A reconstruction algorithm for full waveform inversion based on sparsity regularization

Hongsun Fu and Mingyue Ma

College of Science, Dalian Maritime University, Dalian 116026, China

Email: mamingyue@dlmu.edu.cn

Abstract. Full waveform inversion (FWI) is an important ill-posed problem in mathematical physics inverse problem. In order to address considerable over-smoothing of the Tikhonov regularization solution, sparsity regularization method is applied to FWI. The sparsity of the ‘inhomogeneity’ with respect to a certain basis is priori assumed. The proposed approach is motivated by a Tikhonov functional incorporating a sparsity-promoting $l_1$-penalty term, and it allows us to obtain quantitative results when the assumption is valid. We have proposed an applicable sparse reconstruction algorithm of iterative soft shrinkage type. This algorithm is compared with a conventional reconstruction approach on the basis of smoothness regularization. The numerical examples showed that very good reconstruction results can be obtained by the proposed algorithms.

1. Introduction

FWI was first proposed by Tarantola [1] in time domain. Since it is able to objectively reflect the law of seismic wave propagation and realize high-precision imaging of complicated model, FWI has gradually gained extensive research and attention in recent years [2, 3, 4]. For all the inverse problems, it can be summed up as the optimization problem [5]. At present, sparsity regularization method has been widely used in FWI [6, 7].

A more important fact is that the ill-posedness is common obstacle encountered in frequency-domain FWI. Tikhonov regularization is an efficient approach to handle the ill-posed problem. However, the classical $l_2$-norm regularization leads to overly smooth minimizers. Therefore, alternatives are required that can perform much better. One effective approach of circumventing this difficulty is to incorporate a sparsity-promoting $l_1$-norm penalty into the Tikhonov functional. Recently, the idea of sparsity has been popularized by [8]. The basic idea is to incorporate a sparsity-promoting $l_1$-penalty into the Tikhonov functional.

The goal of this paper is to develop a new sparse reconstruction algorithm [9] and to demonstrate its excellent performances in seismic imaging. We shall consider an approach that combines a ‘data fitting’ term $F(m) = \frac{1}{2} \sum_{k,j} \|K(m) - f^\delta\|_{H_2}^2$ with an $l_1$-penalty. We aim at deriving an applicable sparse reconstruction algorithm of iterative soft shrinkage type by approximately minimizing this functional. Let $m_0$ be a known background. Then the standard iterative soft shrinkage algorithm [10, 11] was adopted to reconstruct the inhomogeneity $\tilde{m}$ ($\tilde{m} = m - m_0$).
\[ m_{i+1} = m_i + \tau S \left( m_i - \tau K'(m_i)^* (K(m_i) - f^\delta) \right), \]
\[ m_{i+1} = m_i + \bar{m}_{i+1}. \tag{1} \]

Where \( S \) is the soft shrinkage operator. It consists of two steps: a gradient descent of the functional \( F(m) \) with a step size \( \tau_i \) and a proximal mapping step. Intuitively, the latter promotes the sparsity of the reconstruction as it zeroes out small coefficients. In practice, the proximal gradient method with a fixed step size favours the classical thresholded Landweber method. This method is known to be slowly convergent. Nevertheless, it is relatively easy to implement and can yield reliable sparse approximations. In this paper, we adopt an adaptive step size selection by the Barzilai–Borwein (BB) rule to accelerate the algorithm.

We capitalize on recent works, notably [12, 13, 14], to propose an efficient sparsity reconstruction algorithm for the inversion problems. The rest of the paper is as follows. In Section 2, we describe the basic mathematical model for full waveform inversion, sparsity constraints, as well as the resulting Tikhonov functional. Then, we describe the intricacies of the proposed reconstruction technique motivated by minimizing such functional in Section 3. In Section 4, we enumerate some numerical examples to prove the feasibility of the algorithm. Finally, we conclude in Section 5.

2. Objective functional with sparsity constraints

We consider operator equations

\[ K(m) = f, \tag{2} \]

Where \( K : H_1 \to H_2 \) is a nonlinear operator between Hilbert spaces \( H_1 \) and \( H_2 \). The related inverse problem involves the computation of an approximation to the solution of this operator equation from given noisy data \( f^\delta \) with

\[ \|f - f^\delta\|_{H_2} \leq \delta. \tag{3} \]

We are particularly interested in the case of ill-posed equations, which need stabilization by regularization methods for computing stable approximations.

Equation (2) and equation (3) lead us to consider the following minimization problem (Tikhonov regularization with sparsity constraint), which consists of a discrepancy term and an \( l_1 \)–penalty. The discrepancy \( F \) under consideration is the standard least squares fitting, that is,

\[ F(m) = \frac{1}{2} \sum_{k,j} \left\| K(m) - f^\delta \right\|^2_{H_2}. \tag{4} \]

Let \( \bar{m}^\dagger = m^\dagger - m_0 \) be the inhomogeneity of the unknown velocity \( m^\dagger \), and we seek an approximation from \( \bar{m} \) to \( \bar{m}^\dagger \). An orthonormal basis is introduced and then the corresponding \( l_1 \)–norm of the inhomogeneity \( \bar{m} \) can be given by

\[ \|\bar{m}\| = \sum_{i \in \mathbb{N}} \omega_i |\langle \bar{m}, \varphi_i \rangle|. \tag{5} \]

The weight sequence \( \omega_i \geq \omega_{\min} > 0 \), which allows regularization for each coefficient independently.

It is now widely accepted that the \( l_1 \)–penalty can effectively enforce the a priori knowledge of a sparse representation of \( \bar{m} \) with respect to \( \{\varphi_i\}_{i \in \mathbb{N}} \) [15]. Consequently, we arrive at the Tikhonov functional

\[ \Phi(m) = F(m) + \alpha \|\bar{m}\|. \tag{6} \]

Where \( \alpha > 0 \) is the regularization parameter.
3. Numerical algorithm

3.1. Soft shrinkage operator

We first define the real valued shrinkage function \( S_\omega(x) = \text{sgn}(x) \max(|x| - \omega, 0), \omega > 0 \), \( (7) \)

Now, we discuss efficient numerical schemes for approximating minimizers of the objective functional (6), which is given by

\[
\min_m \Phi(m) := F(m) + \alpha \| m - m_0 \| = F(m) + J(\tilde{m}). \quad (8)
\]

The minimization problem (8) is complex because of nonsmoothness of the \( l_1 \)-norm penalty, high-degree nonlinearity of the discrepancy \( F(m) \). The classical approach for designing gradient descent methods can be conveniently derived from the first-order optimality condition for the minimizer, which turns the minimization problem into a fixed point equation. The first-order optimality condition for a minimizer \( m \) for the functional \( \Phi(m) \) is given by

\[
0 \in \partial \Phi(m) = F'(m) + \partial J(\tilde{m}) \quad (9)
\]

Multiplying by \( \tau \) and adding \( m - m_0 \) on both sides yields a fixed point relation which has to be satisfied for a minimizer \( m \) and for all \( \tau \in \mathbb{R} \)

\[
m - m_0 + \tau F'(m) + \tau \partial J(\tilde{m}) \quad (10)
\]

Reordering the fixed point relation yields

\[
m - m_0 - \tau F'(m) \in m - m_0 + \tau \partial J(\tilde{m}) \quad (11)
\]

Then we get

\[
\tilde{m} - \tau F'(m) \in \tilde{m} + \tau \partial J(\tilde{m}) = (I + \tau \partial J)(\tilde{m}) \quad (12)
\]

We can turn this into an iteration by requesting

\[
\tilde{m}_{i+1} = (I + \tau \partial J)(\tilde{m}_i) \quad (13)
\]

The expression on the right-hand side is inverted by the proximal mapping for \( J \), and hence yields the iteration

\[
\tilde{m}_{i+1} = S_\omega (\tilde{m}_i - \tau F'(m_i)) \quad (14)
\]

With an appropriately chosen step size \( \tau \) and a shrinkage operator \( S \).

We have the following formula for the gradient of \( F \).

\[
F'(m_i) = K'(m_i) (K(m_i) - f^\delta) \quad (15)
\]

So its solution involves a gradient descent step [17] followed by a soft shrinkage step. The complete procedure is listed in Algorithm 1.

3.2. Barzilai-Borwein step length

Usual gradient algorithms, for example, steepest descent methods, suffer from slow convergence. One of the recent breakthroughs in enhancing its convergence behaviour is due to Barzilai and Borwein [18, 19], who, in 1988, developed an ingenious step size selection rule (BB rule). It performs much better than the standard steepest descent algorithm. The basic idea is to mimic the Hessian with \( \tau H \) over the most recent steps so that \( \tau H (\tilde{m}_i - \tilde{m}_{i-1}) \approx F'(m_i) - F'(m_{i-1}) \) approximately holds. This equation may not have a solution, so it is solved in a least squares sense, that is,

\[
\tau_i = \arg \min_\tau \| \tau (m_i - m_{i-1}) - (F'(m_i) - F'(m_{i-1})) \|^2, \quad (16)
\]

This gives rise to
\[ \tau_i = \frac{\langle \overline{m}_i - \overline{m}_{i-1}, F'(m_i) - F'(m_{i-1}) \rangle}{\langle \overline{m}_i - \overline{m}_{i-1}, \overline{m}_i - \overline{m}_{i-1} \rangle}. \] (17)

In addition, the initial step size is constrained to \([\tau_{\text{min}}, \tau_{\text{max}}]\) and the iteration is terminated when \(\tau_i\) falls below \(\tau_{\text{stop}}\). Alternatively, when the maximum number \(I\) of iterations is reached, the algorithm is terminated [20].

**Table 1. Sparse reconstruction algorithm.**

| Step | Description |
|------|-------------|
| 1:   | Give \(m_0\) and an orthonormal basis \(\{\varphi_i\}_{i \in \mathbb{N}}\), take the weights \(\omega_i\), the regularization parameter \(\alpha\). |
| 2:   | for \(i=1,2,\ldots, I\) do |
| 3:   | Compute \(m_i = \overline{m}_i + m_0\); |
| 4:   | Compute \(F'(m_i)\); |
| 5:   | Determine the step size \(\tau_i\); |
| 6:   | Update inhomogeneity \(\overline{m}_{i+1}\) by \(\overline{m}_{i+1} = \overline{m}_i - \tau_i F'(m_i)\); |
| 7:   | Threshold \(\overline{m}_{i+1}\) by \(S_{\alpha}(\overline{m}_{i+1})\); |
| 8:   | Check stopping criterion. |
| 9:   | end for |
| 10:  | output an approximate minimizer \(\overline{m}\). |

### 4. Numerical experiments

In this section, we present some numerical results to illustrate the power of sparsity regularization and the new algorithms. These examples are generated using a finite-difference discretization of the Helmholtz equation on a regular grid with PML-boundaries on all sides and we invert for the velocity field \(m\). Here, the sparse representation of the velocity model \(m\) is carried in the curvelet domain. The step size bounds \(\tau_{\text{min}}\) and \(\tau_{\text{max}}\) is set to \(10^{-2}\) and \(10^{2}\), respectively, with \(\tau_{\text{stop}} = 10^{-3}\), the weights \(\omega_i = \omega_{\text{min}} = 10^{-2}\), and the maximum number of iterations \(I = 100\). All numerical experiments are conducted by adding \(\frac{\|\text{noise}\|}{\|\text{data}\|} = 3\%\) Gaussian white noise to the data. The value of the regularization parameter \(\alpha\) is set to \(\alpha : \frac{\|\text{noise}\|_2}{\|\text{data}\|_2}\). We consider a 2D synthetic experiment with a roughly 51 by 51 sized model and a mesh width \(h\) equal to 20m. There are 51 sources and 51 receivers located near the surface and equally distributed over the horizontal coordinate. The data are generated at 5 different frequencies ranging from 5 to 25 Hertz [21].

#### 4.1. Single rectangular inclusion

The exact velocity field consists of a homogeneous background and one single rectangular inclusion. We use an estimate of the smooth background as our initial guess \(m_0\). The velocity of the background and inhomogeneities are 2000m/s and 2100m/s, respectively.

The exact velocity model is shown in figure 1(a), and the reconstructions by the smoothness and sparsity (table 1) regularization are shown in figure 1(b) and (c), respectively. The former is overall smooth, which is typical of such smoothness regularization. In contrast, we observe that the sparsity reconstruction is more localized at the correct location, and the estimate is close to the true value, which is quantitatively correct.
Figure 1. Reconstructions for example 4.1. Here and later, we use the notation (a) exact, (b) smoothness, and (c) sparsity to denote the exact velocity, the reconstruction by smoothness regularization, and the reconstruction by the proposed reconstruction algorithm, respectively.

4.2. Multiple rectangular inclusions

The exact velocity field consists of a homogeneous background and two convex inclusions. The velocity of the background and inhomogeneities are 2000m/s, 2100m/s and 1900m/s, respectively.

Figure 2. Reconstructions for example 4.2. (a) exact, (b) smoothness, and (c) sparsity.

The exact velocity model is shown in figure 2(a), and the reconstructions by the smoothness and sparsity (Algorithm 1) regularization are shown in figure 2(b) and (c), respectively. Multiple inclusions are challenging for some existing numerical algorithms that can only recover the convex envelope. Nonetheless, both the smoothness and sparsity regularization can give reasonable
reconstructions. In the smoothness reconstruction, the two inclusions are correctly identified. However, their shape is grossly smoothed. These drawbacks are partially remedied by the proposed reconstruction technique with sharper localization of the inclusions and with more accurate estimate of their velocity.

4.3. Single rectangular inclusion

The exact velocity field consists of a homogeneous background and two circular rectangular inclusions. The velocity of the background and inhomogeneities are 2000 m/s, 2100 m/s, and 1900 m/s, respectively.

Circular rectangular inclusions present one of the most challenging objects to recover. The exact velocity model is shown in figure 3(a), and the reconstructions by the smoothness and sparsity (Algorithm 1) regularization are shown in figure 3(b) and (c), respectively. We observe that two circular rectangular inclusions stand out clearly in the sparsity reconstruction than in the smooth reconstruction. Further, the magnitude of the velocity is reasonably reconstructed in the sparsity reconstruction.

![Figure 3](image.png)

**Figure 3.** Reconstructions for example 4.3. (a) exact, (b) smoothness, and (c) sparsity.

5. Conclusions

We have presented a novel image reconstruction technique for full waveform inversion based on sparsity concepts. It is adapted from the classical iterative soft shrinkage algorithm, and its main ingredients include soft shrinkage iteration, and an adaptive step size selection based on the BB rule. Finally, in order to prove the feasibility of this algorithm, this algorithm is compared with a conventional reconstruction approach on the basis of smoothness regularization. The results indicate that the proposed technique can yield quantitatively acceptable reconstructions in terms of the locations as well as the velocity magnitudes of the inclusions.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (Grant Nos. 41474102, 41304092), and the Fundamental Research Funds for the Central Universities (Grant No. 3132017051). Authors sincerely thank the editors and the anonymous reviewers for hard work and careful review of this article.
References
[1] Tarantola A 2005 Inverse problem theory and methods for model parameter estimation Society for Industrial and Applied Mathematics
[2] Zhang W S, Luo J and Teng J W 2015 Frequency multiscale full-waveform velocity inversion Chinese Journal of Geophysics 58(1) 216-228
[3] Liu Y, Yang J, Chi B and Dong L 2015 An improved scattering-integral approach for frequency-domain full waveform inversion Geophysical Journal International 202(3) 1827-1842
[4] Metivier L, Brossier R, Virieux J and Operto S 2013 Full waveform inversion and the truncated newton method SIAM Journal on Scientific Computing 35(2) B401-B437
[5] Virieux J and Operto S 2009 An overview of full waveform inversion in exploration geophysics Geophysics 74(6) WCC1-WCC26
[6] Bonesky T, Bredies K, Lorenz D A and Maass P 2007 A generalized conditional gradient method for nonlinear operator equations with sparsity constraints Inverse Problems 23(5) 2041-2058
[7] Herrmann F J and Li X 2012 Efficient least-squares imaging with sparsity promotion and compressive sensing Geophysical Prospecting 60(4) 696-712
[8] Daubechies I, Defrise M and Mol C D 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint Communications on Pure & Applied Mathematics 57(11) 1413-1457
[9] Jin B, Khan T and Maass P 2012 A reconstruction algorithm for electrical impedance tomography based on sparsity regularization International Journal for Numerical Methods in Engineering 89(3) 337-353
[10] Bredies K and Lorenz D A 2008 Linear Convergence of Iterative Soft-Thresholding Journal of Fourier Analysis & Applications 14(5-6) 813-837
[11] Jin B and Maass P 2012 Sparsity regularization for parameter identification problems Inverse Problems 28(12) 3001
[12] Bredies K, Lorenz D A and Maass P 2009 A generalized conditional gradient method and its connection to an iterative shrinkage method Computational Optimization and Applications 42(2) 173-193
[13] Fu H S, Liu H B, Han B, et al. 2017 A proximal iteratively regularized Gauss-Newton method for nonlinear inverse problems Journal of Inverse and Ill-posed Problems 25(3) 341-356
[14] Wright S J, Nowak R D and Figueiredo M 2016 Sparse reconstruction by separable approximation IEEE Transactions on Signal Processing 57(7) 2479-2493
[15] Daubechies I, Defrise M and De Mol C 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint Communications on Pure & Applied Mathematics 57(11) 1413-1457
[16] Lorenz D A, Maass P and Muoi P Q 2012 Gradient descent for Tikhonov functionals with sparsity constraints: theory and numerical comparison of step size rules Journal of Volcanology & Geothermal Research 39(3) 237-263
[17] Daubechies I, Fornasier M and Loris I 2008 Accelerated Projected Gradient Method for Linear Inverse Problems with Sparsity Constraints Journal of Fourier Analysis & Applications 14(5-6) 764-792
[18] Barzilai J and Borwein J M 1988 Two-point step size gradient methods IMA Journal of Numerical Analysis 8(1) 141-148
[19] Dai Y-H, Hager WW, Schittkowski K and Zhang H 2006 The cyclic Barzilai-Borwein method for unconstrained optimization IMA Journal of Numerical Analysis 26(3) 604-627
[20] Grippo L, Lampariello F and Lucidi S 1986 A nonmonotone line search technique for Newton’s method SIAM Journal on Numerical Analysis 23(4) 707-716
[21] Leeuwen T V and Herrmann F J 2013 Mitigating local minima in full-waveform inversion by expanding the search space Geophysical Journal International 195(1) 661-667