} \left[\begin{aligned} & -(1/\tilde{u}_{ij}^2) (\partial \tilde{u}_{ij} / \partial m_l) (\partial \tilde{u}_{ij} / \partial m_k) [\ln(\tilde{u}_{ij} / \tilde{d}_{ij})]^* \\ & + (1/\tilde{u}_{ij}) (\partial \tilde{u}_{ij} / \partial m_k) [(1/\tilde{u}_{ij}) (\partial \tilde{u}_{ij} / \partial m_l)]^* \end{aligned} \right] \right\}. \quad (12)$$

Unlike the approximate Hessian of the l_2 objective function in (7), that of the logarithmic objective function contains two terms. Generally, the second term dominates the approximate Hessian due to the residual in the first term. When performing a full waveform inversion, we first estimate the source wavelet information - both amplitude and phase - by the Newton method (Shin et al., 2007). If we estimate the amplitude of the source wavelet accurately, the first term can be small enough to be ignored when compared to the second term. Therefore, we omit the first term and the approximate Hessian can be approximated as

$$\mathbf{H}_{kl}^a \approx \Re \left\{ \sum_i^{N_s} \sum_j^{N_r} \left[(1/\tilde{u}_{ij}) (\partial \tilde{u}_{ij} / \partial m_k) [(1/\tilde{u}_{ij}) (\partial \tilde{u}_{ij} / \partial m_l)]^* \right] \right\}. \quad (13)$$

We obtain the approximate Hessian as a multiplication of matrices and vectors following the procedure we used to calculate that of the l_2 objective function:

$$\mathbf{H}_{kl}^p = \Re \left\{ \sum_i^{N_s} \left[[-(\partial S / \partial m_k) \tilde{\mathbf{u}}_i]^T (\mathbf{S}^{-1})^T \mathbf{D}_i^T \mathbf{D}_i^* (\mathbf{S}^{-1})^* [-(\partial S / \partial m_k) \tilde{\mathbf{u}}_i]^* \right] \right\}. \quad (14)$$

where \mathbf{D}_i is a diagonal matrix with components:

$$\mathbf{D}_{i,kk} = 1/\tilde{u}_{ik} . \quad (15)$$

We can obtain a pseudo-Hessian matrix by an approximation suggested by Shin et al., (2001) as

$$(\mathbf{S}^{-1})^T \mathbf{D}_i^T \mathbf{D}_i^* (\mathbf{S}^{-1})^* \approx \mathbf{D}_i^T \mathbf{D}_i^* . \quad (16)$$

The resulting pseudo-Hessian matrix can be expressed as

$$\begin{aligned} \mathbf{H}_{kk}^p &= \Re \left\{ \sum_i^{N_s} [[-(\partial \mathbf{S} / \partial m_k) \tilde{\mathbf{u}}_i]^T \mathbf{D}_i^T \mathbf{D}_i^* [-(\partial \mathbf{S} / \partial m_k) \tilde{\mathbf{u}}_i]^*] \right\} \\ &= \Re \left\{ \sum_i^{N_s} [[-(\partial \mathbf{S} / \partial m_k) \mathbf{c}]^T [-(\partial \mathbf{S} / \partial m_k) \mathbf{c}]^*] \right\} \\ &= N_s \Re [[-(\partial \mathbf{S} / \partial m_k) \mathbf{c}]^T [-(\partial \mathbf{S} / \partial m_k) \mathbf{c}]^*] . \end{aligned} \quad (17)$$

where $\mathbf{c}^T = [1, 1, \dots, 1]$. Note that the pseudo-Hessian matrix of the logarithmic objective function does not contain the modeled wavefield, unlike that of the l_2 objective function in (8). Accordingly, the k -th model parameter can be updated by using (10) with the diagonal elements of the pseudo-Hessian matrix in (17).

NUMERICAL EXAMPLES

We now compare the behaviour of the two different pseudo-Hessian matrices, where the first is derived from the l_2 objective function, and the second is derived from the logarithmic objective function. We perform a full waveform inversion using the two different pseudo-Hessian matrices in both the Laplace and Frequency domains. The Marmousi model (Versteeg 1994) is used for our synthetic tests; Fig. 1 shows the Marmousi P-wave velocity model.

First, we test the inversions in the Laplace domain (Shin and Cha, 2008). In this model, 71 shots are used with 289 receivers. The interval between shots is 128 m, whereas that between receivers is 32 m. The model grid used a 16 m grid spacing. We inverted the data using 6 Laplace damping constants ranging from 2 to 12 (with an interval of 2).

Fig. 2 shows the inverted P-wave velocities obtained using the pseudo-Hessian of the l_2 objective function and logarithmic objective function after 50 iterations. The results from both inversions are similar; however, the error curve shows that the inversion using the pseudo-Hessian of the logarithmic objective function converged faster (Fig. 3).

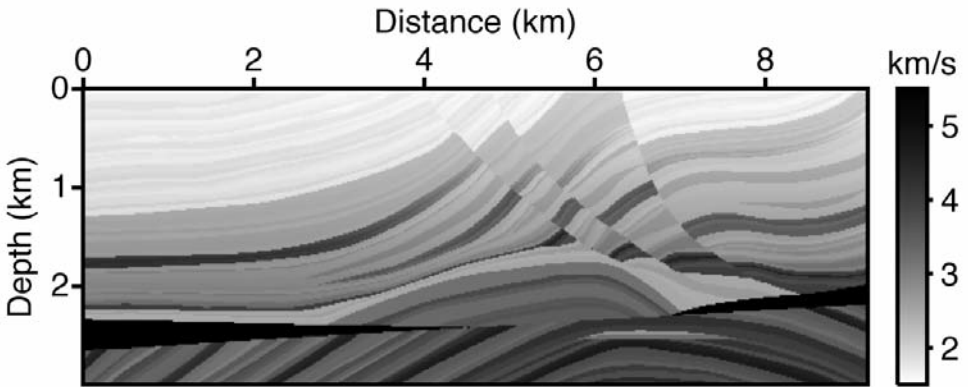


Fig. 1. The Marmousi velocity model.

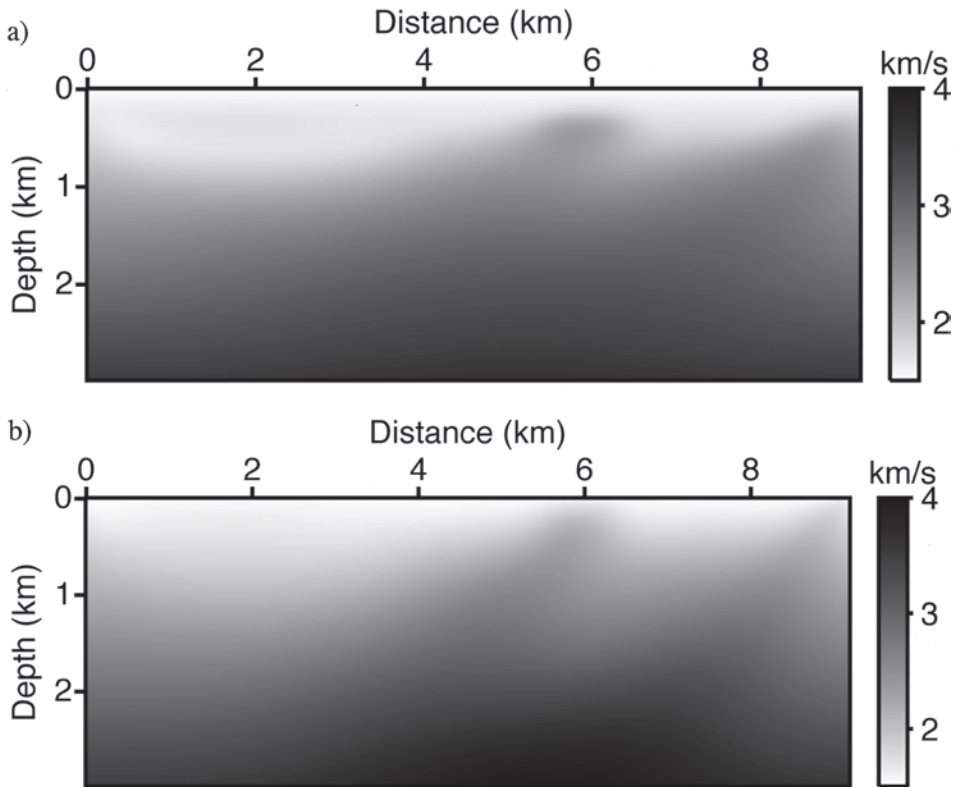


Fig. 2. Inversion results obtained in the Laplace domain using (a) the pseudo-Hessian of the l_2 objective function and (b) that of the logarithmic objective function.

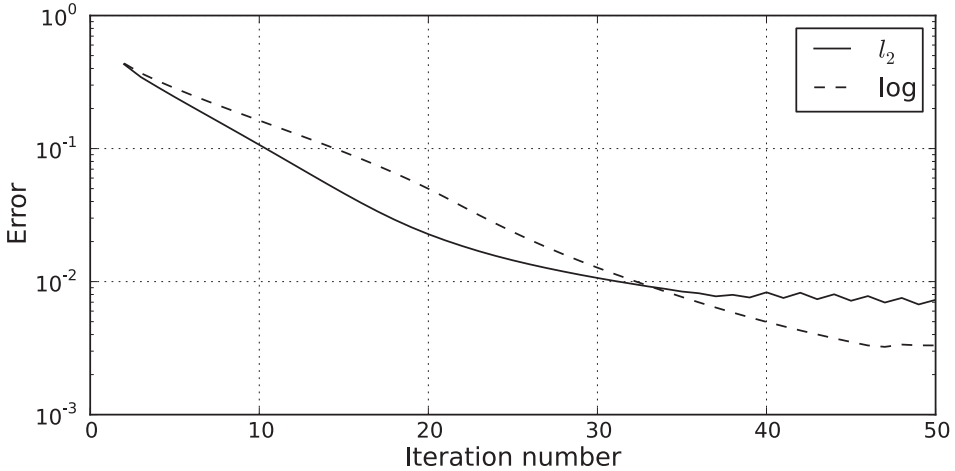


Fig. 3. The error histories of the Laplace-domain inversions.

The difference in behaviour of the two pseudo-Hessian matrices in the Laplace domain originates from the modeled wavefield in the pseudo-Hessian of the l_2 objective function (9). Because the wavefield in the Laplace domain decays exponentially as the offset between the source and receiver increases (Table 1), the pseudo-Hessian of the l_2 objective function has large values only near source locations (Fig. 4a). Conversely, the pseudo-Hessian of the logarithmic objective shows the appropriate shape, which can scale correctly in the gradient direction (Fig. 4b). Profiles extracted from the two pseudo-Hessians show significantly different behaviour from each other (Fig. 5).

Table 1. Green's functions of the 2D and 3D acoustic wave equations in the Laplace domain. s is the damping constant, c is the P-wave velocity in the media, \mathbf{x} and \mathbf{x}' are receiver and source positions.

Dimension	Green's function
2D	$(1/2\pi)K_0[(s/c) \mathbf{x} - \mathbf{x}']$
3D	$(1/4\pi \mathbf{x} - \mathbf{x}')\exp[-(s/c) \mathbf{x} - \mathbf{x}']$

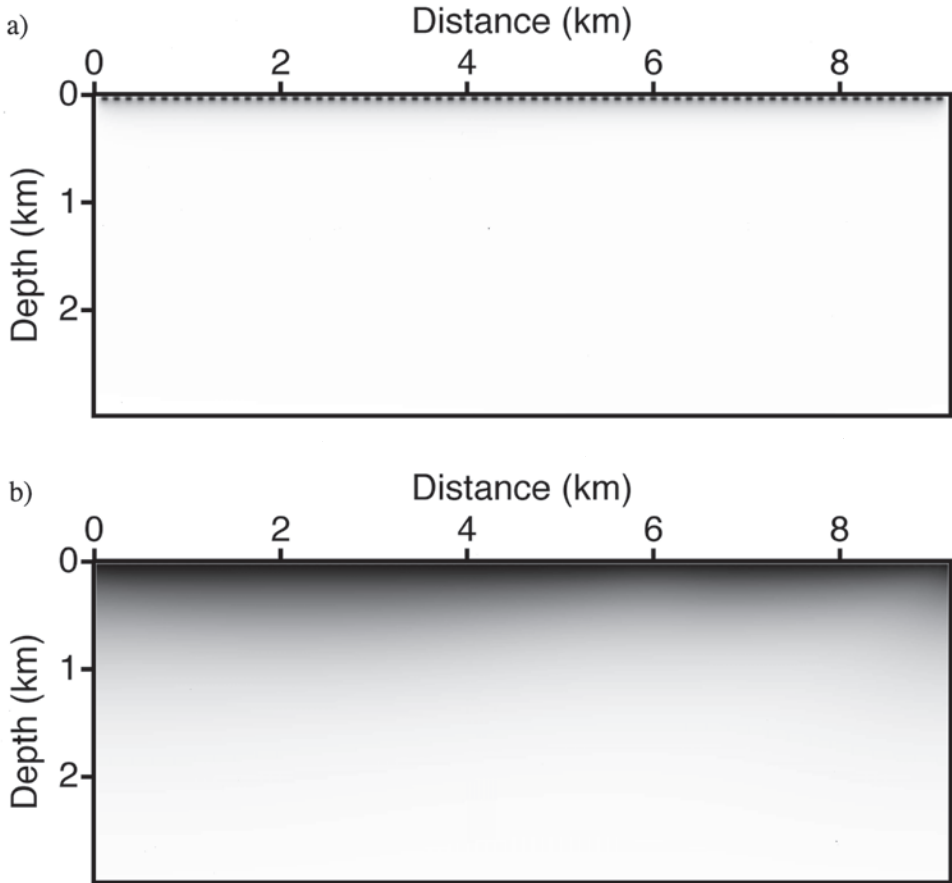


Fig. 4. The pseudo-Hessians of the (a) l_2 objective function and (b) logarithmic objective function after 10 iterations using a Laplace damping constant of 4.

We have also tested the pseudo-Hessians in the frequency domain (Shin and Min 2006). We generated observed data in the frequency domain and inverted the data using the logarithmic objective function. In this model, 575 shots were used in 577 receivers. The interval between both shots and receivers was 16 m. The grid spacing size was also 16 m. We inverted the data using 25 frequencies ranging from 0.4 to 10.0 Hz (with an interval of 0.4 Hz) using the result of the Laplace-domain inversion (Fig. 2b) as the starting model.

Fig. 6 shows the inversion results obtained using the pseudo-Hessians of the l_2 objective function and logarithmic objective function after 500 iterations. The results are very similar in both cases. The error curve shows that the difference between the two pseudo-Hessians during convergence is not significant (Fig. 7).

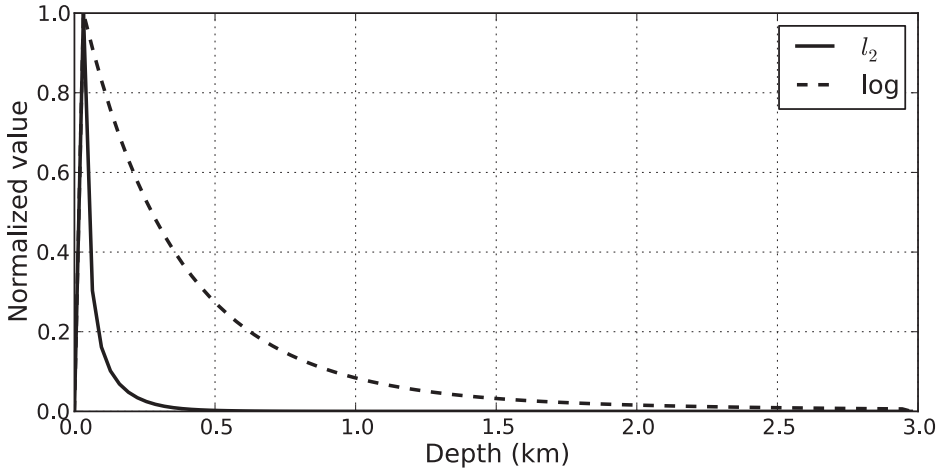


Fig. 5. Normalized profiles of the pseudo-Hessians shown in Fig. 4. The profiles were extracted at a location 6.4 km from the left edge.

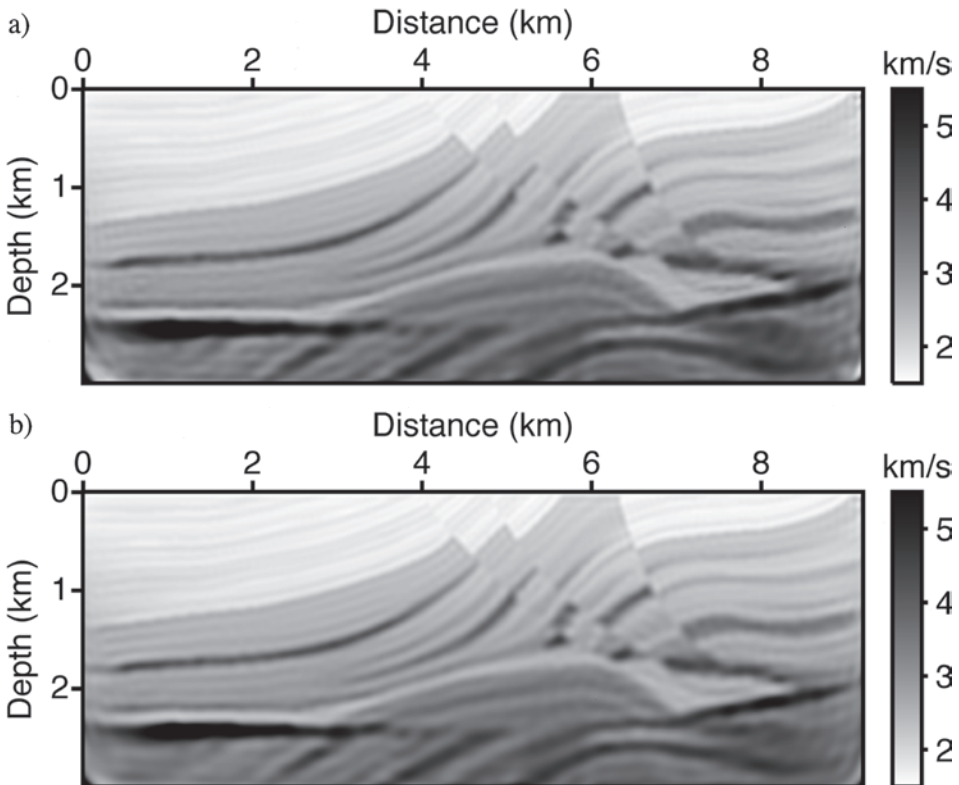


Fig. 6. Inversion results obtained in the frequency domain using (a) the pseudo-Hessian of the l_2 objective function and (b) that of the logarithmic objective function.

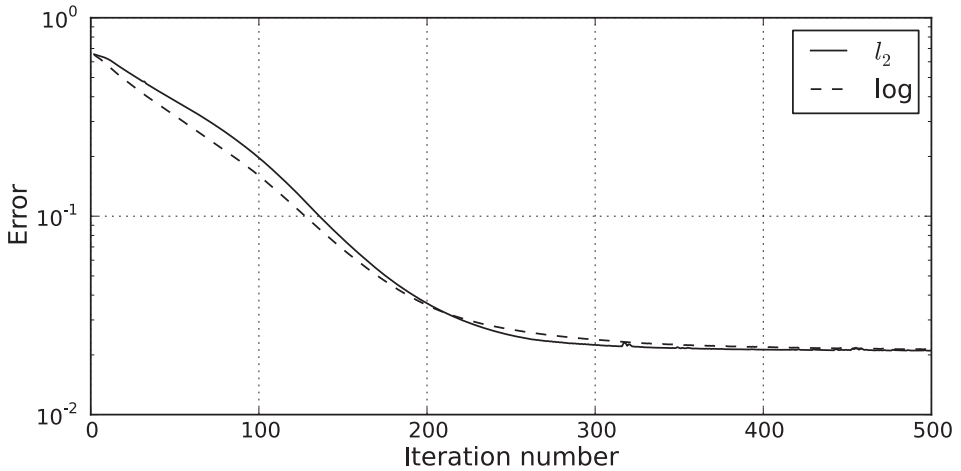


Fig. 7. The error histories of the frequency-domain inversions.

Fig. 8 shows the two pseudo-Hessians in the frequency domain. Unlike those of the Laplace domain, they show similar variations. Because the Green's function in the frequency domain is inversely proportional to the offset between the shot and receiver (Table 2), the effect of the modeled wavefield in the pseudo-Hessian of the l_2 objective function (9) is limited. Profiles extracted from the pseudo-Hessians in the frequency domain show similar tendencies (Fig. 9).

Table 2. Green's functions of the 2D and 3D acoustic wave equations in the frequency domain. ω is the angular frequency, c is the P-wave velocity in the media, \mathbf{x} and \mathbf{x}' are receiver and source positions.

Dimension	Green's function
2D	$(i/4)H_0^{(1)}[(\omega/c) \mathbf{x} - \mathbf{x}']$
3D	$(1/4\pi \mathbf{x} - \mathbf{x}')\exp[i(\omega/c) \mathbf{x} - \mathbf{x}']$

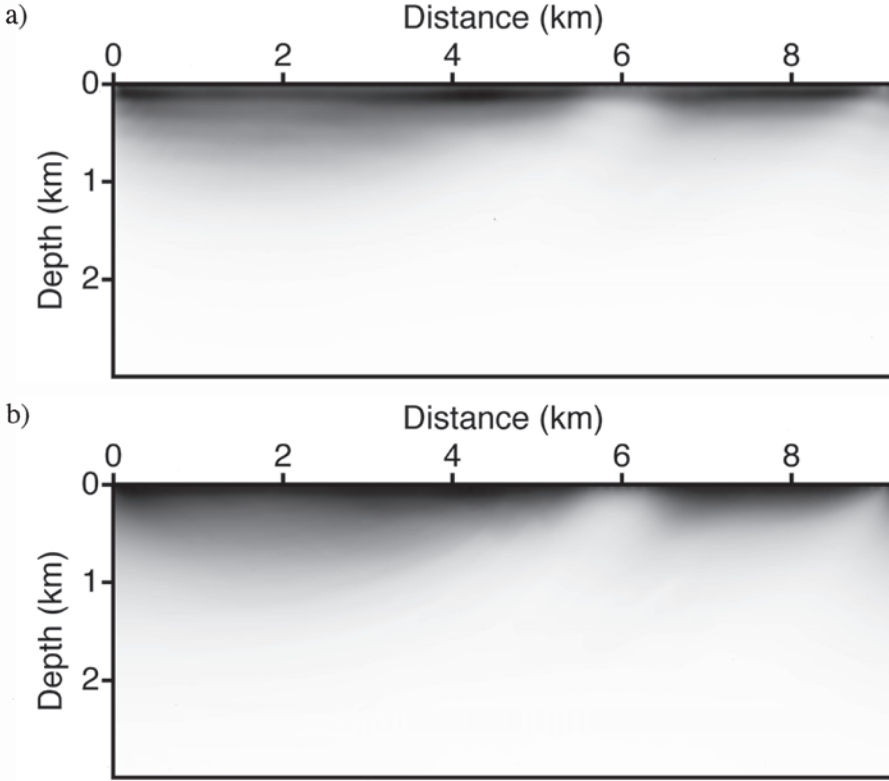


Fig. 8. The pseudo-Hessians of the (a) l_2 objective function and (b) logarithmic objective function after 10 iterations at the frequency of 5.0 Hz.

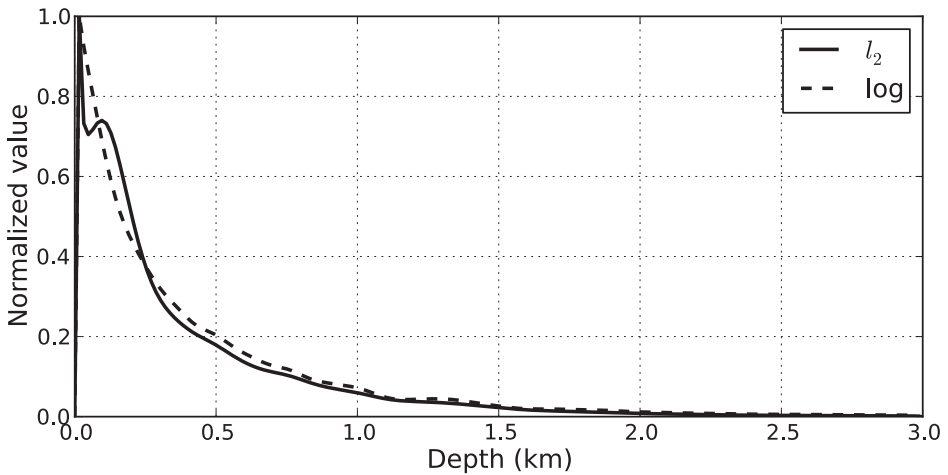


Fig. 9. Normalized profiles of the pseudo-Hessians shown in Fig. 8. The profiles were extracted at a location 6.4 km from the left edge.

CONCLUSIONS

We derived the pseudo-Hessian of the logarithmic objective function for full waveform inversion. This can easily be applied and implemented to a logarithmic full waveform inversion in the frequency or Laplace domains. This workflow requires no additional computational cost for calculation when compared with that of the l_2 objective function. Numerical examples of logarithmic inversion in the Laplace domain show that the pseudo-Hessian of the logarithmic objective function yields a better result than that of the l_2 objective function. On the other hand, the two pseudo-Hessians give similar results in the frequency domain. The difference between the results of Laplace and frequency domains originates from the modeled wavefield in the pseudo-Hessian of the l_2 objective function.

Because the pseudo-Hessian is an approximation to the approximate Hessian, inversion using the pseudo-Hessian requires more iterations to converge when compared to inversions using the approximate or full Hessian. However, the pseudo-Hessian is a more efficient choice than the approximate or full Hessian due to the huge computational cost required to calculate the latter. In summary, the pseudo-Hessian derived in this study can be widely and efficiently used in inversion using the logarithmic objective function, both in the Laplace and frequency domains.

ACKNOWLEDGEMENTS

This work was supported by the Energy Efficiency & Resources of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Knowledge Economy (No. 2010T100200376).

REFERENCES

- Amundsen, L., 1991. Comparison of the least-squares criterion and the Cauchy criterion in frequency-wavenumber inversion. *Geophysics*, 56: 2027-2035.
- Bae, H.S., Shin, C., Cha, Y.H., Choi, Y. and Min, D.J., 2010. 2D acoustic-elastic coupled waveform inversion in the Laplace domain. *Geophys. Prosp.*, 58: 997-1010.
- Ben-Hadj-Ali, H., Operto, S. and Virieux, J., 2008. Velocity model-building by 3D frequency domain, full-waveform inversion of wide-aperture seismic data. *Geophysics*, 73: VE101-VE117.
- Brenders, A. and Pratt, R., 2007. Efficient waveform tomography for lithospheric imaging: Implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophys. J. Internat.*, 168: 152-170.
- Bube, K.P. and Langan, R.T., 1997. Hybrid $l(1)/l(2)$ minimization with applications to tomography. *Geophysics*, 62: 1183-1195.

- Chung, W., Shin, C. Pyun, S. and Calandra, H., 2010. 2D elastic waveform inversion in the Laplace domain. Expanded Abstr., 80th Ann. Internat. SEG Mtg., Denver: 1059-1064.
- Claerbout, J. and Muir, F., 1973. Robust modeling with erratic data. *Geophysics*, 38: 826-844.
- Gauthier, O., Virieux, J. and Tarantola, A., 1986. Two-dimensional nonlinear inversion of seismic waveforms: numerical results. *Geophysics*, 51: 1387-1403.
- Ha, W., Pyun, S., Yoo, J. and Shin, C., 2010. Acoustic full waveform inversion of synthetic land and marine data in the Laplace domain. *Geophys. Prosp.*, 58: 1033-1047.
- Levenberg, K., 1944. A method for the solution of certain nonlinear problems in least squares. *Quart. J. Appl. Mathemat.*, 2: 164-168.
- Marquardt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Industr. Appl. Mathemat.*, 11: 431-441.
- Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, 52: 1211-1228.
- Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A. and Dellapos-Aversana, P., 2004. Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale travelttime and waveform inversions: A case study. *Geophys. Prosp.*, 52: 625-51.
- Pratt, R., Shin, C. and Hicks, G., 1998. Gauss-Newton and full Newton methods in frequencyspace seismic waveform inversion. *Geophys. J. Internat.*, 133: 341-62.
- Shin, C., Jang, S. and Min, D., 2001. Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophys. Prosp.*, 49: 592-606.
- Shin, C. and Min, D., 2006. Waveform inversion using a logarithmic wavefield. *Geophysics*, 71: R31-R42.
- Shin, C. and Cha, T., 2008. Waveform inversion in the Laplace domain. *Geophys. J. Internat.*, 173: 922-931.
- Shin, C. and Cha, T., 2009. Waveform inversion in the Laplace-Fourier domain. *Geophys. J. Internat.*, 177: 1067-1079.
- Shin, C., Pyun, S. and Bednar, J.B., 2007. Comparison of waveform inversion, Part 1: Conventional wavefield vs. logarithmic wavefield. *Geophys. Prosp.*, 55: 449-464.
- Symes, W. and Carazzone, J., 1991. Velocity inversion by differential semblance optimization. *Geophysics*, 56: 654-63.
- Tarantola, A., 1984. Inversion of seismic-reflection data in the acoustic approximation. *Geophysics*, 49: 1259-1266.
- Versteeg, R., 1994. The Marmousi experience: Velocity model determination on a synthetic complex data set. *The Leading Edge*, 13: 927-936.