

doi: 10.11720/wtyht.2022.1320

姚含,徐海.基于梯度投影法的全变差正则化全波形反演[J].物探与化探,2022,46(4):977-981.http://doi.org/10.11720/wtyht.2022.1320

Yao H, Xu H. Total variation regularized full waveform inversion based on gradient projection method[J]. Geophysical and Geochemical Exploration, 2022, 46(4): 977-981. http://doi.org/10.11720/wtyht.2022.1320

# 基于梯度投影法的全变差正则化全波形反演

姚含,徐海

(贵州省地质环境监测院, 贵州 贵阳 550001)

**摘要:**为降低地震全波形反演的不稳定性,常用方法是引入未知模型的先验信息,从而将反演问题正则化。但是,传统正则化方法在包含多个先验信息的情况下,仍然面临挑战。本文提出一种扩展的全波形反演公式,其中包含对模型的凸集约束。本文以慢度平方作为反演的模型参数,展示了如何在施加全变差约束的同时,施加边界约束令其保持在一个物理意义上的可行范围内。为验证本文所提算法的适用性,分别开展简单模型及国际标准地质模型数值实验研究,结果表明,全变差正则化的引入可以提高光滑背景模型下高速扰动体的重构效果。

**关键词:** 梯度投影法;全变差正则化;全波形反演

**中图分类号:** P631.4

**文献标识码:** A

**文章编号:** 1000-8918(2022)04-0977-05

## 0 引言

我国油气资源勘探从构造勘探阶段逐步走向岩性勘探阶段,勘探难度逐渐加大。地震勘探具有勘探深度大、精度高、数据量大等特点,是目前油气勘探最有效的方法,应用广泛。为进一步提高地震勘探精度,地震全波形反演方法迅速发展起来,成为国内外学者研究的热点。

全波形反演(full waveform inversion, FWI)是一种高精度、高分辨率的速度反演方法<sup>[1-2]</sup>。FWI利用地震资料中的运动学和动力学信息,通过最小化模拟数据与实际观测数据之间的误差来重构地下介质模型<sup>[3-4]</sup>,能得到地下介质中更多细节的参数变化特征<sup>[5]</sup>。由于实际观测中,子波未知,数据频带有限,且通常存在一定程度的噪声干扰,波场与介质参数之间呈现较强的非线性关系,FWI体现为高度欠定问题,对初始模型的精度要求较高,导致FWI在实际应用中达不到预期的效果。

为了减少FWI的不稳定性,加入先验模型约束的正则化项,能使反演具有更好的稳定性<sup>[6-7]</sup>。全变差(total variation, TV)约束方法是一种非光滑约

束,由Rudin等<sup>[8]</sup>在1992年提出,主要用于图像的去噪处理,能有效处理不连续性解的求解问题,重构图像中的一些间断点,保留边缘信息。近几年来,全变差约束已经应用在电法成像<sup>[9]</sup>、图像去噪<sup>[10-11]</sup>、偏移速度分析<sup>[12-13]</sup>、混合数据偏移<sup>[14]</sup>等领域。地下地层通常存在较多复杂的构造变化带及岩性突变带,导致地层物理参数在这些地方产生不连续变化,应用基于TV约束的FWI可以有效保留这些不连续特征<sup>[15]</sup>。

本文采用梯度投影法<sup>[16]</sup>,构建基于全变差约束的全波形反演目标函数,实现多约束集下的声波方程全波形反演,逐步搜索到全局极值点,避免陷入局部极值,提高全波形速度反演的精度与速度。

## 1 方法原理

FWI是一个最优化问题,通过对观测波场与理论波场的残差求取极小值来构造目标函数进行迭代更新,最终重构地下介质的参数模型。当前常用的目标函数主要是通过数据残差的 $l_2$ 范数来构建,即:

$$G(\mathbf{m}) = \frac{1}{2} \|\mathbf{d}_{\text{cal}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2, \quad (1)$$

收稿日期: 2021-6-22; 修回日期: 2022-4-13

基金项目: 贵州高层次人才科研启动专项资金资助项目(0203001018040)

第一作者: 姚含(1972-),男,2006年毕业于贵州工业大学,主要研究方向包括水文地质、工程地质及环境地质。Email:124121080@qq.com

我们考虑模型参数  $\mathbf{m}$ , 它对应于速度平方的倒数, 是一个实向量  $\mathbf{m} \in \mathbb{R}^N$ , 其中  $N$  是空间离散化中的点数。对于每个源和频率, 波场和观测数据分别用  $u_{sv} \in \mathbb{C}^N$  和  $\mathbf{d}_{\text{obs}} \in \mathbb{C}^{N_r}$  表示, 其中,  $s=1, \dots, N_s$  指源,  $v=1, \dots, N_v$  指频率,  $N_r$  为接收器的数量。  $\mathbf{d}_{\text{cal}} = P u_{sv}$ ,  $P$  是将波场投射到接收器位置的算子。

目标函数对  $\mathbf{m}$  求一阶导数可以得到梯度矩阵, 公式为:

$$\begin{aligned} \nabla G(\mathbf{m}) &= \frac{\partial G(\mathbf{m})}{\partial \mathbf{m}} \\ &= \text{Re} \left[ \left( \frac{\partial \mathbf{d}_{\text{cal}}(\mathbf{m})}{\partial \mathbf{m}} \right)^T (\mathbf{d}_{\text{cal}}(\mathbf{m}) - \mathbf{d}_{\text{obs}})^H \right] \\ &= \text{Re}[\mathbf{J}^T \Delta \mathbf{d}^H], \end{aligned} \quad (2)$$

其中,  $\mathbf{J}$  为 Frechet 导数矩阵, 上标 T 和 H 分别表示矩阵的转置和共轭。对目标函数求二阶导数可以得到近似 Hessian 矩阵, 公式为:

$$\mathbf{H}_a = \text{Re}[\mathbf{J}^T \mathbf{J}^H]. \quad (3)$$

根据高斯牛顿法可以得到模型的更新量为:

$$\Delta \mathbf{m} = -\mathbf{H}_a^{-1} \nabla G(\mathbf{m})_{m_0}. \quad (4)$$

利用梯度下降法<sup>[17]</sup>最小化目标函数  $G$ , 模型更新量可以写成:

$$\Delta \mathbf{m} = \arg \min_{\Delta \mathbf{m}} \Delta \mathbf{m}^T \nabla G(\mathbf{m}) + \frac{1}{2} \Delta \mathbf{m}^T \mathbf{H}_a \Delta \mathbf{m}. \quad (5)$$

## 2 全变差正则化

如果把  $m$  表示成  $N_1 \times N_2$  的速度模型, 网格间距为  $h$ , 用  $(i, j)$  代表各个网格点坐标, 其中  $i=1, 2, \dots, N_1, j=1, 2, \dots, N_2$ , 则在纵向和横向两个方向的梯度分别为:

$$\begin{aligned} (\nabla \mathbf{m})_{i,j}^z &= \begin{cases} \frac{m_{i+1,j} - m_{i,j}}{h} & \text{if } i < N_1 \\ 0 & \text{if } i = N_1 \end{cases} \\ (\nabla \mathbf{m})_{i,j}^x &= \begin{cases} \frac{m_{i,j+1} - m_{i,j}}{h} & \text{if } j < N_2 \\ 0 & \text{if } j = N_2 \end{cases} \end{aligned} \quad (6)$$

全变差定义式如下:

$$\begin{aligned} \|\mathbf{m}\|_{TV} &= \sum_{i \leq N_1, j \leq N_2} \|(\nabla \mathbf{m})_{i,j}\| \\ &= \frac{1}{h} \sum_{ij} \sqrt{(m_{i+1,j} - m_{i,j})^2 + (m_{i,j+1} - m_{i,j})^2}, \end{aligned} \quad (7)$$

式(7)代表离散模型中每一点的离散梯度的  $l_2$  范数之和。假设 Neumann 边界条件使这些差值在边界处为零。我们可以通过定义一个差分算子  $\mathbf{D}$  来更紧凑地表示  $\|\mathbf{m}\|_{TV}$ , 这样  $\mathbf{Dm}$  就是离散梯度的一个

连接,  $(\mathbf{Dm})_n$  表示在  $n$  索引位置对应的离散梯度向量,  $n=1, \dots, N, N=N_1 \times N_2$ , 可以定义:

$$\|\mathbf{m}\|_{TV} = \sum_{n=1}^N \|(\mathbf{Dm})_n\|. \quad (8)$$

将全变差约束项应用于常规全波形反演算法中, 可以得到基于全变差约束的目标函数, 公式为:

$$G(\mathbf{m}) = \frac{1}{2} \|\mathbf{d}_{\text{cal}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2 + \lambda \|\mathbf{m}\|_{TV}, \quad (9)$$

对于式(9), 如果添加约束  $\mathbf{m} \in [B_1, B_2]$  和  $\|\mathbf{m}\|_{TV} \leq \tau$ , 则式(9)具有等价的优化形式:

$$\begin{aligned} \min G(\mathbf{m}) &= \frac{1}{2} \|\mathbf{d}_{\text{cal}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2 \quad \text{约束于} \\ &\mathbf{m}^n + \Delta \mathbf{m} \in [B_1, B_2], \|\mathbf{m}\|_{TV} \leq \tau, \end{aligned} \quad (10)$$

其中,  $\tau$  是与  $\lambda$  有关的一个正常数,  $\tau$  值越小, 全变差约束项在目标函数中的权重作用越大。此时反问题可以表示为:

$$\mathbf{m}^{n+1} = \mathbf{m}^n + \Delta \mathbf{m} \quad (11a)$$

$$\Delta \mathbf{m} = \arg \min_{\Delta \mathbf{m}} \Delta \mathbf{m}^T \nabla G(\mathbf{m}^n) + \frac{1}{2} \Delta \mathbf{m}^T \mathbf{H}_a \Delta \mathbf{m} \quad (11b)$$

根据改进的原始对偶混合梯度(PDHG)方法, 求解式(11)的全变差约束问题相当于求解以下问题的鞍点  $(\Delta \mathbf{m}, \mathbf{p})$ :

$$\begin{aligned} \min_{\Delta \mathbf{m}} \max_{\mathbf{p}} &\left[ \Delta \mathbf{m}^T \nabla G(\mathbf{m}^n) + \frac{1}{2} \Delta \mathbf{m}^T \mathbf{H}_a \Delta \mathbf{m} + \right. \\ &\left. \mathbf{p}^T \mathbf{D}(\mathbf{m}^n + \Delta \mathbf{m}) - \tau \|\mathbf{p}\|_{\infty, 2} \right], \end{aligned} \quad (12)$$

其中,  $\mathbf{p}$  是拉格朗日乘子,  $\mathbf{m}^n + \Delta \mathbf{m} \in [B_1, B_2]$ ,  $\|\cdot\|_{\infty, 2}$  为  $\|\cdot\|_{1,2}$  的对偶模, 相当于求最大值, 而不是  $l_2$  模的和。

修改后的 PDHG 方法需要迭代:

$$\begin{aligned} \mathbf{p}^{k+1} &= \arg \min_{\mathbf{p}} \tau \|\mathbf{p}\|_{\infty, 2} - \mathbf{p}^T \mathbf{D}(\mathbf{m}^n + \Delta \mathbf{m}^k) + \\ &\frac{1}{2\delta} \|\mathbf{p} - \mathbf{p}^k\|^2 \\ \Delta \mathbf{m}^{k+1} &= \arg \min_{\Delta \mathbf{m}} \Delta \mathbf{m}^T \nabla G(\mathbf{m}^n) + \frac{1}{2} \Delta \mathbf{m}^T \mathbf{H}_a \Delta \mathbf{m} + \\ &\Delta \mathbf{m}^T \mathbf{D}^T(2\mathbf{p}^{k+1} - \mathbf{p}^k) + \frac{1}{2\alpha} \|\Delta \mathbf{m} - \Delta \mathbf{m}^k\|^2, \end{aligned} \quad (13)$$

$$\mathbf{m}^n + \Delta \mathbf{m} \in [B_1, B_2]$$

迭代公式可以更明确地写成:

$$\begin{aligned} \mathbf{p}^{k+1} &= \mathbf{p}^k + \delta \mathbf{D}(\mathbf{m}^n + \Delta \mathbf{m}^k) - \\ &\prod_{\|\cdot\|_{1,2} \leq \tau \delta} [\mathbf{p}^k + \delta \mathbf{D}(\mathbf{m}^n + \Delta \mathbf{m}^k)] \\ \Delta \mathbf{m}^{k+1} &= \left( \mathbf{H}_a + \frac{1}{\alpha} \mathbf{I} \right)^{-1} \cdot \\ &\left[ -\nabla G(\mathbf{m}^n) + \frac{\Delta \mathbf{m}^k}{\alpha} - \mathbf{D}^T(2\mathbf{p}^{k+1} - \mathbf{p}^k) \right], \end{aligned} \quad (14)$$

其中,  $\delta$  和  $\alpha$  为迭代步长, 满足  $0 < \alpha\delta < 1/\|D^T D\|$ ,  $\Pi_{\|\cdot\|_{1,2} \leq \tau\delta}(z)$  表示  $z$  在半径为  $\tau\delta$  的球上  $\|\cdot\|_{1,2}$  范数的正交投影。终止条件为:

$$\max\left(\frac{\|p^{k+1} - p^k\|}{\|p^{k+1}\|}, \frac{\|\Delta m^{k+1} - \Delta m^k\|}{\|\Delta m^{k+1}\|}\right) \leq 10^{-5}。 \quad (15)$$

### 3 数值模拟

考虑一个 2D 合成实验, 模拟区域网格点大小为  $200 \times 200$ , 网格宽度  $h$  为 10 m。图 1 所示的合成速度模型有一个由较慢的平滑背景包围的恒定的高速区域。将该平滑背景作为反演的初始模型  $m^0$ 。类似于 van Leeuwen 和 Herrmann<sup>[18]</sup> 中的例 1, 将  $N_s = 34$  个源置于矩形高速异常体左侧,  $N_r = 81$  个接收器置于矩形高速异常体右侧。源  $q_{sv}$  对应于主频为 30 Hz 的 Ricker 小波。

在 5~35 Hz 共 31 个不同频率上开展反演实验。本文考虑无噪声和微噪声两类数据, 在有噪声的情况下, 对每个频率指标  $v$  分别加入随机高斯噪声  $d_v$ , 标准差为  $0.05 \|d_v\|/\sqrt{N_s N_r}$ 。这可能不是一个真实的噪声模型, 但它至少可以表明该方法对数据中的少量噪声是鲁棒的。

正则化参数  $\tau$  有 3 种不同的选择: ①  $\tau_{opt}$ , 即慢度平方的总变化量; ②  $\tau_{large} = 1000\tau_{opt}$ , 此时  $\tau_{large}$  足

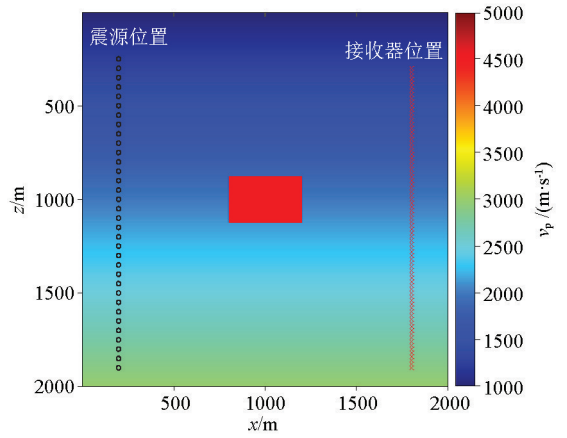


图 1 真实速度模型

Fig.1 True velocity model

够大, 全变差约束没有影响; ③  $\tau_{small} = 0.875\tau_{opt}$ , 比理想的选择稍微小一点。注意, 通过使用从式 (5) 开始的高斯牛顿步骤作为初始模型, 式 (11) 中的凸子问题在  $\tau_{large}$  情况下立即收敛。对于所有实验, 惩罚参数  $\lambda$  固定为 1。从 5 Hz 数据和 6 Hz 数据开始, 使用计算出的  $m$  作为反演 6 Hz 和 7 Hz 数据的初始模型, 按照由低频到高频的顺序, 以此类推。每一个频段都进行 50 次外部迭代, 以确保凸子问题收敛, 达到式 (15) 所示终止条件时, 迭代停止。

图 2 为开展的 6 次数值实验结果。在无噪声和有噪声的情况下, TV 正则化均减少了反演模型中的振荡现象, 并使得高速异常体得到更好的恢复。

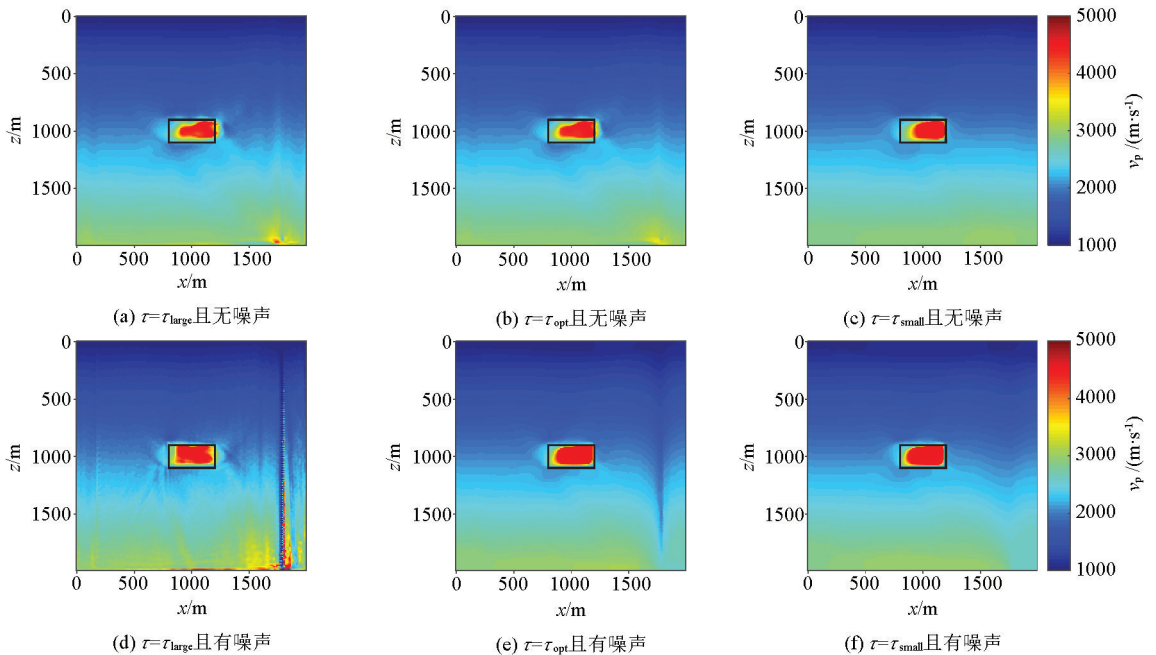


图 2 不同正则化参数  $\tau$  在有噪声和无噪声情况下反演的速度模型

Fig.2 Inversed velocity model from noise free data and noisy data under different regularization parameter  $\tau$

从噪声数据中反演得到的不连续点的位置要比从无噪声数据中反演的更准确。这是因为噪声在接收器上造成了更大的不连续,从而加强了其他地方TV正则化的效果。 $\tau_{small}$ 实验证明了这一点,增加TV正则化的权重会导致较大不连续点的分辨率更高。

选取国际标准复杂地质模型开展全波形反演实验。真实速度模型如图3a所示,取自墨西哥湾典型复杂岩体的合成数据——SEAM模型。模拟区域网

格点大小为  $267 \times 384$ , 网格宽度  $h$  为 20 m。在靠近地面处放置总共 72 个源和 191 个接收器。初始模型采用图 3b 所示平滑模型。图 3c 及图 3d 为无噪声数据下的反演结果,图 3e 及图 3f 为噪声数据下的反演结果。图 3d 及图 3f 为选取  $\tau_{small}$  正则化参数下的反演结果,图 3c 及图 3e 为选取  $\tau_{large}$  正则化参数下的反演结果。由图可知,全变差正则化有助于消除伪影,主要是减少盐下的噪声。

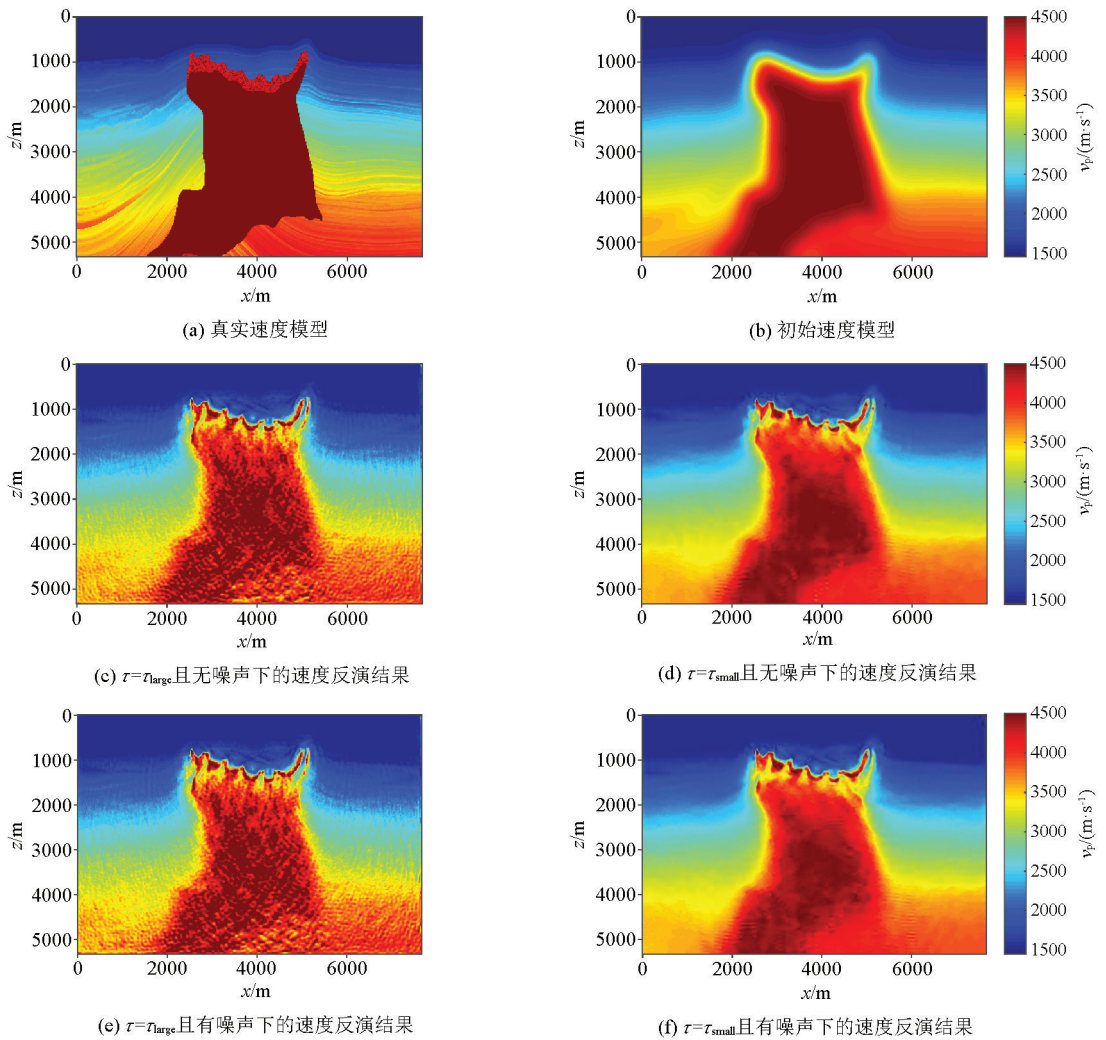


图3 Seam模型的真实速度模型、初始速度模型,以及不同正则化参数 $\tau$ 在有噪声和无噪声情况下反演的速度模型  
 Fig.3 True and initial velocity model of Seam model, and inverted velocity model from noise free data and noisy data under different regularization parameter  $\tau$

#### 4 结论及讨论

本文提出了一种在计算上可行的梯度投影算法,该算法可将附加凸约束条件下的全波形反演二次惩罚公式最小化。本文展示了在模型上添加全变

差约束和边界约束时,全波形反演中凸子问题的求解过程。数值实验表明,在数据有限或含有噪声的情况下,TV正则化可以通过消除伪影和精确识别不连续点的位置,提高模型反演效果。

本文围绕的梯度投影法,其本质依赖于目标函数的一阶偏微分信息,这使得收敛速度受到一定制

约。因此,理论上,我们可以采用二阶方法进行非线性优化,以探索收敛速度是否有提升。

## 参考文献 (References):

- [1] Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics [J]. *Geophysics*, 2009, 74 (6): WCC1 – WCC26.
- [2] Andrew P V, Malcolm S. Elastic versus acoustic inversion for marine surveys [J]. *Geophysical Journal International*, 2018, 215 (1): 1003 – 1021.
- [3] Lailly P. The seismic inverse problem as a sequence of before-stack migrations [C] // *Conference on Inverse Scattering: Theory and Applications, Expanded Abstracts*, 1983: 206 – 220.
- [4] Tarantola A. Inversion of seismic reflection data in the acoustic approximation [J]. *Geophysics*, 1984, 49 (8): 1259 – 1266.
- [5] Mora P, Wu Z. Elastic versus acoustic inversion for marine surveys [J]. *Geophysical Journal International*, 2018, 214 (1): 596 – 622.
- [6] Lin Y Z, Huang L J. Acoustic and elastic-waveform inversion using a modified total-variation regularization scheme [J]. *Geophysical Journal International*, 2015, 200: 489 – 502.
- [7] Du Z, Liu D, Wu G, et al. A high-order total-variation regularization method for full-waveform inversion [J]. *Journal of Geophysics and Engineering*, 2021, 18 (2): 241 – 252.
- [8] Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms [J]. *Physica D: Nonlinear Phenomena*, 1992, 60: 259 – 268.
- [9] Chung E T, Chan T F, Tai X C. Electrical impedance tomography using level set representation and total variational regularization [J]. *Journal of Computational Physics*, 2005, 205: 357 – 372.
- [10] 屈勇, 曹俊兴, 朱海东, 等. 一种改进的全变分地震图像去噪技术 [J]. *石油学报*, 2011, 32 (5): 815 – 819.
- Qu Y, Cao J X, Zhu H D, et al. An improved total variation technique for seismic image denoising [J]. *Acta Petrolei Sinica*, 2011, 32 (5): 815 – 819.
- [11] 公成敏. 全变分原理在地震数据去噪中的应用 [J]. *计算机与数字工程*, 2014, 42 (7): 1271 – 1274.
- Gong C M. Application of variational principle in seismic data denoising [J]. *Computer & Digital Engineering*, 2014, 42 (7): 1271 – 1274.
- [12] Anagaw A Y. Total variation and adjoint stat methods for seismic wavefield imaging [M]. Alberta: Physics Department of University of Alberta, 2009.
- [13] Anagaw A Y, Sacchi M D. Edge-preserving seismic imaging using the total variation method [J]. *Journal of Geophysics and Engineering*, 2012, 9 (2): 138 – 146.
- [14] 卢昕婷, 韩立国, 张盼, 等. 基于全变分原理的多震源混合数据直接偏移方法 [J]. *地球物理学报*, 2015, 58 (9): 3335 – 3345.
- Lu X T, Han L G, Zhang P, et al. Direct migration method of multi-source blended data based on total variation [J]. *Chinese Journal of Geophysics*, 2015, 58 (9): 3335 – 3345.
- [15] Esser E, Guasch L, Leeuwen T V, et al. Total variation regularization strategies in full-waveform inversion [J]. *SIAM Journal on Imaging sciences*, 2018, 11 (1): 376 – 406.
- [16] Birgin E G, Martínez J M, Raydan M. Nonmonotone spectral projected gradient methods on convex sets [J]. *SIAM Journal on Control and Optimization*, 2000, 10 (4): 1196 – 1211.
- [17] Bertsekas D P. *Nonlinear programming; 2<sup>nd</sup> edition* [M]. Belmont: Athena Scientific, 1999.
- [18] van Leeuwen T, Herrmann F J. Mitigating local minima in full-waveform inversion by expanding the search space [J]. *Geophysical Journal International*, 2013, 195: 661 – 667.

## Total variation regularized full waveform inversion based on gradient projection method

YAO Han, XU Hai

(Guizhou Geological Environment Monitoring Institute, Guiyang 550001, China)

**Abstract:** To reduce the ill-posedness of seismic full waveform inversion, a common method is to introduce prior information to regularize the inversion problem. Traditional regularization methods still face challenges even when they contain multiple prior information. This study proposed an extended full waveform inversion formula, which includes the convex set constraints on models. Specifically, this study showed how to constrain the total variation of the slowness square while forcing the constraint to keep it within a physical reality range. To verify the applicability of the algorithm proposed in this study, numerical experiments on simple models and international standard geological models were carried out. The results show that the introduction of total variation regularization can improve the reconstruction of high-speed disturbances under smooth background models.

**Key words:** gradient projection method; total variation regularization; full waveform inversion

(本文编辑:叶佩)