Robust reconstruction via Group Sparsity with Radon operators

Ji Li and Daniel Trad

University of Calgary, CRWES

Summary

Sparse solutions of linear systems play an essential role in seismic data processing, including denoising and interpolation. An additional structure called group sparsity can be used to improve the performance of the sparse inversion. We propose a robust Orthogonal Matching Pursuit algorithm with the Radon operators in the frequency slowness $(w-p)$ domain to solve the strong-group sparsity problem, which is used to interpolate seismic data. A modified ADMM solver is used to solve the $\ell_1 - \ell_1$ problem, which simultaneously makes the algorithm robust to the erratic noise. We compare the performance of the same method with and without the group sparsity constraint. The result shows that the strong group sparsity inversion performs better than the traditional sparsity inversion. Both synthetic and real seismic data are being tested to examine the performance of the proposed algorithm.

Introduction

Sparse representation is an important tool for signal processing, including seismic data processing. Sparse representation means the seismic signal is sparse (the number of nonzero coefficients k is much smaller than the total number of coefficients) in some transform domains like Radon, Fourier and Wavelet.

On the other hand, an additional group structure can be added to the problem, often referred to as group sparsity or group Lasso (Bach, 2008; Yuan and Lin, 2006; Nardi and Rinaldo, 2008) to improve the accuracy of the sparse estimation. Many methods have been proposed in seismic data processing with the idea of group sparsity. Li and Sacchi (2022) and Naghizadeh (2012) group the coefficients into different dips in the frequency wavenumber $(f-k)$ domains for seismic data denoising and interpolation. Vera Rodriguez et al. (2012) use group sparsity for microseismic data denoising. Trad et al. (2002) use a similar idea for fast calculation of the hyperbolic Radon transform. Chen et al. (2019) combine the group sparsity and total variation for seismic signal denoising.

One interesting goal is to develop a robust sparse solver to solve the inversion problem for seismic data corrupted by erratic noise. For instance, Guitton and Symes (2003) replaces the ℓ_2 norm with $Huber$ norm for the residual term to cope with the seismic data with outliers.Trickett et al. (2012) uses a rank reduction filter to remove the erratic noise in the seismic data. More recently, Li and Sacchi (2021) developed a sparse and robust Radon transform, estimated via Matching Pursuit, to solve the simultaneous source separation problem. In this paper, we cooperate the robust inversion with group sparsity to get a better sparse estimation.

Theory

A seismic signal can be represented as follows:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{1}$$

Where $\mathbf{y}$ denotes the signal, $\mathbf{x}$ is the coefficients, and $\mathbf{A}$ is a synthesis operator or matrix which can transfer the coefficients into the seismic signal. $\mathbf{e}$ represents additive noise in the signal. In this situation, the sparse approximation problem can be expressed as

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 \leq \delta, \tag{2}$$

where $\|\mathbf{x}\|_0$ is the ℓ_0 norm, which measure the number of nonzero coefficients of $\mathbf{x}$. and $\|.\|_2^2$ symbolizes the ℓ_2 norm (equal to the square root of the the inner product of a vector with itself) used to fit the residuals. This problem is a combinatorial, nondeterministic polynomial time (NP)-hard problem for which finding an exact solution is prohibitively expensive. Generally, two groups of methods can be used to solve the problem 2. One is the convex relaxation, replacing the ℓ_0 norm with the ℓ_1 norm Chen et al. (2001). Which transfers the problem into a convex optimization problem.

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1. \tag{3}$$

Which is also known as Lasso (Tibshirani, 1996), and can be solved by many methods like FISTA (Beck and Teboulle, 2009), IRLS (Scales and Gersztenkorn, 1988) and ADMM boyd2011distributed. Another approach is the greedy method like Matching Pursuit (MP) (Mallat and Zhang, 1993) and Orthogonal Matching Pursuit (OMP) (Tropp and Gilbert, 2007; Pati et al., 1993).

In some cases, adding a group structure constraint can yield a better estimation (Majumdar and Ward, 2009; Elhamifar and Vidal, 2011). Unlike the standard sparsity assumption, the sparsity is measured by the number of nonzero coefficients k. A group-sparse vector can be divided into groups so that few groups contain nonzero coefficients, but the groups do not need to be sparse.

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \beta \sum_{i=1}^{N} \|\mathbf{x}[i]\|_2. \tag{4}$$

A further refinement, called strong group sparsity, can be made to improve performance (Vincent and Hansen, 2014; Simon et al., 2013). For the strong sparsity problem, the coefficients are not only located within a few groups, but the nonzero groups are also sparse.

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 + \beta \sum_{i=1}^{N} \|\mathbf{x}[i]\|_2. \tag{5}$$

When the erratic noise corrupts the data, we can use M-estimators like $Gaussian-norm$,$Huber-norm$ or ℓ_1-norm to replace the ℓ_2-norm for the error term. For our case, we will use the ℓ_1-norm to replace the ℓ_2-norm. Then the problem changes to

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_1^1 + \lambda \|\mathbf{x}\|_1 + \beta \sum_{i=1}^{N} \|\mathbf{x}[i]\|_2. \tag{6}$$

To solve the problem 6, we use the linear Radon transform to synthesize the seismic signal. And we use the orthogonal matching pursuit with the $\ell_1-\ell_1$ ADMM solver (Wen et al., 2016) to estimate the sparse Radon coefficients.

By using the linear Radon operators in the $f-p$ domain, the problem becomes to minimize the following cost function.

$$\operatorname{argmin} \|\mathscr{L}^*\mathbf{m}-\mathbf{d}\|_1 + \lambda \|\mathbf{m}\|_1 + \beta \sum_{i=1}^{N} \|\mathbf{m}[i]\|_2. \tag{7}$$

where $\mathbf{L}^*$ is the forward Radon operator, which transfers the Radon coefficients in the $f-p$ domain to the $t-x$ domain. **m** is the Radon coefficients in the $f-p$ domain. The $\|\mathscr{L}^*\mathbf{m}-\mathbf{d}\|_1$ make the cost function robust to the erratic noise. $\beta\sum_{i=1}^{N}\|\mathbf{m}[i]\|_2$ is use to promote the group sparsity, and $\lambda\|\mathbf{m}\|_1$ can enhance the sparsity within the groups. To solve this problem, we can use the greedy method; in our case, we use the orthogonal Matching Pursuit (OMP)(Tropp and Gilbert, 2007; Pati et al., 1993).

To solve the minimization problem 7, we need to modify the OMP algorithm. Instead of picking one coefficient each time, we pick a group of coefficients for the basis function selection part. We divide our coefficients into different groups based on the slowness p. Then we pick the group which has the maximum ℓ_2 norm. And then, we utilize all the coefficients located with all currently pick slowness p_{T^k} by minimizing the following cost function

$$\widehat{\mathbf{m}}_{T^k}^k = \underset{\widetilde{\mathbf{m}}_{T^k}}{\operatorname{argmin}} \|\mathbf{d} - \mathscr{L}_n^T(\widetilde{\mathbf{m}}_{T^k})\|_1 + \lambda \|\mathbf{m}\|_1. \tag{8}$$

This $\ell 1 - \ell 1$ minimization problem can be solved by the $\ell 1 - \ell 1$ ADMM solver (Yang and Zhang, 2011; Wen et al., 2016).

Example

Figure 1 (a) represents a simple 2D synthetic example with three linear events. Part (b) is the shot gather after we removed 75% of the traces randomly and added the erratic noise manually. Figure 2

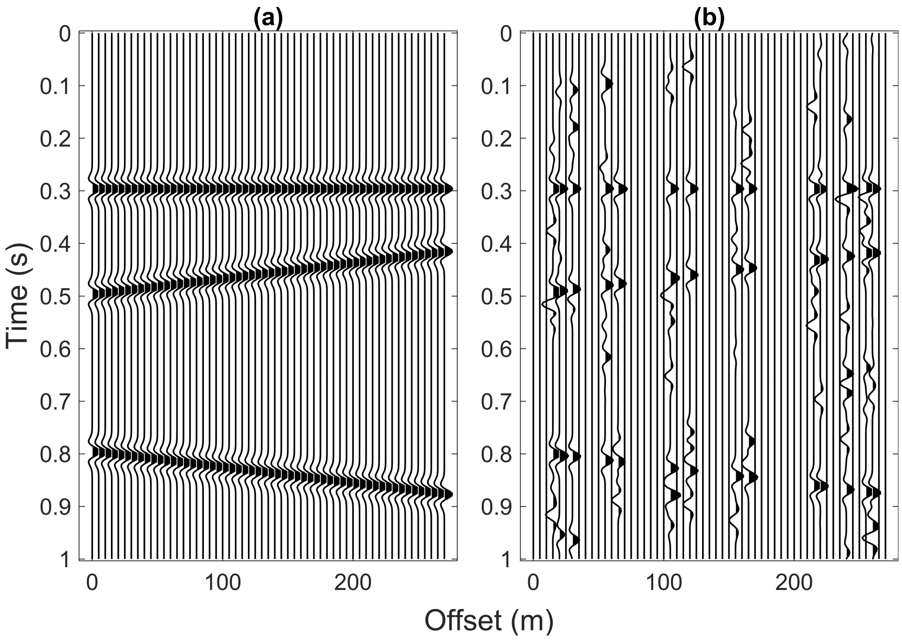

Figure 1: Reconstructed result in the $t-x$ domain. (a) Clean data. (b) Synthetic data with erratic noise and 75% of traces missed.

shows the denoised results of the same method with different constraints. As expected, the $\ell_2 - \ell_1$ inversion, which is not robust to the erratic noise, has the worst performance. The $\ell_1 - \ell_1$ inversion performs better than the $\ell_2 - \ell_1$. Since the data includes a lot of erratic noise with many missed traces, the $\ell_1 - \ell_1$ inversion method still has much noise. Part (c) is the $\ell_2 - \ell_1$ with group sparsity. As we can see, even with the $\ell_2 - \ell_1$ inversion, the performance is better than using the $\ell_1 - \ell_1$ inversion without group sparsity. And finally, part (g) is the denoised result with $\ell_1 - \ell_1$ inversion and group sparsity, which has the best performance.

We also test the proposed algorithm with a 3D data patch from the Western Canadian Basin survey. This example shows a single window extracted from a prestack volume. The 3D patch includes 301-time samples and 30 x 30 midpoints extracted from a constant x- and y- offset volume. The prestack volum was NMO corrected to limit the number of dips. We also apply a low-pass filter with a frequency cutoff of 70Hz to avoid interpolating high-frequency noise. Then we applied the proposed method to interpolate this single patch. For this example, we use the 3D Radon operators and compare the interpolation result with another conventional multidimensional interpolation method. In our case, we

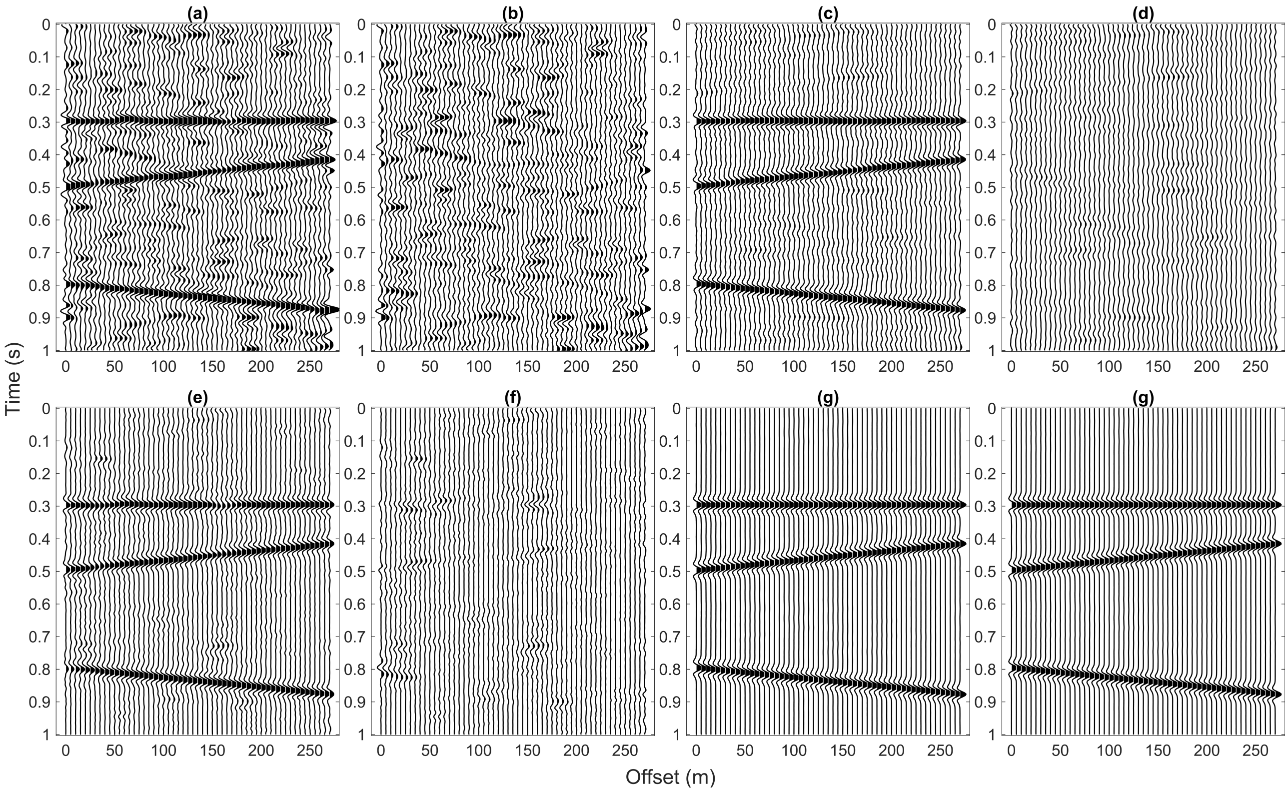

Figure 2: Reconstructed result in the $t-x$ domain. (a) $\ell_2-\ell_1$ without group sparsity, snr=2.35 dB. (b) Errors between a and clean data. (c) $\ell_2-\ell_1$ with group sparsity, snr=6.9 dB. (d) Errors between c and clean data. (e) $\ell_1-\ell_1$ without group sparsity, snr=5.22 dB. (f) Errors between e and clean data. (g) $\ell_1-\ell_1$ with group sparsity, snr=14.3 dB. (h) Errors between g and clean data.

compare it with the POCS (Abma and Kabir, 2006). Figure 3 presents the final results. Figure 3 (a) shows six slides of the 3D patch. The original data has many missing traces with large gaps. It also includes significant noise of unknown distribution and some bad traces; both can be treated as erratic noise. Figure 3 (b) presents the results after POCS reconstruction. This result shows that when too many traces are missing, the gaps between the traces become too large and the algorithm's output deteriorates. And also, POCS can not remove the erratic noise at the same time. Figure 3 (c) shows the reconstruction result using the proposed robust method. The reconstruction result of the proposed robust strong group sparse method shows significant improvement compared to POCS.

Conclusion

We proposed a robust group sparse inversion algorithm which can provide a better sparse estimation and be robust to the erratic noise simultaneously. The core of the proposed method is based on the orthogonal Matching Pursuit. We divide the Radon coefficients in the $f-p$ domain into different slowness groups. In each iteration, the algorithm selected the group with the maximum norm. And then, the Radon coefficients located within all currently selected groups are directly fitting to the seismic data in the $t-x$ domain. The coefficient optimization part is solved by the modified $\ell_1-\ell_1$ ADMM solver, which makes the algorithm robust to the erratic noise. The tests on both 2D and 3D synthetic examples prove the effectiveness and robustness of the proposed method. Furthermore, compared with the traditional MP and OMP algorithm, the proposed algorithm can save the total costs a lot since it selects one group with multiple coefficients instead of picking just one best-correlated coefficient like MP and OMP.

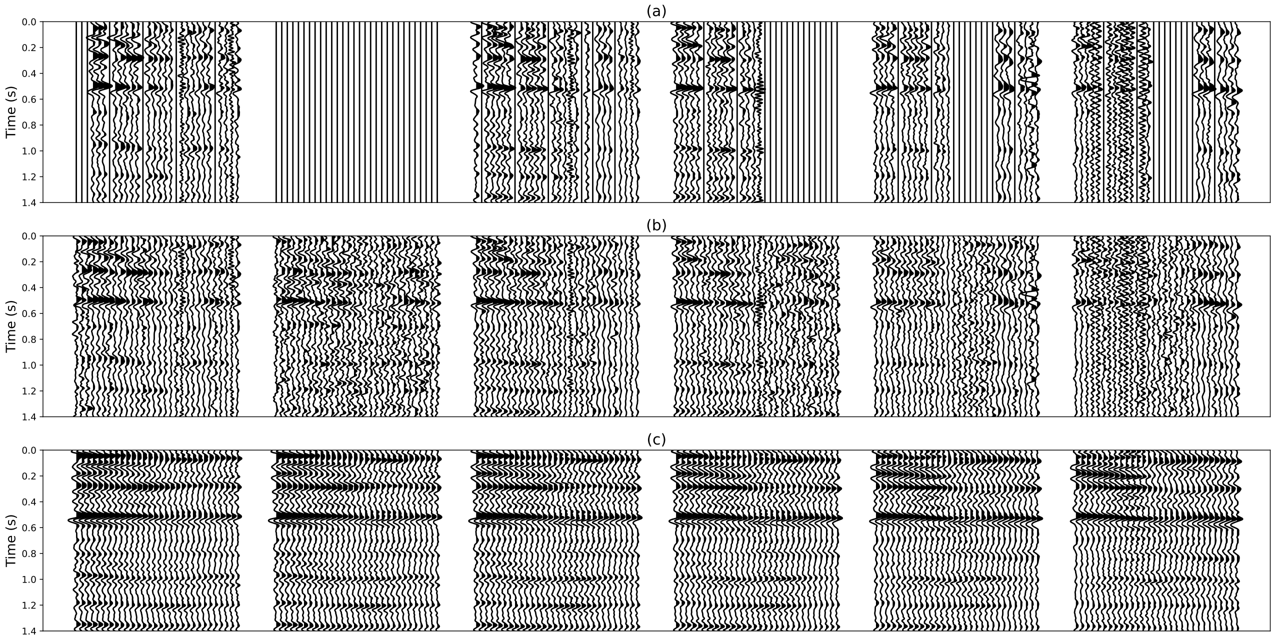

Figure 3: (a) Six slides of a 3D cube real data. (b) POCS interpolation. (c) Robust strong group sparsity interpolation

Acknowledgements

We thank the sponsors of CREWES for their continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19.

References

Abma, R., and N. Kabir, 2006, 3d interpolation of irregular data with a pocs algorithm: Geophysics, **71**, E91–E97.
Bach, F. R., 2008, Consistency of the group lasso and multiple kernel learning.: Journal of Machine Learning Research, **9**.
Beck, A., and M. Teboulle, 2009, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: SIAM journal on imaging sciences, **2**, 183–202.
Chen, S. S., D. L. Donoho, and M. A. Saunders, 2001, Atomic decomposition by basis pursuit: SIAM review, **43**, 129–159.
Chen, Y., Z. Peng, M. Li, F. Yu, and F. Lin, 2019, Seismic signal denoising using total generalized variation with overlapping group sparsity in the accelerated admm framework: Journal of Geophysics and Engineering, **16**, 30–51.
Elhamifar, E., and R. Vidal, 2011, Robust classification using structured sparse representation: CVPR 2011, IEEE, 1873–1879.
Guitton, A., and W. W. Symes, 2003, Robust inversion of seismic data using the huber norm: Geophysics, **68**, 1310–1319.
Li, J., and M. D. Sacchi, 2021, An lp-space matching pursuit algorithm and its application to robust seismic data denoising via time-domain radon transforms: Geophysics, **86**, V171–V183.
——, 2022, Robust reconstruction via orthogonal matching pursuit with fourier operators: Second International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists and American Association of Petroleum ..., 2947–2951.
Majumdar, A., and R. K. Ward, 2009, Classification via group sparsity promoting regularization: 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 861–864.
Mallat, S. G., and Z. Zhang, 1993, Matching pursuits with time-frequency dictionaries: IEEE Transactions on signal processing, **41**, 3397–3415.
Naghizadeh, M., 2012, Seismic data interpolation and denoising in the frequency-wavenumber domain: Geophysics, **77**, V71–V80.
Nardi, Y., and A. Rinaldo, 2008, On the asymptotic properties of the group lasso estimator for linear models: Electronic Journal of Statistics, **2**, 605–633.
Pati, Y. C., R. Rezaiifar, and P. S. Krishnaprasad, 1993, Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition: Proceedings of 27th Asilomar conference on signals, systems and computers, IEEE, 40–44.
Scales, J. A., and A. Gersztenkorn, 1988, Robust methods in inverse theory: Inverse problems, **4**, 1071.
Simon, N., J. Friedman, T. Hastie, and R. Tibshirani, 2013, A sparse-group lasso: Journal of computational and graphical statistics, **22**, 231–245.
Tibshirani, R., 1996, Regression shrinkage and selection via the lasso: Journal of the Royal Statistical Society: Series B (Methodological), **58**, 267–288.
Trad, D. O., T. J. Ulrych, and M. D. Sacchi, 2002, Accurate interpolation with high-resolution time-variant radon transforms: Geophysics, **67**, 644–656.

Trickett, S., L. Burroughs, and A. Milton, 2012, Robust rank-reduction filtering for erratic noise: Presented at the 2012 SEG Annual Meeting, OnePetro.
Tropp, J. A., and A. C. Gilbert, 2007, Signal recovery from random measurements via orthogonal matching pursuit: IEEE Transactions on information theory, **53**, 4655–4666.
Vera Rodriguez, I., D. Bonar, and M. Sacchi, 2012, Microseismic data denoising using a 3c group sparsity constrained time-frequency transform: Geophysics, **77**, V21–V29.
Vincent, M., and N. R. Hansen, 2014, Sparse group lasso and high dimensional multinomial classification: Computational Statistics & Data Analysis, **71**, 771–786.
Wen, F., P. Liu, Y. Liu, R. C. Qiu, and W. Yu, 2016, Robust sparse recovery in impulsive noise via $\ell_p - \ell_1$ optimization: IEEE Transactions on Signal Processing, **65**, 105–118.
Yang, J., and Y. Zhang, 2011, Alternating direction algorithms for ℓ_1-problems in compressive sensing: SIAM journal on scientific computing, **33**, 250–278.
Yuan, M., and Y. Lin, 2006, Model selection and estimation in regression with grouped variables: Journal of the Royal Statistical Society: Series B (Statistical Methodology), **68**, 49–67.