APPLIED GEOPHYSICS, Vol.19, No. 2 (June 2022), P. 185–196, 11 Figures.
DOI:10.1007/s11770-022-0934-6

Efficient seismic data reconstruction based on Geman function minimization

Li Yan-Yan[1], Fu Li-Hua[1]*, Cheng Wen-Ting[1], Niu Xiao[1], and Zhang Wan-Juan[1]

Abstract: Seismic data typically contain random missing traces because of obstacles and economic restrictions, influencing subsequent processing and interpretation. Seismic data recovery can be expressed as a low-rank matrix approximation problem by assuming a low-rank structure for the complete seismic data in the frequency–space (f–*x*) domain. The nuclear norm minimization (NNM) (sum of singular values) approach treats singular values equally, yielding a solution deviating from the optimal. Further, the log-sum majorization–minimization (LSMM) approach uses the nonconvex log-sum function as a rank substitution for seismic data interpolation, which is highly accurate but time-consuming. Therefore, this study proposes an efficient nonconvex reconstruction model based on the nonconvex Geman function (the nonconvex Geman low-rank (NCGL) model), involving a tighter approximation of the original rank function. Without introducing additional parameters, the nonconvex problem is solved using the Karush–Kuhn–Tucker condition theory. Experiments using synthetic and field data demonstrate that the proposed NCGL approach achieves a higher signal-to-noise ratio than the singular value thresholding method based on NNM and the projection onto convex sets method based on the data-driven threshold model. The proposed approach achieves higher reconstruction efficiency than the singular value thresholding and LSMM methods.

Keywords: Seismic data reconstruction, low rank, Geman function, nonconvex, Karush–Kuhn–Tucker condition

Introduction

Due to limitations, such as environmental interference and instrument conditions, collected seismic data are frequently irregular or incomplete and can severely affect the subsequent data processing and interpretation (Fu et al., 2018; Liu et al., 2017a). Therefore, missing data reconstruction has been incorporated into seismic exploration technology, and improving its accuracy is crucial for interpretation.

Several seismic data reconstruction methods have been proposed, such as sparse transform methods employing interpolation in a transform domain such as Radon (Kabir and Verschuur, 1995), Fourier (Zhang et al., 2016), Curvelet (Wang et al., 2015a), Dreamlet (Wang et al., 2015b), and Seislet (Liu et al., 2017b). Prediction filter methods recover seismic data via predictive filters, achieving superior antialiasing performance (Biondi et al., 1998). Wave equation approaches exploit the underground medium's velocity information for reconstructing seismic data, which is computation intensive (Witten and Shragge, 2015). Machine learning (ML) techniques automatically learn the mapping

Manuscript received by the Editor August 1, 2021; revised manuscript received March 1, 2022.
1. School of Mathematics and Physics, China University of Geosciences, Wuhan 430074, China.
*Corresponding author: Fu Li-Hua (E-mail: lihuafu@cug.edu.cn).

between incomplete and complete data (Fang et al., 2021; Liu et al., 2021). Consequently, they can interpolate the missing trace and are more efficient than traditional methods. However, ML approaches require a large amount of data for optimal performance.

Rank-reduction (RR) approaches are also popular for seismic recovery (Chen et al., 2017; Gao et al., 2017; Zhang et al., 2017). In these approaches, complete and noise-free seismic data have a low-rank structure, whereas missing traces increase the rank or slow the decay of singular values. Therefore, missing traces can be interpolated using RR techniques. Trickett et al. (2010) proposed Cadzow filtering via a truncated singular value decomposition method, but it is computation intensive. Therefore, Oropeza and Sacchi (2011) proposed a multichannel singular-spectrum analysis method based on the random singular value decomposition (SVD) algorithm, decreasing the computation time. Yang et al. (2013) introduced the nuclear norm as the rank function's convex approximation and Ma (2013) extended it to a three-dimensional (3D) situation.

Because the nuclear norm is the rank function's convex approximation, convex optimization methods can solve the nuclear norm minimization (NNM) method. These include singular value thresholding (SVT) (Cai et al., 2010), accelerated proximal gradient (Toh and Yun, 2010), augmented Lagrangian multiplier (Yang and Yuan, 2013), and fixed-point continuation (Ma et al., 2011). However, NNM involves minimizing the sum of singular values, different from rank minimization, providing a suboptimal solution to the optimization problem.

The nonconvex function has been widely employed as a rank function surrogate because of its rank approximation superiority. Zhang et al. (2020) proposed a truncated nuclear norm regularization method for adequate seismic reconstruction; however, it requires estimating the matrix rank and two-layer SVD steps, making it time-consuming. They also achieved seismic interpolation via a nonconvex function log-sum majorization–minimization (LSMM) framework (Zhang et al., 2019), with additional regularization parameters introduced in the corresponding optimization method. However, the algorithm converges slowly by minimizing the surrogate function, and choosing a reasonable parameter is challenging.

This study uses the nonconvex Geman function to approximate the original rank function and applies the Karush–Kuhn–Tucker (KKT) condition theory to iteratively solve the problem. This approach requires no additional parameters, eliminating the time required for tuning parameters. Experimental results showed that the proposed method (the nonconvex Geman low-rank (NCGL) method) outperforms classical methods, such as the SVT and projection onto convex sets (POCS) based on sparse transform (Gao et al., 2013) in terms of reconstruction accuracy, ensuring a better reconstruction effect. Further, NCGL significantly improves the seismic data reconstruction efficiency compared to SVT and LSMM approaches.

Seismic data reconstruction based on the low-rank theory

A two-dimensional (2D) seismic record $\mathbf{Z}_t = \mathbf{Z}(t,x) \in \mathbb{R}^{M\times N} (t = 1,...,M, x = 1,...,N)$ includes N traces and M time-sampling points at each trace. Let $\mathbf{Z}_f(\omega,x) \in \mathbb{C}^{M\times N}$ be the data in the f–x domain obtained by applying the Fourier transform to each trace of $\mathbf{Z}_t$. The fixing frequency is $\omega = \omega_j$ (ω_j is a certain frequency value), and each frequency slice vector of $\mathbf{Z}_f$ is expressed as (Oropeza and Sacchi, 2011)

$$\mathbf{z}_{\omega_j} = (\mathbf{Z}_f(\omega_j,1), \mathbf{Z}_f(\omega_j,2),..., \mathbf{Z}_f(\omega_j,N)) \in \mathbb{C}^{1\times N}, \quad (1)$$

and can be rearranged into a Hankel matrix as

$$\mathbf{M}_{\omega_j} = P_h(\mathbf{z}_{\omega_j}) = \begin{pmatrix} \mathbf{Z}_f(\omega_j,1) & \mathbf{Z}_f(\omega_j,2) & ... & \mathbf{Z}_f(\omega_j,n) \\ \mathbf{Z}_f(\omega_j,2) & \mathbf{Z}_f(\omega_j,3) & ... & \mathbf{Z}_f(\omega_j,n+1) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{Z}_f(\omega_j,m) & \mathbf{Z}_f(\omega_j,m+1) & ... & \mathbf{Z}_f(\omega_j,N) \end{pmatrix}, \quad (2)$$

where $P_h : \mathbb{C}^{1\times N} \to \mathbb{C}^{m\times n}$ denotes the Hankel transform operator, $m = floor(N/2+1)$. $floor(\cdot)$ is the integer down function, and $n = N - m + 1$. Thus, the Hankel matrix approximates a square matrix and $m \geq n$.

The Hankel matrix at a given frequency slice of complete seismic data is of low rank. If seismic records are missing or corrupted by noise, the Hankel matrix's rank increases or produces slow-decaying singular values (Oropeza and Sacchi, 2011). Therefore, seismic data recovery can be solved using rank minimization. Figure 1 shows the basic procedure for seismic recovery based on the Hankel pretransform.

First, we acquire a 2D dataset $\mathbf{Z}_f$ in the f–x domain using a 1D Fourier transform of the original seismic data $\mathbf{Z}_t$. Each frequency slice $\mathbf{z}_{\omega_j}$ ($\omega_j \in [\omega_l, \omega_h]$) of data $\mathbf{Z}_f$ is rearranged into a Hankel matrix $\mathbf{M}_{\omega_j}$, followed by calculating the low-rank approximation matrix $\mathbf{X}^*_{\omega_j}$ of $\mathbf{M}_{\omega_j}$ using rank minimization algorithms. We then obtain the recovery frequency slice vector $\mathbf{z}^*_{\omega_j} \in \mathbb{C}^{1\times N}$, expressed as

$$\mathbf{z}^*_{\omega_j} = P_h^{-1}(\mathbf{X}^*_{\omega_j}). \quad (3)$$

Accordingly, the inverse Hankel transform operator $P_h^{-1}:\mathbb{C}^{m\times n}\rightarrow\mathbb{C}^{l\times N}$ transforms the recovery 2D matrix $\mathbf{X}_{\omega_j}^*$ into a 1D frequency slice vector $\mathbf{z}_{\omega_j}^*$. Finally, by reconstructing each slice $\mathbf{z}_{\omega_j}^*$, seismic records $\mathbf{Z}_t^*$ in the t–x domain are recovered using the 1D inverse Fourier transform of the data $\mathbf{Z}_f^*\in\mathbb{C}^{M\times N}$ (with $\mathbf{Z}_f^*$ containing $\mathbf{z}_{\omega_j}^*$), yielding the following transformation operator

$$\mathbf{Z}_t^* = ifft(\mathbf{Z}_f^*). \tag{4}$$

According to the above theory, we model each frequency slice of observed data as

$$\min_{\mathbf{X}_\omega} rank(\mathbf{X}_\omega) \qquad s.t.\ \mathbf{W}\odot\mathbf{X}_\omega = \mathbf{M}_\omega, \tag{5}$$

where $\mathbf{X}_\omega$ is the matrix to be restored. $rank(\mathbf{X}_\omega)$ is the rank function of $\mathbf{X}_\omega$ (the number of nonzero singular values). $\mathbf{M}_\omega$ is the Hankel matrix constructed from the frequency slice, and $\odot$ represents the matrix's Hadamard product. The variable $\mathbf{W}=P(\mathbf{M}_\omega)\in\mathbb{R}^{m\times n}$ can be expressed as

$$\mathbf{W}_{ij}=\begin{cases}1 & if\ (i,j)\in\Omega\\ 0 & if\ (i,j)\notin\Omega\end{cases}, \tag{6}$$

where $P:\mathbb{R}^{m\times n}\rightarrow\mathbb{R}^{m\times n}$ represents the masking operator and Ω denotes an index subset corresponding to the observed entries of $\mathbf{M}_\omega$.

The rank function is nonconvex and discontinuous, making solving equation (5) difficult. Yang et al. (2013) used the nuclear norm substitute rank, which resulted in the relation of the rank minimization model into the NNM model, given as

$$\min_{\mathbf{X}_\omega}\|\mathbf{X}_\omega\|_* \qquad s.t.\ \mathbf{W}\odot\mathbf{X}_\omega=\mathbf{M}_\omega, \tag{7}$$

where $\|\mathbf{X}_\omega\|_*=\sum_{i=1}^{n}\sigma_i(\mathbf{X}_\omega)(i=1,...,n)$ and $\sigma_i(\mathbf{X}_\omega)$ is the ith largest singular value of $\mathbf{X}_\omega$.

In the NNM approach, the sums of the singular values are simultaneously minimized, ignoring the differences between the singular values that frequently leads to a biased solution.

Considering that the data matrix's rank corresponds to the number of nonzero singular values, the rank minimization problem equates to the singular value vectors' l_0 norm minimization. Conventional compressive sensing uses the l_1 norm as a surrogate l_0 norm to efficiently solve the convex problem. However, the l_1 minimization solution is typically suboptimal relative to that involving the original l_0 minimization because the former is a loose approximation of the latter. Many nonconvex surrogates have been proposed for the l_0 norm, and nonconvex sparse optimization numerically outperforms convex models in signal recovery, error correction, and image processing (Candès et al., 2008). This study employs the nonconvex Geman function as a surrogate rank function instead of the nuclear norm (l_1 norm of matrix singular value vectors) for seismic reconstruction.

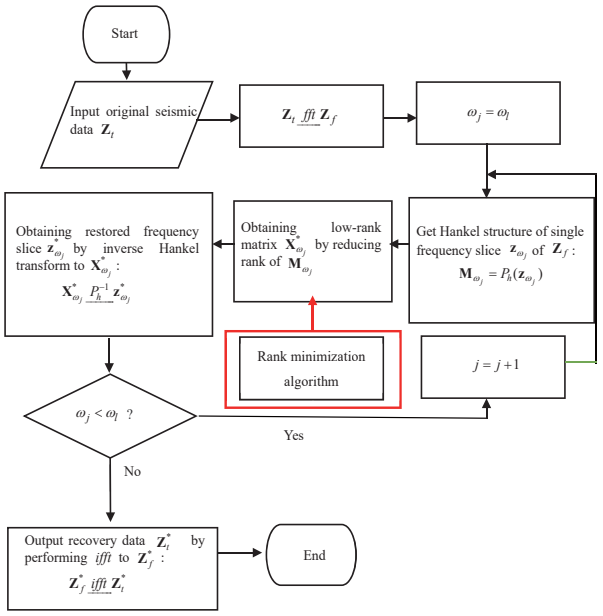


Figure 1. Flowchart for seismic data reconstruction based on the Hankel transform approach.

Nonconvex Geman function-based approach for seismic reconstruction

Nonconvex Geman function

The continuous, differentiable, and nonconvex Geman function (Geman and Yang, 1995) can be expressed as

$$f(x)=\frac{x}{x+\gamma},\ (x\geq 0), \tag{8}$$

where γ is a constant. The Geman function is often used as an l_0 norm surrogate function. According to many studies, rank is approximated better using this function; thus, it is widely applied in image processing and signal recovery (Geman and Yang, 1995; Lu et al., 2016).

Figure 2 shows that the Geman function approximates rank closer than the nuclear norm, comparing its results at different γ (γ = 1 and 10) with those from the nuclear norm. The horizontal axis displays the singular values, and the vertical axis represents the corresponding function values. The black starred line represents the true rank, fixed at 1. The lines with red asterisks and blue circles exhibit the Geman function's contribution when γ = 1 and 10 for varying singular values, respectively, and the line with green triangles depicts varying nuclear norm results with singular values. The nonconvex Geman function results are closer to the true rank regarding the variation of the singular values, whereas the nuclear norm results display significant deviation from the true rank.

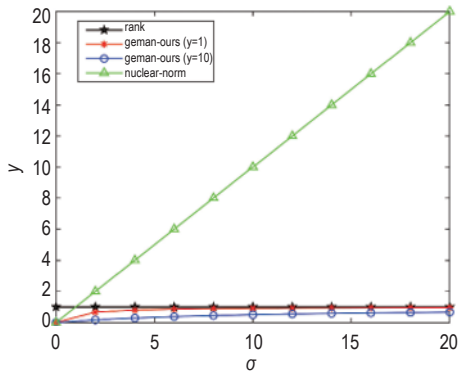


Figure 2. Plot of γ versus singular values showing rank approximation involving the Geman function and nuclear norm, for a true rank of 1.

If the Geman function is used as a surrogate rank function, the optimization problem can be formulated as

$$\min_{\mathbf{X}_\omega} \sum_{i=1}^{n} \frac{\sigma_i(\mathbf{X}_\omega)}{\sigma_i(\mathbf{X}_\omega)+\gamma} \qquad s.t.\ \mathbf{W}\odot\mathbf{X}_\omega = \mathbf{M}_\omega. \tag{9}$$

However, because equation (9) represents a nondeterministic polynomial time (NP)-hard problem, this study uses the KKT condition theory to derive an iterative approximate solution to the original optimization problem.

NCGL algorithm

The left-hand side of equation (9) can be expressed as

$$\sum_{i=1}^{n} \frac{\sigma_i(\mathbf{X}_\omega)}{\sigma_i(\mathbf{X}_\omega)+\gamma} = \sum_{i=1}^{n} \frac{(\lambda_i(\mathbf{X}_\omega^T\mathbf{X}_\omega))^{0.5}}{(\lambda_i(\mathbf{X}_\omega^T\mathbf{X}_\omega))^{0.5}+\gamma}, \tag{10}$$

where $\lambda_i(\cdot)$ is the matrix's ith eigenvalue. For notation ease, we replace $\lambda_i(\mathbf{X}_\omega^T\mathbf{X}_\omega)$ and $\mathbf{X}_\omega^T\mathbf{X}_\omega \in \mathbb{C}^{n\times n}$ with λ_i and $h(\mathbf{X}_\omega)$, respectively. Then, equation (9) can be rewritten as

$$\min_{\mathbf{X}_\omega} \sum_{i=1}^{n} \frac{\lambda_i^{0.5}}{\lambda_i^{0.5}+\gamma} \qquad s.t.\ \mathbf{W}\odot\mathbf{X}_\omega = \mathbf{M}_\omega. \tag{11}$$

However, the optimization problem based on equation (11) remains NP-hard because the equation is nonconvex. We use the KKT condition of equation (11) to obtain a solution for the nonconvex regularizer by formulating a corresponding Lagrangian function.

The Lagrangian function of equation (11) is given as

$$L(\mathbf{X}_\omega, \mathbf{Y}) = \sum_{i=1}^{n} \frac{\lambda_i^{0.5}}{\lambda_i^{0.5}+\gamma} - Tr(\mathbf{Y}^T(\mathbf{W}\odot\mathbf{X}_\omega - \mathbf{M}_\omega)), \tag{12}$$

where $\mathbf{Y}$ denotes the Lagrangian multiplier and $Tr(\cdot)$ is the matrix's trace function.

Let $T(\mathbf{X}_\omega) = Tr(\mathbf{Y}^T(\mathbf{W}\odot\mathbf{X}_\omega - \mathbf{M}_\omega))$.The KKT condition of equation (11) produces the following equations. The optimal local solution to equation (11) satisfies equations. (13) and (14).

$$\frac{\partial \sum_{i=1}^{n} \frac{\lambda_i^{0.5}}{\lambda_i^{0.5}+\gamma}}{\partial \mathbf{X}_\omega} - \frac{\partial T(\mathbf{X}_\omega)}{\partial \mathbf{X}_\omega} = \mathbf{0}, \tag{13}$$

$$\mathbf{W}\odot\mathbf{X}_\omega = \mathbf{M}_\omega. \tag{14}$$

Then, we simplify equation (13). According to the matrix differentiation property, equation (13) can be rewritten as

$$\frac{Tr((\frac{\partial \sum_{i=1}^{n} \frac{\lambda_i^{0.5}}{\lambda_i^{0.5}+\gamma}}{\partial \mathbf{X}_\omega})^T \partial h(\mathbf{X}_\omega))}{\partial \mathbf{X}_\omega} - \frac{\partial T(\mathbf{X}_\omega)}{\partial \mathbf{X}_\omega} = \mathbf{0}. \tag{15}$$

Due to the relationship between the symmetric matrix's eigenvalues and the corresponding eigenvectors (Magnus, 1985), we can obtain equation (16),

$$\frac{\partial \sum_{i=1}^{n} \frac{\lambda_i^{0.5}}{\lambda_i^{0.5}+\gamma}}{\partial \mathbf{X}_\omega} = \sum_{i=1}^{n} \frac{\gamma}{2\lambda_i^{0.5}(\lambda_i^{0.5}+\gamma)} \frac{\partial \lambda_i}{\partial h(\mathbf{X}_\omega)} = \sum_{i=1}^{n} \frac{\gamma}{2\lambda_i^{0.5}(\lambda_i^{0.5}+\gamma)} \mathbf{u}_i\mathbf{u}_i^T, \tag{16}$$

where $\mathbf{u}_i \in \mathbb{R}^{n\times 1}$ is the corresponding eigenvector of the ith eigenvalue of $h(\mathbf{X}_\omega)$.

Equation (15) can be updated as

$$\frac{Tr((\sum_{i=1}^{n} \frac{\gamma}{2\lambda_i^{0.5}(\lambda_i^{0.5}+\gamma)} \mathbf{u}_i\mathbf{u}_i^T)^T \partial h(\mathbf{X}_\omega))}{\partial \mathbf{X}_\omega} - \frac{\partial T(\mathbf{X}_\omega)}{\partial \mathbf{X}_\omega} = \mathbf{0}. \tag{17}$$

Let

$$\mathbf{D}=\sum_{i=1}^{n}\frac{\gamma}{2\lambda_i^{0.5}(\lambda_i^{0.5}+\gamma)}\mathbf{u}_i\mathbf{u}_i^T=\mathbf{U}[\mathbf{v}]\mathbf{U}, \tag{18}$$

where $\mathbf{u}_j$ is the ith column of $\mathbf{U}\in\mathbb{R}^{n\times n}$, $[\mathbf{v}]=diag(\mathbf{v})$ is a diagonal matrix, and

$\mathbf{v}=[v_1,...,v_n]\in\mathbb{R}^{1\times n}(v_i=\gamma/2\lambda_i^{0.5}(\lambda_i^{0.5}+\gamma),\ i=1,...,n).$

Based on the definition of $\mathbf{D}$, equation (17) can be reduced to

$$\frac{Tr(\mathbf{D}^T\partial h(\mathbf{X}_\omega))}{\partial\mathbf{X}_\omega}-\frac{\partial T(\mathbf{X}_\omega)}{\partial\mathbf{X}_\omega}=\mathbf{0}. \tag{19}$$

Then, equation (19) can be rewritten as

$$2\mathbf{X}_\omega\mathbf{D}-\mathbf{W}\odot\mathbf{Y}=\mathbf{0}. \tag{20}$$

Finally, equation (13) is simplified to equation (21) as

$$\mathbf{X}_\omega=\frac{1}{2}(\mathbf{W}\odot\mathbf{Y})\mathbf{D}^{-1}. \tag{21}$$

In summary, optimal local solutions $\mathbf{X}_\omega$ and $\mathbf{Y}$ to equation (11) satisfy equations (21) and (14). Because the definition of $\mathbf{D}$ relies on $\mathbf{X}_\omega$, equation (21) represents a fixed equation that can be solved iteratively. Therefore, in the kth iteration, if $\mathbf{D}^k$ is fixed, $\mathbf{X}_\omega^{k+1}$ and $\mathbf{Y}^{k+1}$ can be iteratively calculated using equations (22) and (23)

$$\mathbf{X}_\omega=\frac{1}{2}(\mathbf{W}\odot\mathbf{Y})(\mathbf{D}^k)^{-1}, \tag{22}$$

$$\mathbf{W}\odot\mathbf{X}_\omega=\mathbf{M}_\omega. \tag{23}$$

It is necessary to ensure that $\mathbf{D}^k$ is a positive definite matrix. Therefore, we set $h(\mathbf{X}_\omega)=\mathbf{X}_\omega^T\mathbf{X}_\omega+\varepsilon\mathbf{I}$, where ε= $0.0001||diag(\mathbf{M}_\omega{}^T\mathbf{M}_\omega)||_\infty$.

Further, based on equations (22) and (23), we derive

$$\mathbf{M}_\omega=\frac{1}{2}\mathbf{W}\odot((\mathbf{W}\odot\mathbf{Y})(\mathbf{D}^k)^{-1}). \tag{24}$$

For equation (24), we cannot directly calculate $\mathbf{Y}$; therefore, we update the rows of $\mathbf{Y}$ individually. We write the row vector form of equation (24) as (j is the number of rows)

$$\mathbf{m}_j=\frac{1}{2}(\mathbf{W}\odot\mathbf{Y})_j(\mathbf{D}^k)^{-1}[\mathbf{w}_j]=\frac{1}{2}\mathbf{y}_j[\mathbf{w}_j](\mathbf{D}^k)^{-1}[\mathbf{w}_j], \tag{25}$$

where $\mathbf{m}_j$ is the jth row of $\mathbf{m}_\omega$, $\mathbf{w}_j$ denotes the jth row of $\mathbf{w}$, $\mathbf{y}_j$ represents the jth row of $\mathbf{Y}$, and $[\mathbf{w}_j]=\mathrm{diag}(\mathbf{w}_j)(j=1,\cdots,m)$ is a diagonal matrix.

Let $\mathbf{A}^k=[\mathbf{w}_j](\mathbf{D}^k)^{-1}[\mathbf{w}_j]$, and we obtain

$$\mathbf{m}_j=\frac{1}{2}\mathbf{y}_j\mathbf{A}^k. \tag{26}$$

$\mathbf{w}_j$ has zero values; therefore, $\mathbf{A}^k$ represents a singular matrix, and $\mathbf{m}_j$ is not directly solvable using equation (26). We obtain the nonzero subvector $\hat{\mathbf{m}}_j$ by obtaining nonzero elements of $\mathbf{m}_j$ and nonzero submatrix $\hat{\mathbf{A}}^k$ by obtaining nonzero elements of $\mathbf{A}^k$. Then, we obtain

$$\hat{\mathbf{y}}_j=2\hat{\mathbf{m}}_j(\hat{\mathbf{A}}^k)^{-1}, \tag{27}$$

where $\hat{\mathbf{y}}_j$ is a nonzero subvector corresponding to the nonzero elements of $\mathbf{y}_j$.

The Lagrangian multiplier $\mathbf{Y}^{k+1}$ is obtained using equation (27). Further, we obtain

$$\mathbf{X}_\omega^{k+1}=\frac{1}{2}(\mathbf{W}\odot\mathbf{Y}^{k+1})(\mathbf{D}^k)^{-1} \tag{28}$$

Algorithm 1 outlines the seismic data reconstruction procedure based on the NCGL method.

Algorithm 1: Nonconvex Geman low-rank (NCGL) method for seismic data reconstruction

Input: Original seismic records $\mathbf{Z}_t$, γ, $err=1e-4$, ε, $K=300$, ω_l, ω_h

1. $\mathbf{Z}_f=fft(\mathbf{Z}_t)$,
2. for $\omega=\omega_l:\omega_h$
3. $\mathbf{M}_\omega=P_h(\mathbf{z}_\omega)$, $\mathbf{X}_\omega^1=\mathbf{M}_\omega$, $\mathbf{W}=P(\mathbf{M}_\omega)$,
4. for $k=1:K$
5. $[\mathbf{U}^k,[\boldsymbol{\lambda}^k],\mathbf{U}^k]=svd(h(\mathbf{X}_\omega^k))=svd((\mathbf{X}_\omega^k)^T\mathbf{X}_\omega^k+\varepsilon\mathbf{I})$, $\boldsymbol{\lambda}^k=[\lambda_1^k,...,\lambda_n^k]$,
6. $v_i^k=\dfrac{\gamma}{2(\lambda_i^k)^{0.5}((\lambda_i^k)^{0.5}+\gamma)}(i=1,...,n)$,
7. Calculate $\mathbf{D}^k$ using equation (18): $\mathbf{D}^k=\mathbf{U}^k[\mathbf{v}](\mathbf{U}^k)^T$ $(\mathbf{v}^k=[v_1^k,...,v_n^k])$,
8. Update the rows of $\mathbf{Y}^{k+1}$ individually using equation (27),
9. Update $\mathbf{X}_\omega^{k+1}$ using equation (28),
10. if $\dfrac{(o_{k+1}-o_k)}{o_k}<err$ break (where o_k denotes the objective function value at the kth iteration)
11. end for
12. $\mathbf{z}_\omega^*=P_h^{-1}(\mathbf{X}_\omega)$,
13. end for

Output: Reconstructed seismic records $\mathbf{Z}_t^*=ifft(\mathbf{Z}_f^*)$ (with $\mathbf{Z}_f^*$ containing $\mathbf{z}_\omega^*$).

From Algorithm 1, in each iteration, the matrix to be recovered $\mathbf{X}_\omega$ is updated using the current weighted matrix $\mathbf{D}$, which is updated using the updated $\mathbf{X}_\omega$. Finally, we used the operators P_h^{-1} and *ifft* to transform the restored low-rank matrix $\mathbf{X}_\omega$ into the reconstructed seismic data $\mathbf{Z}_t^*$. Furthermore, the only parameter requiring tuning in Algorithm 1 is γ, performed empirically to achieve the best results. Appendix A presents the convergence analysis of Algorithm 1.

Experiments

We consider a synthetic seismic dataset comprising three curving events and a field prestack dataset to verify the proposed reconstruction approach's performance. To evaluate the NCGL method's results, we define the signal-to-noise (SNR) as

$$SNR = 10\,log_{10}\left(\frac{\left\|\tilde{\mathbf{Z}}\right\|_F^2}{\left\|\tilde{\mathbf{Z}} - \mathbf{Z}_t^*\right\|_F^2}\right), \tag{29}$$

where $\tilde{\mathbf{Z}}$ represents the complete and noiseless data, and $\mathbf{Z}_t^*$ is the recovered data. A high SNR value indicates a better reconstruction performance. Furthermore, the POCS, SVT, and LSMM methods were also used for comparison.

We selected parameters that could yield the best results in our experiments. Therefore, for the NCGL method, we selected the Geman function with γ belonging to the set {10, 20, 30, 40, 50}, with the convergence error set to $1e-4$. Conversely, for the LSMM method, we varied the ε, *tol*, and λ_0 parameters within previous studies' recommendations (Zhang et al., 2019). For the SVT method, we set the regularization parameter $\tau = 200$. For the POCS method, we set $P_i = 98$, $P_i = 1$, and $N = 100$. Further, we empirically selected the best parameter values for all methods.

Synthetic data example

Figure 3(a) shows the synthetic seismic dataset comprising 200 traces, with 600 temporal sampling points for each trace, a temporal sampling interval of 4 ms, and a spatial sampling interval of 10 m. The corresponding corrupted data were characterized by 60% of the traces missing, with randomly selected indices between 1 and 200 [Figure 3(b)]. The missing trace values were set to zero, and the NCGL algorithm was applied to each frequency slice's Hankel matrix of the 2D data.

Figures 4(a), (b), (c), and (d) show the POCS, SVT, LSMM, and the proposed NCGL methods' reconstruction results, respectively. The reconstructed seismic data's SNR values are 20.68, 22.97, 25.73, and 25.83 dB, respectively. Figures 4(e)–(h) show the corresponding reconstruction errors, illustrating the

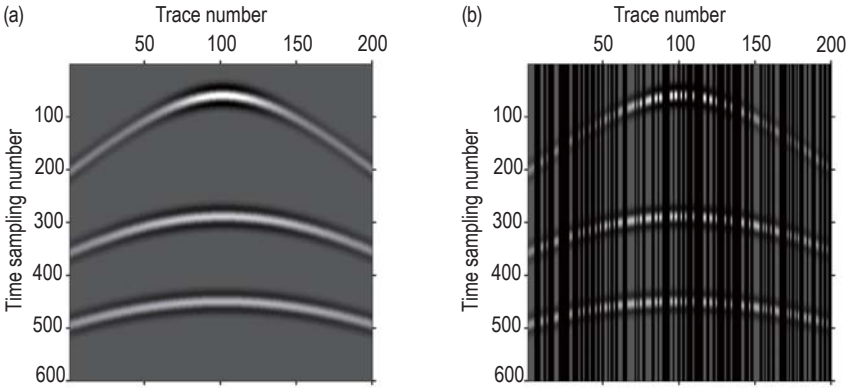
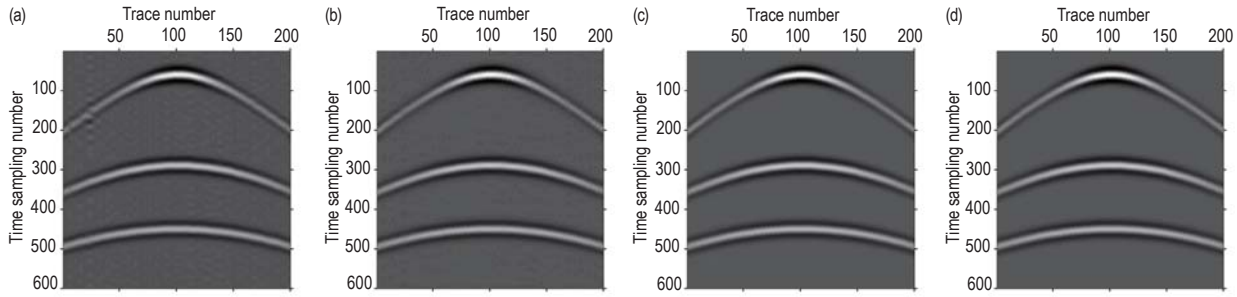


Figure 3. Images showing (a) original seismic data and (b) the corresponding corrupted data with 60% of traces missing.



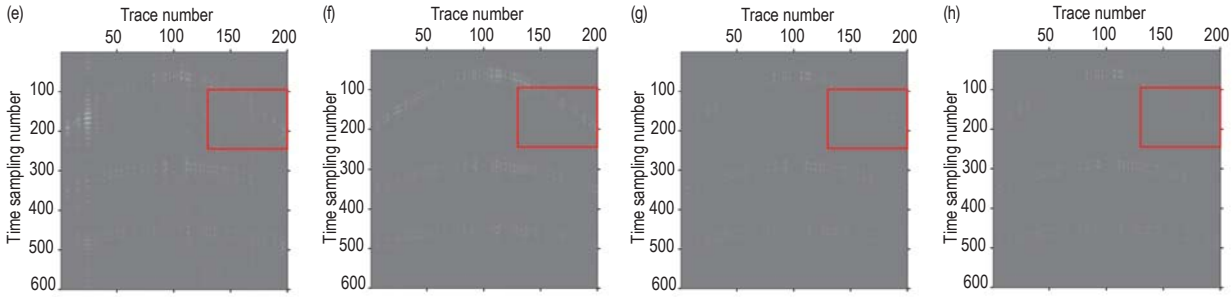


Figure 4. Reconstructed results for 60% randomly missing traces using (a) POCS (SNR = 20.68 dB), (b) SVT (SNR = 22.97 dB), (c) LSMM (SNR = 25.73 dB), and (d) NCGL (SNR = 25.83 dB), (e–h) Reconstructed errors for (e) POCS, (f) SVT, (g) LSMM, and (h) NCGL.

differences between the original and reconstructed data, and the areas in the red box highlight the differences between the NCGL approach and the other three approaches. The reconstruction images, errors, and SNR values demonstrate a more accurate reconstruction using nonconvex regularization than convex relaxation methods.

To illustrate the reconstruction performance of our approach, we compared the images of 198th trace with complete seismic data and the reconstruction results using the different methods [Figures 5(a)–(d)]. The solid green lines represent the complete data, whereas the solid blue lines denote the reconstruction results. Figures 5(e)–(h) show the differences between the 198th trace of the complete and reconstruction data. Figures 5(a)-(h) demonstrate improved reconstruction performance using our method compared with the SVT approach and POCS approach, and are comparable to the LSMM method.

Figure 6 shows the variations of the reconstruction data's SNR values using the different methods against percentages of missing traces varying from 40% to 80%. For all approaches, the higher the missing ratios are, the worse the reconstruction results, although the NCGL method still exhibits better results than the other methods.

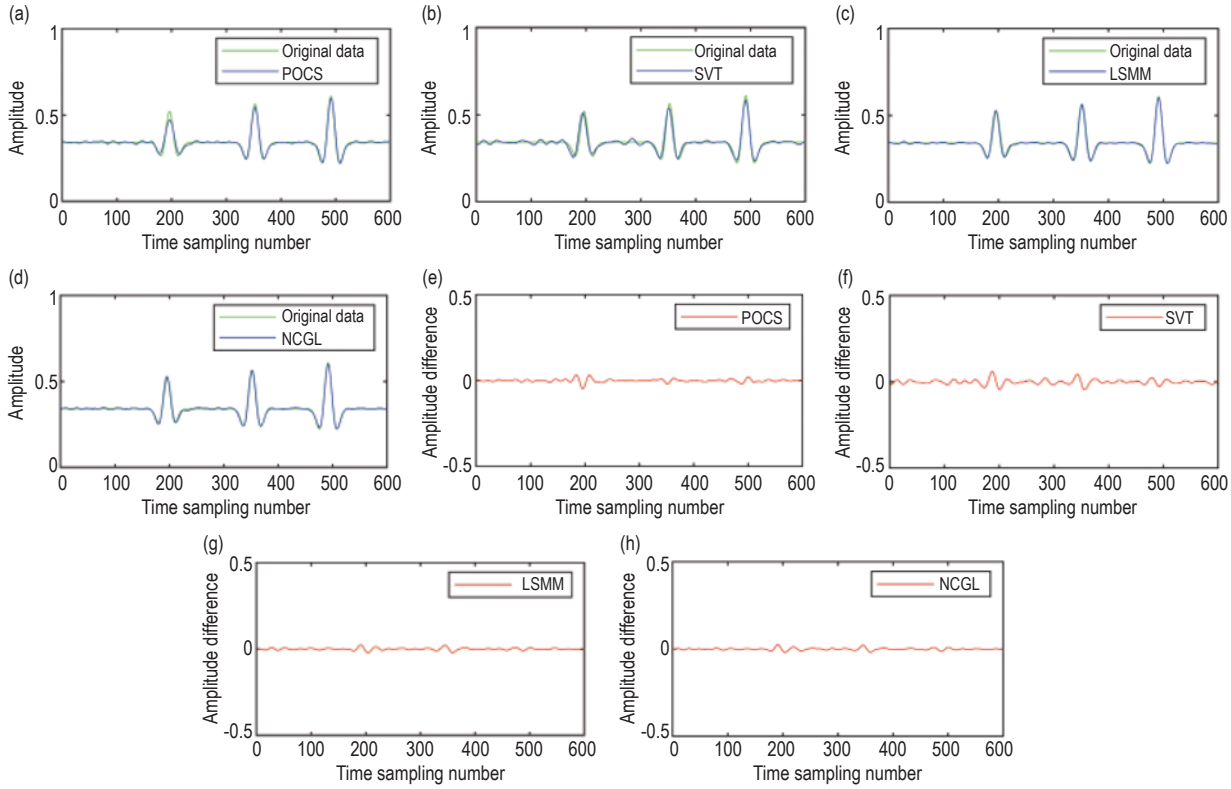


Figure 5. Reconstruction results for the original seismic data set with 60% missing traces showing (a)–(d) comparison of the 198th trace reconstructed using the POCS, SVT, LSMM, and NCGL methods, respectively, and (e)–(h) the corresponding reconstruction errors.

Field data example

We evaluated three methods' performance on a field prestack seismic dataset, comprising 120 traces with 512 sampling points for each trace, a temporal sampling interval of 4 ms, and a spatial sampling interval of 25 m. Figure 7(a) shows the complete seismic data, and Figure 7(b) shows the corresponding corrupted seismic data with 50% of the traces randomly removed. Figures 7(c), (d), (e), and (f) show the reconstruction results for the POCS, SVT, LSMM, and NCGL methods, respectively. Additionally, Figures 7(g), (h), (i), and (j) show the methods' reconstruction errors, respectively, and their corresponding SNR values are 4.22, 7.24, 8.11, and 8.21 dB, respectively. Furthermore, we compared the 97th trace of the original and reconstructed data for the POCS, SVT, LSMM, and NCGL methods [Figures 8(a), (b), (c), and (d), respectively]. Figures 8(e)–(h) show the differences between the 97th trace of the original data and in the reconstructions. Figure 9 shows the reconstruction performance based on the POCS, SVT, LSMM, and NCGL methods for missing traces varying from 10% to 50%. The experimental results reveal that our approach outperforms the convex relaxation method,

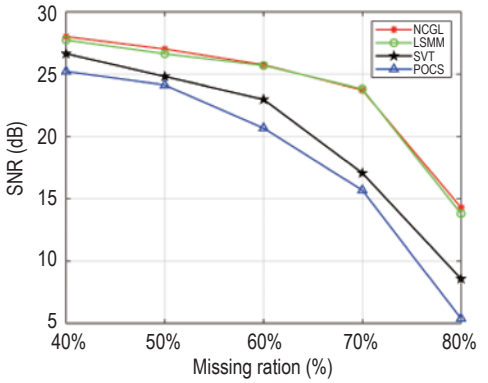


Figure 6. Plot comparing the SNR values for reconstruction based on the POCS, SVT, LSMM, and NCGL methods, respectively, for percentages of missing traces.

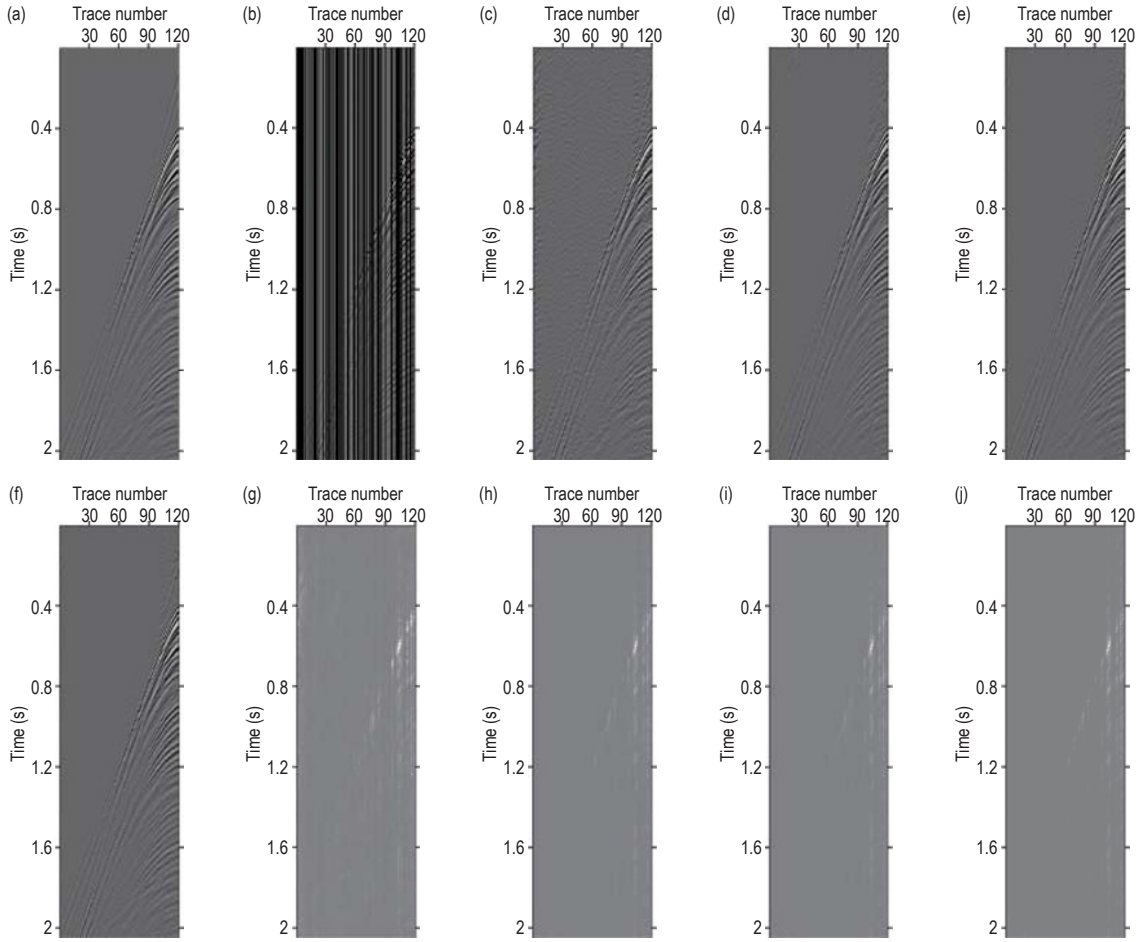


Figure 7. Images displaying (a) original prestack seismic data, (b) corrupted data with 50% randomly missing traces, (c) reconstructed results using POCS (SNR = 4.22 dB), (d) SVT (SNR = 7.24 dB), (e) LSMM (SNR = 8.11 dB), and (f) NCGL (SNR = 8.21 dB); (g–j) Reconstructed errors for (g) POCS, (h) SVT, (i) LSMM, and (j) NCGL.

highlighting a better reconstruction performance of the nonconvex regularization methods. We demonstrate that the NCGL approach's reconstruction efficiency is better than the LSMM method in computation time and convergence speed comparisons.

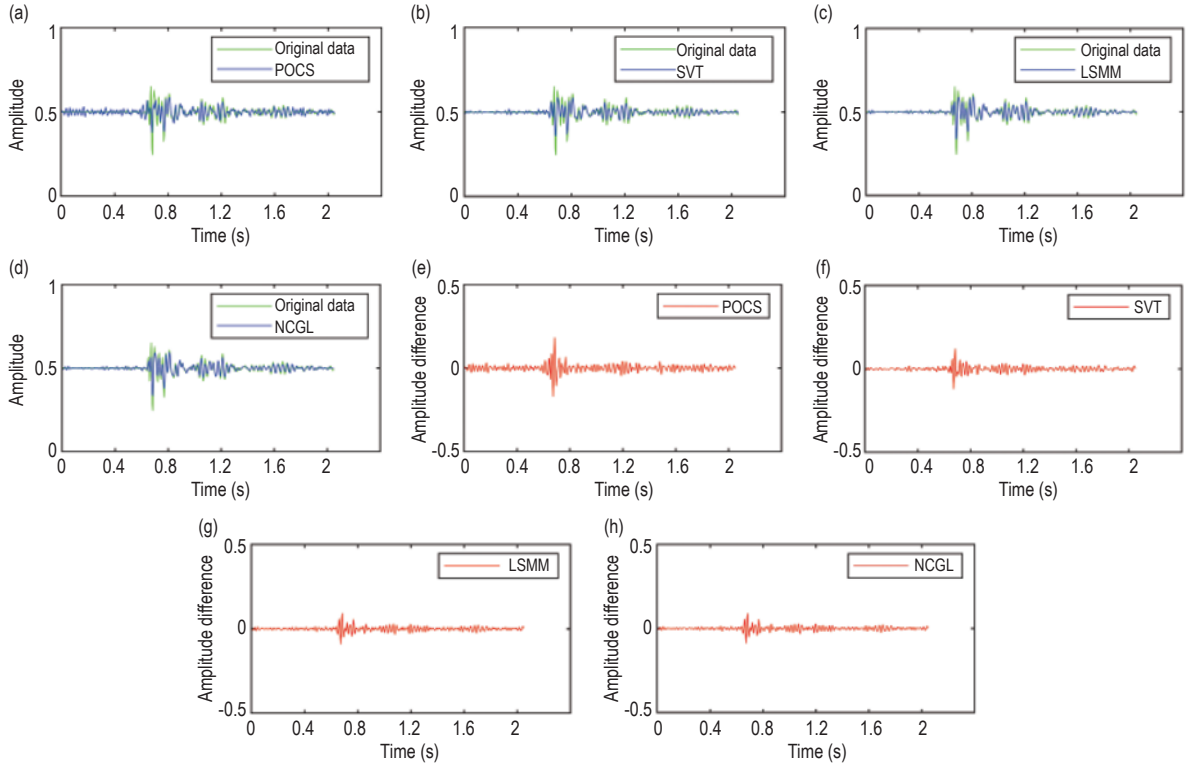


Figure 8. Reconstruction results for an original seismic dataset with 50% missing traces showing (a)–(d) comparison of the 97th trace reconstructed using the POCS, SVT, LSMM, and NCGL methods, respectively, and (e)–(h) the corresponding reconstruction errors.

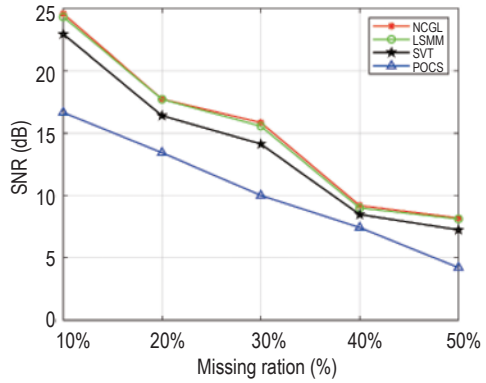


Figure 9. Plots of SNR values associated with reconstruction using three methods versus the percentage of missing traces.

Computation time and convergence speed comparisons

Table 1 shows the computational complexity of the primary steps in Algorithm 1.

Table 1: Summary of the computational complexity of the primary steps in Algorithm 1

Primary steps	Computational complexity
Step 5	$O(n^3)$
Step 7	$O(n^3)$
Step 8	$m(O((\eta n)^3))$ $(0 < \eta < 1)$
Step 9	$O(n^2 m)$

In Step 5, SVD is conducted, and the computational complexity is $O(n^3)$, whereas Step 7 is for updating $\mathbf{D}^k$, including only the matrix products, with computational complexity represented as $O(n^3)$. In Step 8, $\mathbf{Y}^{k+1}$ is updated by updating the rows $\mathbf{y}_j^{k+1}$ $(j,\ldots,n)$ of $\mathbf{Y}^{k+1}$ successively, and the computational complexity depends on the nonzero numbers of $\mathbf{m}_j$. Therefore, the average computational complexity in Step 8 is expressed as $m(O((\eta n))^3)$ when summing the missing ratios $\eta(0<\eta<1)$. The calculation method involves the for loop in MATLAB, which can be replaced by parallel computing in practice to reduce time. In Step 9, $\mathbf{X}_{\omega}^{k+1}$,

which involves a computational complexity represented by $O(n^2m)$, is updated in each iteration.

After a comprehensive analysis, the algorithm converges within J iterations; then, the NCGL method's computational complexity is expressed as $J\max(O(n^2m),m(O((\eta n)^3)))$. Figures 10(a) and (b) show that the root-mean-square error (RMSE) of the adjacent iteration values $\mathbf{X}_{\omega}^{k}$ and $\mathbf{X}_{\omega}^{k+1}$ varies with the number of iterations when the synthetic and field data are fixed at a certain frequency slice. From the figures, NCGL method can converge to a smaller J value faster than LSMM and SVT methods. Therefore, the algorithm presented in this study has lower computational complexity. The POCS method is based on the Fourier transform and excludes the Hankel pretransform. Therefore, this study only compares the convergence speed and time consumption of SVT, LSMM, and NCGL methods.

Furthermore, we evaluated the proposed NCGL method's reconstruction efficiency by ascertaining the time consumed in reconstructing the synthetic and real seismic data. Figure 11 shows the time consumed using each reconstruction approach for varying percentages of missing traces. The proposed method achieved the highest computational efficiency compared with the SVT and LSMM.

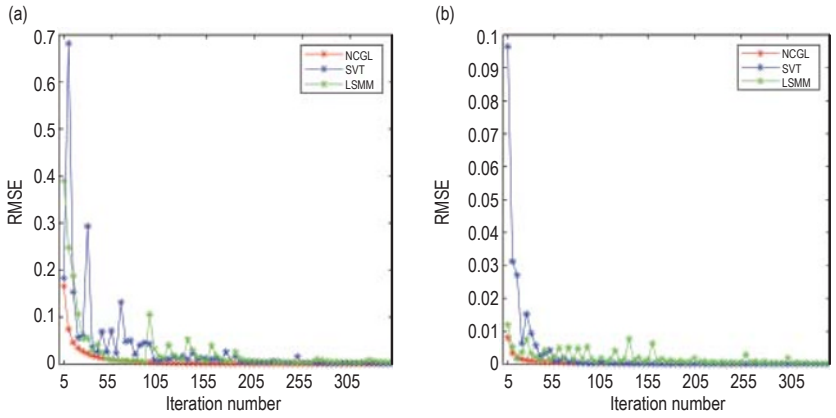


Figure 10. Based on different approaches, the root-mean-square error contributions of $\mathbf{X}_{\omega}^{k}$ and $\mathbf{X}_{\omega}^{k+1}$ for varying iteration numbers when fixed at a certain frequency slice: (a) experimental results of synthetic data, (b) experimental results of field data.

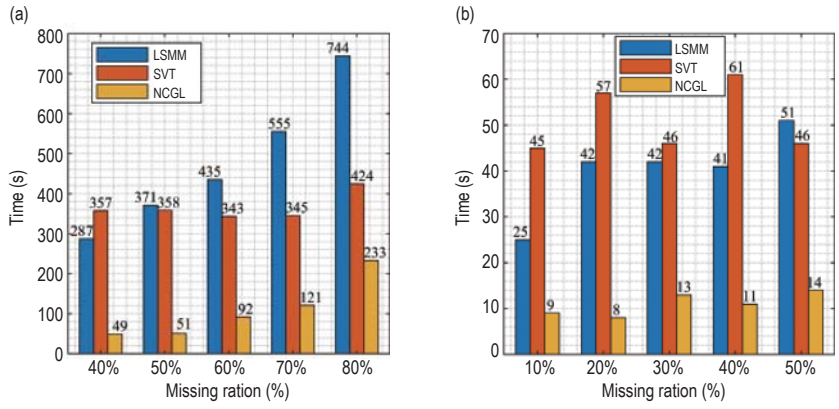


Figure 11. Reconstruction time comparison for three methods based on different data involving varying percentages of missing seismic traces: (a) experimental records on synthetic data; (b) experimental records on field data.

Conclusions

This study proposed a nonconvex regularization method based on the Geman function for reconstructing missing traces in seismic data. The proposed method provides more accurate reconstruction data than conventional convex nuclear norm-based models because the Geman function approximates rank better. Due to its nonconvexity, we developed an efficient optimization model to provide the solution. Further, we used the KKT condition to obtain an iterative solution of the original optimization problem without introducing additional parameters, which is more advantageous than existing nonconvex methods because of the reduced time for tuning parameters. Experiments on synthetic and field data demonstrated that the proposed method has a faster convergence speed and better reconstruction efficiency than SVT and LSMM.

Acknowledgements

This research is financially supported by the National Key R&D Program of China (No. 2018YFC1503705), the Science and Technology Research Project of Hubei Provincial Department of Education (No. B2017597), the Hubei Subsurface Multiscale Imaging Key Laboratory (China University of Geosciences) (No. SMIL-2018-06), and the Fundamental Research Funds for the Central Universities (No. CCNU19TS020).

References

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: Geophysics, **63**(2), 574–588.

Cai, J. F., Candes, E. J., and Shen, Z. W., 2010, A singular value thresholding algorithm for matrix completion: SIAM Journal on Optimization, **20**(4), 1956–1982.

Candès, E. J., Wakin, M. B., and Boyd, S. P., 2008, Enhancing sparsity by reweighted ℓ1 minimization: Journal of Fourier Analysis and Applications, **14**(5), 877–905.

Chen, Y., Zhou, Y., Chen, W., Zu, S., Huang, W., and Zhang, D., 2017, Empirical low-rank approximation for seismic noise attenuation: IEEE Transactions on Geoscience and Remote Sensing, **55**(8), 4696–4711.

Fang, W. Q., Fu, L. H., Zhang, M., and Li, Z. M., 2021, Seismic data interpolation based on U-net with texture loss: Geophysics, **86**(1), V41–V54.

Fu, L. H., Zhang, M., Liu, Z. H., and Li, H. W., 2018, Reconstruction of seismic data with missing traces using normalized Gaussian weighted filter: Journal of Geophysics and Engineering, **15**(5), 2009–2020.

Gao, J. J., Cheng, J. K., and Sacchi, M. D., 2017, Five-dimensional seismic reconstruction using parallel square matrix factorization: IEEE Transactions on Geoscience and Remote Sensing, **55**(4), 2124–2135.

Gao, J. J., Stanton, A., Naghizadeh, M., Sacchi, M. D., and Chen, X. H., 2013, Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction: Geophysical Prospecting, **61**(S1), 138–151.

Geman, D., and Yang, C. D., 1995, Nonlinear image recovery with half-quadratic regularization, IEEE Transactions on Image Processing: **4**(7), 932–946.

Kabir, M. M., and Verschuur, D. J., 1995, Restoration of missing offsets by parabolic Radon transform: Geophysical Prospecting, **43**(3), 347–368.

Liu, L., Plonka, G., and Ma, J. W., 2017a, Seismic data interpolation and denoising by learning a tensor tight frame: Inverse Problems, **33**(10), 1–32.

Liu, Q., Fu, L. H., and Zhang, M., 2021, Deep-seismic-prior-based reconstruction of seismic data using convolutional neural networks: Geophysics, **86**(2), V131–V142.

Liu, Y., Zhang, P., and Liu, C., 2017b, Seismic data interpolation using generalised velocity-dependent seislet transform: Geophysical Prospecting, **65**(S1), 82–93.

Lu, C. Y., Tang, J. H., Yan, S. C., and Lin, Z. C., 2016, Nonconvex nonsmooth low rank minimization via Iteratively reweighted nuclear norm: IEEE Transactions on Image Processing, **25**(2), 829–839.

Ma, J. W., 2013, Three-dimensional irregular seismic data reconstruction via low-rank matrix completion: Geophysics, **78**(5), V181–V192.

Ma, S. Q., Goldfarb, D., and Chen, L. F., 2011, Fixed point and Bregman iterative methods for matrix rank minimization: Mathematical Programming, **128**(1), 321–353.

Magnus, J. R., 1985, On differentiating eigenvalues and eigenvectors: Econometric Theory, **1**(2), 179–191.

Marshall, A. W., Olkin, I., and Arnold, B. C., 1979, Inequalities: theory of majorization and its applications: Academic Press, New York, 1–909.

Oropeza, V., and Sacchi, M., 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**(3), V25–V32.

Toh, K. C., and Yun, S. W., 2010, An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems: Pacific Journal of Optimization, **6**(15), 615–640.

Trickett, S., Burroughs, L., Milton, A., Walton, L., and Dack. R., 2010, Rank-reduction-based trace interpolation: Society of Exploration Geophysicists, 3829–3833.

Wang, B. F., Li, J. Y., Chen, X. H., and Cao, J. J., 2015a, Curvelet-based 3D reconstruction of digital cores using the POCS method: Chinese Journal of Geophysics-Chinese Edition, **58**(5), 486–495.

Wang, B. F., Wu, R. S., Chen, X. H., and Li, J. Y., 2015b, Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform: Geophysical Journal International, **201**(2), 1182–1194.

Witten, B., and Shragge, J., 2015, Extended wave-equation imaging conditions for passive seismic data: Geophysics, **80**(6), WC61–WC72.

Yang, J. F., and Yuan, X. M., 2013, Linearized augmented lagrangian and alternating direction methods for nuclear norm minimization: Mathematics of Computation, **82**(281), 301–329.

Yang, Y., Ma, J. W., and Osher, S., 2013, Seismic data reconstruction via matrix completion: Inverse Problems & Imaging, **7**(4), 1379–1392.

Zhang, D., Chen, Y. K., Huang, W. L., and Gan, S. W., 2016, Multi-step damped multichannel singular spectrum analysis for simultaneous reconstruction and denoising of 3D seismic data: Journal of Geophysics and Engineering, **13**(5), 704–721.

Zhang, D., Zhou , Y. T., Chen, H. M., Chen, W., Zu S. H., and Chen Y. K., 2017, Hybrid rank-sparsity constraint model for simultaneous reconstruction and denoising of 3D seismic data: Geophysics, **82**(5), V351–V367.

Zhang, W. J., Fu, L. H., and Liu, Q., 2019, Nonconvex log-sum function-based majorization–minimization framework for seismic data reconstruction: IEEE Geoscience and Remote Sensing Letters, **16**(11), 1776–1780.

Zhang, W. J., Fu, L. H., Zhang, M., and Cheng, W. T., 2020, 2-D Seismic data reconstruction via truncated nuclear norm regularization: IEEE Transactions on Geoscience and Remote Sensing, **58**(9), 6336–6343.

Appendix

Convergence proof of the proposed NCGL

Before describing our theoretical proof, we introduce two Lemmas. **Lemma 1** (Marshall et al., 1979): Assuming $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n\times n}$ are two Hermitian matrices with eigenvalues in identical order, we can consider that the following inequality is valid

$$\sum_{i=1}^{n}\lambda_i(\mathbf{X})\lambda_{n-i+1}(\mathbf{Y}) \le Tr(\mathbf{XY}) \le \sum_{i=1}^{n}\lambda_i(\mathbf{X})\lambda_i(\mathbf{Y}), \quad (30)$$

Further, equality exists if, and only if, two orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$ satisfying $\mathbf{U}\mathbf{\Lambda}_{\mathbf{X}}\mathbf{V}^T = \mathbf{X}, \mathbf{U}\mathbf{\Lambda}_{\mathbf{Y}}\mathbf{V}^T = \mathbf{Y}$ can be found, where $\mathbf{\Lambda}_{\mathbf{X}}$ and $\mathbf{\Lambda}_{\mathbf{Y}}$ are the ordered eigenvalue matrices.

Lemma 2: The $\mathbf{X}_{\omega}^{k+1}$ and $\mathbf{Y}^{k+1}$ in equation (28) satisfy the KKT condition of the following convex optimization problem

$$\min_{\mathbf{X}_{\omega}} Tr((\mathbf{D}^k)^T h(\mathbf{X}_{\omega})) \qquad s.t.\ \mathbf{W}\odot\mathbf{X}_{\omega} = \mathbf{M}_{\omega}. \quad (31)$$

Theorem: In the kth iteration of Algorithm 1, assume $\mathbf{X}_{\omega}^{k}$ is the kth iterative optimal solution, and we obtain

$$\sum_{i=1}^{n}\frac{\sigma_i(\mathbf{X}_{\omega}^{k+1})}{\sigma_i(\mathbf{X}_{\omega}^{k+1})+\gamma} \le \sum_{i=1}^{n}\frac{\sigma_i(\mathbf{X}_{\omega}^{k})}{\sigma_i(\mathbf{X}_{\omega}^{k})+\gamma}.$$

Proof: By adding Lemma 2, assuming $\mathbf{X}_{\omega}^{k+1}$ is the (k+1)th iterative solution, we obtain

$$Tr((\mathbf{D}^k)^T h(\mathbf{X}_{\omega}^{k+1})) \le Tr((\mathbf{D}^k)^T h(\mathbf{X}_{\omega}^{k})), \quad (32)$$

that is,

$$Tr(\mathbf{U}^k[\mathbf{v}^k](\mathbf{U}^k)^T\mathbf{U}^{k+1}[\boldsymbol{\lambda}^{k+1}](\mathbf{U}^{k+1})^T) \le Tr(\mathbf{U}^k[\mathbf{v}^k](\mathbf{U}^k)^T\mathbf{U}^k[\boldsymbol{\lambda}^k](\mathbf{U}^k)^T). \quad (33)$$

Considering $\lambda_1 \ge \lambda_1 \ge \cdots \ge \lambda_n$, according to the Geman properties, we obtain $v_1 \ge v_1 \ge \cdots \ge v_n$. Further, according to Lemma 1 and equation (33), the following inequality is valid,

$$\sum_{i=1}^{n}v_i^k\lambda_i^{k+1} = Tr([\mathbf{v}^k][\boldsymbol{\lambda}^{k+1}]) \le Tr(\mathbf{U}^k[\mathbf{v}^k](\mathbf{U}^k)^T\mathbf{U}^{k+1}[\boldsymbol{\lambda}^{k+1}](\mathbf{U}^{k+1})^T)$$
$$\le Tr(\mathbf{U}^k[\mathbf{v}^k](\mathbf{U}^k)^T\mathbf{U}^k[\boldsymbol{\lambda}^k](\mathbf{U}^k)^T) = Tr([\mathbf{v}^k][\boldsymbol{\lambda}^k]) = \sum_{i=1}^{n}v_i^k\lambda_i^k. \quad (34)$$

The Geman function is nonconvex, therefore,

$$\frac{(\lambda_i^{k+1})^{0.5}}{(\lambda_i^{k+1})^{0.5}+\gamma} - \frac{(\lambda_i^{k})^{0.5}}{(\lambda_i^{k})^{0.5}+\gamma} \le v_i^k(\lambda_i^{k+1}-\lambda_i^k). \quad (35)$$

Then,

$$\sum_{i=1}^{n}\frac{(\lambda_i^{k+1})^{0.5}}{(\lambda_i^{k+1})^{0.5}+\gamma} - \sum_{i=1}^{n}v_i^k\lambda_i^{k+1} \le \sum_{i=1}^{n}\frac{(\lambda_i^{k+1})^{0.5}}{(\lambda_i^{k+1})^{0.5}+\gamma} - \sum_{i=1}^{n}v_i^k\lambda_i^k. \quad (36)$$

Further, based on equations (35) and (36), we obtain

$$\sum_{i=1}^{n}\frac{(\lambda_i^{k+1})^{0.5}}{(\lambda_i^{k+1})^{0.5}+\gamma} \le \sum_{i=1}^{n}\frac{(\lambda_i^{k})^{0.5}}{(\lambda_i^{k})^{0.5}+\gamma}., \quad (37)$$

which can be rewritten as

$$\sum_{i=1}^{n}\frac{\sigma_i^{k+1}}{\sigma_i^{k+1}+\gamma} \le \sum_{i=1}^{n}\frac{\sigma_i^{k}}{\sigma_i^{k}+\gamma}, \quad (38)$$

where σ_i^k denotes the ith singular value of $\mathbf{X}_{\omega}^k$.

Theorem demonstrates that Algorithm 1 monotonically reduces the objective function value, making the algorithm convergent.

Li Yan-Yan received a B.S. degree in mathematics and applied mathematics from Henan Normal University, China, in 2016. She is currently pursuing an M.S. degree at the School of Mathematics and Physics, China University of Geosciences, Wuhan, China. Her research interests include seismic data processing.

Fu Li-Hua received B.S. and M.Sc. degrees in mathematics from Hubei University, Wuhan, China, in 2000 and 2003, respectively, and a Ph.D. degree in geoexploration and information technology from China University of Geosciences, Wuhan, in 2009. She is currently a professor at the School of Mathematics and Physics, China University of Geosciences. Her research interests include seismic data processing, deep learning, and computer vision.