Robust seismic waveform inversion

Taeyoung Ha [1] and Wookeen Chung [2] and Changsoo Shin [2]

1) *National Institute for Mathematical Sciences 385-16, Doryong-dong, Yuseong-gu,Daejeon 305-340, Korea*

2) *School of Civil, Urban and Geosystem Engineering, Seoul National Univ., San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea*

Corresponding Author : Taeyoung Ha, tha@nims.re.kr

ABSTRACT

For seismic imaging and inversion, the inverted image depends on how we define the object function. ℓ^1-norm is more robust than ℓ^2-norm. However, it is difficult to apply the Newton-type algorithm directly because the partial derivative for ℓ^1-norm has a singularity. To overcome the difficulties of singularities, Huber function given by hybrid ℓ^1/ℓ^2-norm is used. We tested the robustness of our new object function with several noisy data set.

INTRODUCTION

It is generally well known that ℓ^1-norm is more robust than ℓ^2-norm for seismic noise[2,6,3]. However, ℓ^1-norm has a singular problem when the values of residual vanish. Even if the values of residuals are not zero, the numerical inversion process goes to failure at very small residual. In order to achieve the robustness and the stability in waveform inversion, we construct the objective function consisting of Huber function. Huber function is composed that ℓ^1-norm is taken when the value of residuals is large and ℓ^2-norm is chosen when the value of residuals is small (choosing ℓ^1-norm or ℓ^2-norm is decided by a criterion[1]). By combination of ℓ^1-norm and ℓ^2-norm, we can obtain more robust results for several various noises than that obtained by minimizing the object function composed of ℓ^2-norm.

HUBER FUNCTION

We define the object function with ℓ^1 and ℓ^2 error function suggested by Huber[4]:

$$M_\epsilon(r) = \begin{cases} \dfrac{|r|^2}{2\varepsilon}, & |r| \le \varepsilon, \\ |r| - \dfrac{\varepsilon}{2}, & |r| > \varepsilon, \end{cases} \tag{1}$$

where ε is threshold between the ℓ^1 and ℓ^2 norms for real value r. In complex value $r = a + bi$, a, b real values, ℓ^1-norm is $|a| + |b|$, and ℓ^2-norm is $\sqrt{a^2 + b^2}$. Therefore, we define the function proposed by Huber in complex plane as follows :

$$M_{\epsilon,c}(r) = \begin{cases} \dfrac{|r|_2^2}{2\varepsilon}, & |r|_2 \le \varepsilon, \\ |r|_1 - \dfrac{\varepsilon}{2}, & |r|_2 > \varepsilon. \end{cases} \tag{2}$$

Since $|r|_1$ is larger than $|r|_2$, the singularities of the function (2) don't appear.

THE OBJECT FUNCTION

Suppose we have r_N experimental observations record at a subset of nodal points corresponding to the receiver locations and define the model in terms of a parameter set $\mathbf{p}$. Then, from the artificial model using finite-element or finite-difference method, we can obtain model response calculated at the receiver locations. In forward modeling of wave equation at frequency domain, the matrix equation can be expressed by

$$\mathbf{K}\hat{\mathbf{u}} = \widehat{\mathbf{f}}, \tag{3}$$

where $\mathbf{K}$ is the complex impedance matrix, $\hat{\mathbf{u}}$ is the Fourier transformed wavefield and $\widehat{\mathbf{f}}$ is the source vector. Also, $\mathbf{K}$ and $\hat{\mathbf{u}}$ are dependent on model parameter $\mathbf{p}$.

Now, we define the object function

$$\mathbf{E}(\mathbf{p}) = \sum_{\omega}^{N_\omega} \sum_{s}^{N_s} \sum_{r}^{N_r} M_{\varepsilon,c}(r_{\omega,s,r}) \tag{4}$$

where $r_{\omega,s,r} = u_{\omega,s,r} - d_{\omega,s,r}$ and N_ω, N_s and N_r are the number of frequencies, sources and receivers. For $|r|_2 \le \varepsilon$, we can easily obtain the gradient of the object function $\mathbf{E}$ as like [5]. The gradient of the objective function with respect to the kth velocity parameter p_k is

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial p_k} &= \mathrm{Re} \sum_{s,r,\omega} \left[\frac{\partial \widehat{u}_{sr}}{\partial p_k} \right] (\widehat{u}_{\omega sr} - \widehat{d}_{\omega sr})^* \\ &= \mathrm{Re} \sum_{s,\omega} \left[\frac{\partial \widehat{u}_{s,r_1}}{\partial p_k} \cdots \frac{\partial \widehat{u}_{s,r_N}}{\partial p_k} \right]^{\mathbf{T}} \begin{bmatrix} r^*_{r_1} \\ \vdots \\ r^*_{r_N} \end{bmatrix}. \end{aligned} \tag{5}$$

which $\cdot^*$ denotes the conjugate complex.

And, if $|r|_2 > \varepsilon$,

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial p_k} &= \sum_{s,\omega,r} \left[\mathrm{Re} \left[\frac{\partial \widehat{u}}{\partial p_k} \right] \mathrm{Re}[r_{sgn}] + \mathrm{Im} \left[\frac{\partial \widehat{u}}{\partial p_k} \right] \mathrm{Im}[r_{sgn}] \right] \\ &= \mathrm{Re} \sum_{s,\omega} \left[\frac{\partial \widehat{u}_{s,r_1}}{\partial p_k} \cdots \frac{\partial \widehat{u}_{s,r_N}}{\partial p_k} \right]^{\mathbf{T}} \begin{bmatrix} r^*_{1,sgn} \\ \vdots \\ r^*_{r_N,sgn} \end{bmatrix}, \end{aligned} \tag{6}$$

where the signed function sgn is defined as follows: for real value a,

$$\mathrm{sgn}(a) = \begin{cases} 1, & \text{if } a > 0, \\ -1, & \text{if } a < 0, \\ 0, & \text{if } a = 0. \end{cases}$$

and, $r_{r_k,sgn} = \text{sgn}(\text{Re}(r_{r_k})) + i\text{sgn}(\text{Im}(r_{r_k}))$. The gradient of object function consisting of Huber function is

$$\frac{\partial \mathbf{E}}{\partial p_k} = \text{Re}\sum_{s,\omega}\left[\frac{\partial \mathbf{u}_{sr}}{\partial p_k}\right]^{\mathbf{T}} \tilde{\mathbf{r}}^* \tag{7}$$

where $\mathbf{r} = (r_1, \cdots, r_{N_r})$ and

$$r_k = \begin{cases} \frac{1}{\varepsilon}(u-d), & |r|_2 \leq \varepsilon, \\ r_{k,sgn}, & |r|_2 > \varepsilon. \end{cases} \tag{8}$$

for $k = 1, \cdots, N_r$.

BACKPROPAGATION ALGORITHM

By zero padding argument in the matrix equation (7),

$$\nabla_{p_k}\mathbf{E} = \text{Re}\sum_{s,\omega}\left[\frac{\partial \widehat{u}_{s,r_1}}{\partial p_k} \cdots \frac{\partial \widehat{u}_{s,r_N}}{\partial p_k} \cdots \frac{\partial \widehat{u}_{s,N_r}}{\partial p_k}\right]^{\mathbf{T}} \begin{bmatrix} \mathbf{r}^* \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{9}$$

And, taking the derivative of the matrix equation (3) resulting from Galerkin approximation process of wave equation[5], we obtain

$$\mathbf{K}\frac{\partial \widehat{\mathbf{u}}}{\partial p_k} = -\frac{\partial \mathbf{K}}{\partial p_k}\widehat{\mathbf{u}} \qquad \text{or} \qquad \frac{\partial \widehat{\mathbf{u}}}{\partial p_k} = \mathbf{K}^{-1}\mathbf{v}_k$$

with

$$\mathbf{v}_k = -\frac{\partial \mathbf{K}}{\partial p_k}\widehat{\mathbf{u}},$$

where $\mathbf{v}_k$ is defined as the virtual source vector with respect to the kth model parameter. Therefore, the equation (9) becomes

$$\begin{aligned} \nabla_{p_k}\mathbf{E} &= \text{Re}\sum_{s,\omega}\left[\frac{\partial \mathbf{u}_{sr}}{\partial p_k}\right]^{\mathbf{T}} \tilde{\mathbf{r}}^* = \text{Re}\sum_{s,\omega}\left[\mathbf{K}^{-1}\mathbf{v}_k(\omega)\right]^{\mathbf{T}} \tilde{\mathbf{r}}^* \\ &= \text{Re}\sum_{s,\omega}\mathbf{v}_k^{\mathbf{T}}\left[\mathbf{K}^{-1}\right]^{\mathbf{T}} \tilde{\mathbf{r}}^* = \text{Re}\sum_{s,\omega}\mathbf{v}_k^{\mathbf{T}}\left[\mathbf{K}^{-1}\tilde{\mathbf{r}}^*\right], \end{aligned} \tag{10}$$

where $\tilde{\mathbf{r}} = (\mathbf{r}^*\ 0\ \cdots\ 0)^T$.

REFERENCES

1. K.P. Bube and R.T. Langan. Hybrid ℓ^1/ℓ^2 minimization with applications to tomography. *Geophysics*, 62:1183–1195, 1997.
2. J. Clarbout and F. Muir. Robust modeling with erratic data. *Geophysics*, 38(5):826–844, 1973.

3. E. Crase, A. Pica, M. Noble, J. McDonald, and A. Tarantola. Robust elastic nonlinear waveform inversion:application to real data. *Geophysics*, 55(5):527–538, 1990.
4. Peter J. Huber. Robust regression: asymptotics, conjectures and Monte Carlo. *Ann. Statist.*, 1:799–821, 1973.
5. C. Shin, K. Yoon, K. J. Marfurt, K. Park, D. Yang, H. Lim, S. Chung, and S. Shin. Efficient calculation of a partial-derivative wavefield using reciprocity for seismic imaging and inversion. *Geophysics*, 66:1856–1863, 2001.
6. A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics (SIAM), 2005.