Efficient Full Waveform Inversion Subject To A Total Variation Constraint

Yudai INADA, Shingo TAKEMOTO and Shunsuke ONO
Institute of Science Tokyo

アブストラクト Full waveform inversion (FWI) aims to reconstruct subsurface properties from observed seismic data. Since FWI is an ill-posed inverse problem, appropriate regularizations or constraints are useful approaches to achieve accurate reconstruction. The total variation (TV) -type regularization or constraint is widely known as a powerful prior that models the piecewise smoothness of subsurface properties. However, the optimization problem of the TV-type regularized or constrained FWI is difficult to solve due to the non-linearity of the observation process and the non-smoothness of the TV-type regularization or constraint. Conventional approaches to solve the problem rely on an inner loop and/or approximations, resulting in high computational cost and/or inappropriate solutions. In this paper, we develop an efficient algorithm with neither an inner loop nor approximations to solve the problem based on a primal-dual splitting method. We also demonstrate the effectiveness of the proposed method through experiments using the SEG/EAGE Salt and Overthrust Models.

1 Introduction

Full waveform inversion (FWI)[1], [2] aims to reconstruct subsurface properties from observed seismic data. Since the observed seismic data are generated from the subsurface properties, FWI is formulated as an inverse problem. However, it is ill-posed, and the quality of the solution depends heavily on the initial model provided[2].

In general, FWI reconstruct the subsurface properties by solving an optimization problem that minimizes the squared error between the observed data and the modeled data. To stabilize the inversion and achieve more accurate reconstruction, other formulations have been proposed[3]–[8]. Adding regularizations or constraints to the objective function has also been proposed, such as Tikhonov regularization[9], Total Variation (TV)[10], and Total Generalized Variation (TGV)[11]. For example, studies have used regularization of Tikhonov[12], TV[13], directional TV[14], high-order TV[15], and TGV[16]. TV has also been used as a constraint[17]–[19].

Recently, neural networks (NNs) have also been proposed to estimate subsurface properties directly from observed data[20]–[23]. However, NNs require a large amount of observed data as training data, and since the dimensions of the observed data may vary depending on the observation method, training data must be prepared not only for each target prior but also for each observation method. Therefore, reconstruction using optimization is still useful.

Also, the value of the squared error between the observed data and the modeled data, which is the objective function of FWI, changes depending on the observation method. Therefore, The parameters of the regularization must be chosen accordingly. However, the TV constraint has the advantage that the parameters can be determined only from prior knowledge of the subsurface properties[24].

In conventional methods that apply the TV constraint to FWI[17], [18], when optimizing the usual FWI objective function using quasi-Newton methods such as the L-BFGS method, the parameter updates are computed to satisfy the constraints. Here, another optimization is required to compute the updates that satisfy the constraints, resulting in a double loop and high computational cost. Moreover, approximations are used in the process of introducing constraints, such as treating nonlinear transformations as linear or enforcing constraints outside the optimization method.

In this paper, we develop an efficient algorithm based on a primal-dual splitting method to solve the TV-constrained FWI problem with neither an inner loop nor approximations. We also demonstrate the effectiveness of the proposed method through experiments using the SEG/EAGE Salt and Overthrust Models.

2 PRELIMINARIES

2.1 Mathematical Tools

Throughout this paper, we denote vector and matrix by boldface lowercase letter (e.g., $\boldsymbol{x}$) and boldface uppercase letter (e.g., $\boldsymbol{X}$), respectively. The operator l_X norm of a vector and matrix is denoted by $\|\cdot\|_X$.

For $\boldsymbol{x} \in \mathbb{R}^{\mathrm{N}}$, the mixed $l_{1,2}$ norm is defined as follows:

$$\|\boldsymbol{x}\|_{1,2} \coloneqq \sum_{\mathfrak{g} \in \mathfrak{G}} \|\boldsymbol{x}_{\mathfrak{g}}\|_2 \tag{1}$$

where $\mathfrak{G}$ is a set of groups, $\boldsymbol{x}_{\mathfrak{g}}$ is the $\mathfrak{g}$-th group of $\boldsymbol{x}$.

For $\boldsymbol{x} \in \mathbb{R}^{\mathrm{N}}$, the total variation (TV)[10] is defined as follows:

$$\mathrm{TV}(\boldsymbol{x}) \coloneqq \|\mathbf{D}\boldsymbol{x}\|_{1,2} = \sum_{i=1}^{\mathrm{N}} \sqrt{d_{h,i}^2 + d_{v,i}^2} \tag{2}$$

here, $d_{h,i}$ and $d_{v,i}$ are the horizontal and vertical differences of the i-th element of $\boldsymbol{x}$, respectively, when vector $\boldsymbol{x}$ is considered as a matrix.

For proper lower-semicontinuous convex function $f \in \mathbb{R}^N \to \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^N$, the convex conjugate function is defined as follows:

$$f^*(\boldsymbol{x}) \coloneqq \sup_{\boldsymbol{y} \in \mathbb{R}^N} \left\{\boldsymbol{y}^T \boldsymbol{x} - f(\boldsymbol{y})\right\} \tag{3}$$

For a set $C \subset \mathbb{R}^N$ and $\boldsymbol{x} \in \mathbb{R}^N$, the indicator function is defined as follows:

$$\iota_C(\boldsymbol{x}) \coloneqq \begin{cases} 0 & \text{if } \boldsymbol{x} \in C \\ \infty & \text{otherwise} \end{cases} \tag{4}$$

For $\gamma > 0$, $f \in \mathbb{R}^N \to \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^N$, the proximity operator is defined as follows:

$$\mathrm{prox}_{\gamma f}(\boldsymbol{x}) \coloneqq \underset{\boldsymbol{y} \in \mathbb{R}^N}{\mathrm{argmin}} \left\{ f(\boldsymbol{y}) + \frac{1}{2\gamma}\|\boldsymbol{y} - \boldsymbol{x}\|_2^2 \right\} \tag{5}$$

The proximity operator for the convex conjugate function is expressed as follows[25, Theorem 3.1 (ii)]:

$$\mathrm{prox}_{\gamma f^*}(\boldsymbol{x}) = \boldsymbol{x} - \gamma \mathrm{prox}_{f/\gamma}(\boldsymbol{x}/\gamma) \tag{6}$$

The proximity operator for the indicator function is expressed as follows.

$$\mathrm{prox}_{\gamma \iota_C(\cdot)}(\boldsymbol{x}) = P_C(\boldsymbol{x}) \coloneqq \underset{\boldsymbol{y} \in C}{\mathrm{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|_2 \tag{7}$$

The proximity operator for the box constraint is expressed as follows.

$$P_{[a,b]^N}(\boldsymbol{x}) = \min(\max(\boldsymbol{x}, a), b) \tag{8}$$

The proximity operator for the l_1 norm upper bound constraint is expressed as follows[26](faster algorithms are also proposed[27]):

$$P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_1 \le \alpha\}}(\boldsymbol{x}) = \mathrm{SoftThrethold}(\boldsymbol{x}, \beta) \tag{9}$$

where

$$\boldsymbol{x}_{\mathrm{abs}} = \mathrm{abs}(\boldsymbol{x})$$
$$\boldsymbol{y} = \mathrm{sort}_{\mathrm{desc}}(\boldsymbol{x}_{\mathrm{abs}})$$
$$\beta' = \max\{\frac{1}{i}((\sum_{j=1}^{i} \boldsymbol{y}_j) - \alpha) \mid i = 1, \dots, N\}$$
$$\beta = \max\{\beta', 0\}$$

The proximity operator for the $l_{1,2}$ norm upper bound constraint is expressed as follows[28]:

$$P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_{1,2} \le \alpha\}}(\boldsymbol{x}) = [\boldsymbol{p}_1^T, \dots, \boldsymbol{p}_N^T]^T \tag{10}$$

where

$$\boldsymbol{p}_i = \begin{cases} 0 & \text{if } \|\boldsymbol{x}_i\|_2 = 0 \\ \boldsymbol{\beta}_i \frac{\boldsymbol{x}_i}{\|\boldsymbol{x}\|_2} & \text{otherwise} \end{cases}$$
$$\boldsymbol{\beta} = P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_1 \le \alpha\}}([\|\boldsymbol{x}_1^T\|_2, \dots, \|\boldsymbol{x}_N^T\|_2])$$

2.2 Primal-Dual Splitting Algorithm

The Primal-Dual Splitting (PDS) Algorithm[29]–[32] is applied to the following problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^N} \{f(\boldsymbol{x}) + g(\boldsymbol{x}) + h(\boldsymbol{L}\boldsymbol{x})\} \tag{11}$$

where $\boldsymbol{L} \in \mathbb{R}^{\mathrm{M} \times \mathrm{N}}$, f is differentiable convex function and g, h are convex functions whose proximity operator can be computed efficiently(proximable).

The PDS algorithm solve by iteratively updating the following:

$$\boldsymbol{x}^{(k+1)} = \mathrm{prox}_{\gamma_1 g}(\boldsymbol{x}^{(k)} - \gamma_1(\nabla f(\boldsymbol{x}^{(k)}) + \boldsymbol{L}^T \boldsymbol{y}^{(k)})) \tag{12}$$

$$\boldsymbol{y}^{(k+1)} = \mathrm{prox}_{\gamma_2 h^*}(\boldsymbol{y}^{(k)} + \gamma_2 \boldsymbol{L}(2\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)})) \tag{13}$$

where γ_1, γ_2 are step sizes. for more details and convergence rates, please refer to[31].

2.3 Full Waveform Inversion

In general, an objective function of FWI is defined as follows:

$$E(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{u}_{\mathrm{obs}} - \boldsymbol{u}_{\mathrm{cal}(\boldsymbol{m})}\|_2^2 \tag{14}$$

where $\boldsymbol{m}$ is velocity model, $\boldsymbol{u}_{\mathrm{obs}}$ is the observed seismic data, and $\boldsymbol{u}_{\mathrm{cal}(\boldsymbol{m})}$ is the calculated seismic data with the velocity model.

3 Proposed Method

We solve the objective function of the TV and box constrained FWI using PDS, which is expressed by the following equation:

$$E(\boldsymbol{m}) \quad \text{s.t.} \quad \mathrm{TV}(\boldsymbol{m}) \leq \alpha \ , \ \boldsymbol{m} \in [a, b]^N \tag{15}$$

where α is the upper bound of the $l_{1,2}$ norm, and a and b are the lower and upper bounds of the velocity model, respectively.

To apply PDS, the constraints are integrated into the objective function as indicator functions:

$$E(\boldsymbol{m}) + \iota_{\|\cdot\|_{1,2} \leq \alpha}(\mathbf{D}\boldsymbol{m}) + \iota_{[a,b]^N}(\boldsymbol{m}) \tag{16}$$

The steps obtained by applying PDS are as follows:

$$\begin{aligned}
\widetilde{\boldsymbol{m}}^{(k+1)} &= \boldsymbol{m}^{(k)} - \gamma_1 (\nabla E(\boldsymbol{m}^{(k)}) + \boldsymbol{D}^T \boldsymbol{y}^{(k)}) \\
\boldsymbol{m}^{(k+1)} &= P_{[a,b]^N}(\widetilde{\boldsymbol{m}}^{(k+1)}) \\
\widetilde{\boldsymbol{y}}^{(k+1)} &= \boldsymbol{y}^{(k)} + \gamma_2 \boldsymbol{D}(2\boldsymbol{m}^{(k+1)} - \boldsymbol{m}^{(k)}) \\
\boldsymbol{y}^{(k+1)} &= \widetilde{\boldsymbol{y}}^{(k+1)} - \gamma_2 P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_{1,2} \leq \alpha\}}(\frac{1}{\gamma_2} \widetilde{\boldsymbol{y}}^{(k+1)})
\end{aligned}$$

Here, $P_{[a,b]^N}(\cdot)$ and $P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_{1,2} \leq \alpha\}}(\cdot)$ are given by(8) and (10), respectively, and the gradient $\nabla E(\cdot)$ can be computed using the adjoint-state method [33].

4 EXPERIMENTS

the SEG/EAGE Salt and Overthrust Models を用いて実験を行います.

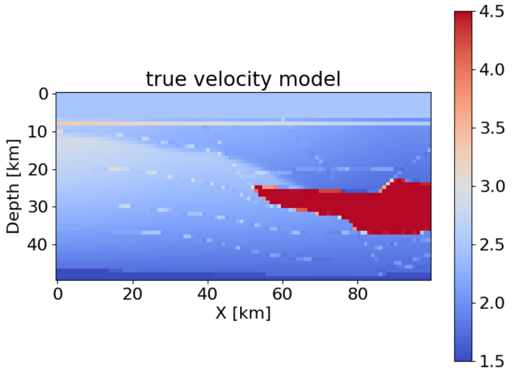

図 1: true

5 CONCLUSION

aaa

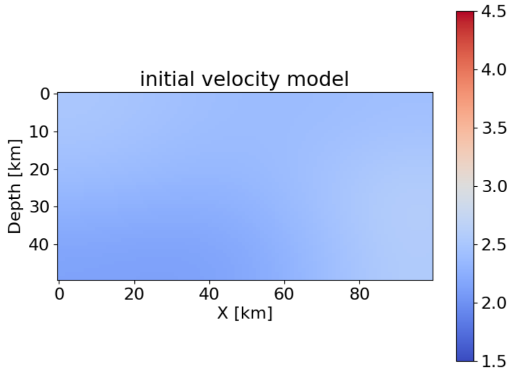

図 2: initial

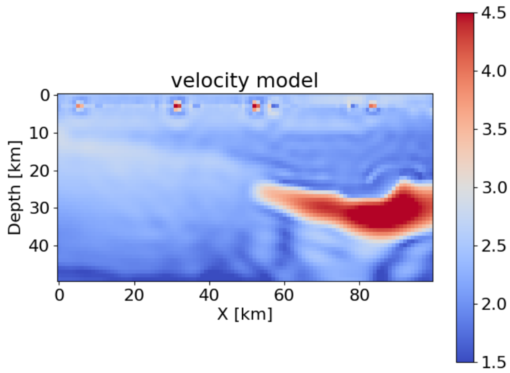

図 3: gradient

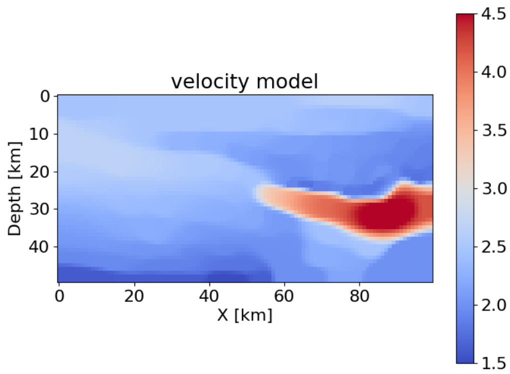

図 4: pds

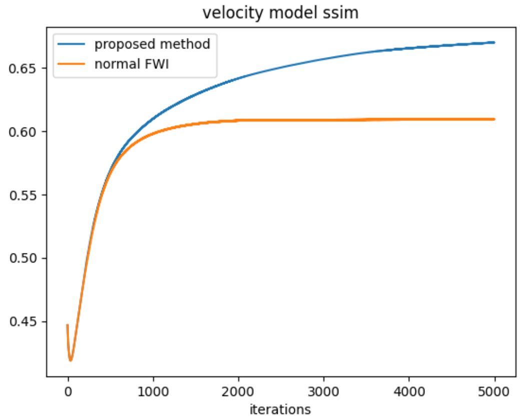

図 5: ssim

参考文献

[1] A. Tarantola, "Inversion of seismic reflection data in the acoustic approximation," *Geophysics*, vol. 49, no. 8, pp. 1259–1266, 1984.

[2] J. Virieux and S. Operto, "An overview of full-waveform inversion in exploration geophysics," *Geophysics*, vol. 74, no. 6, pp. WCC1–WCC26, 2009.

[3] C. Shin and D.-J. Min, "Waveform inversion using a logarithmic wavefield," *Geophysics*, vol. 71, no. 3, pp. R31–R42, 2006.

[4] E. Bozdağ, J. Trampert, and J. Tromp, "Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements," *Geophysical Journal International*, vol. 185, no. 2, pp. 845–870, 2011.

[5] J. Luo and R.-S. Wu, "Seismic envelope inversion: reduction of local minima and noise resistance," *Geophysical Prospecting*, vol. 63, no. 3, pp. 597–614, 2015.

[6] B. Engquist and B. D. Froese, "Application of the wasserstein metric to seismic signals," *arXiv preprint arXiv:1311.4581*, 2013.

[7] L. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, "Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion," *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, vol. 205, no. 1, pp. 345–377, 2016.

[8] M. Warner and L. Guasch, "Adaptive waveform inversion: Theory," *Geophysics*, vol. 81, no. 6, pp. R429–R445, 2016.

[9] A.-i. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, A. G. Yagola, A. Tikhonov, A. Goncharsky, V. Stepanov, and A. Yagola, *Numerical methods for the approximate solution of ill-posed problems on compact sets*. Springer, 1995.

[10] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[11] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[12] A. Asnaashari, R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, "Regularized seismic full waveform inversion with prior model information," *Geophysics*, vol. 78, no. 2, pp. R25–R36, 2013.

[13] A. Y. Anagaw and M. D. Sacchi, "Full waveform inversion with total variation regularization," in *Recovery-CSPG CSEG CWLS Convention*, 2011, pp. 1–4.

[14] S. Qu, E. Verschuur, and Y. Chen, "Full-waveform inversion and joint migration inversion with an automatic directional total variation constraint," *Geophysics*, vol. 84, no. 2, pp. R175–R183, 2019.

[15] Z. Du, D. Liu, G. Wu, J. Cai, X. Yu, and G. Hu, "A high-order total-variation regularisation method for full-waveform inversion," *Journal of Geophysics and Engineering*, vol. 18, no. 2, pp. 241–252, 2021.

[16] K. Gao and L. Huang, "Acoustic-and elastic-waveform inversion with total generalized p-variation regularization," *Geophysical Journal International*, vol. 218, no. 2, pp. 933–957, 2019.

[17] E. Esser, L. Guasch, T. van Leeuwen, A. Y. Aravkin, and F. J. Herrmann, "Total variation regularization strategies in full-waveform inversion," *SIAM Journal on Imaging Sciences*, vol. 11, no. 1, pp. 376–406, 2018.

[18] E. Esser, L. Guasch, F. J. Herrmann, and M. Warner, "Constrained waveform inversion for automatic salt flooding," *The Leading Edge*, vol. 35, no. 3, pp. 235–239, 2016.

[19] P. Yong, W. Liao, J. Huang, and Z. Li, "Total variation regularization for seismic waveform inversion using an adaptive primal dual hybrid gradient method," *Inverse Problems*, vol. 34, no. 4, p. 045006, 2018.

[20] Y. Wu and Y. Lin, "Inversionnet: An efficient and accurate data-driven full waveform inversion," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 419–433, 2019.

[21] S. Li, B. Liu, Y. Ren, Y. Chen, S. Yang, Y. Wang, and P. Jiang, "Deep-learning inversion of seismic data," *arXiv preprint arXiv:1901.07733*, 2019.

[22] F. Yang and J. Ma, "Deep-learning inversion: A next-generation seismic velocity model building method," *Geophysics*, vol. 84, no. 4, pp. R583–R599, 2019.

[23] V. Kazei, O. Ovcharenko, P. Plotnitskii, D. Peter, X. Zhang, and T. Alkhalifah, "Mapping full seismic waveforms to vertical velocity profiles by deep learning," *Geophysics*, vol. 86, no. 5, pp. R711–R721, 2021.

[24] B. Peters and F. J. Herrmann, "Constraints versus penalties for edge-preserving full-waveform inversion," *The Leading Edge*, vol. 36, no. 1, pp. 94–100, 2017.

[25] P. L. Combettes and N. N. Reyes, "Moreau ' s decomposition in banach spaces," *Mathematical Programming*, vol. 139, no. 1, pp. 103–114, 2013.

[26] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the l 1-ball for learning in high dimensions," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 272–279.

[27] L. Condat, "Fast projection onto the simplex and the l 1 ball," *Mathematical Programming*, vol. 158, no. 1, pp. 575–585, 2016.

[28] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," *Signal, Image and Video Processing*, vol. 9, pp. 1737–1749, 2015.

[29] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, pp. 120–145, 2011.

[30] P. L. Combettes and J.-C. Pesquet, “Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators,” *Set-Valued and variational analysis*, vol. 20, no. 2, pp. 307–330, 2012.

[31] L. Condat, “A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms,” *Journal of optimization theory and applications*, vol. 158, no. 2, pp. 460–479, 2013.

[32] B. C. Vũ, “A splitting algorithm for dual monotone inclusions involving cocoercive operators,” *Advances in Computational Mathematics*, vol. 38, no. 3, pp. 667–681, 2013.

[33] R.-E. Plessix, “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications,” *Geophysical Journal International*, vol. 167, no. 2, pp. 495–503, 2006.