Application of Robust Principal Component Analysis (RPCA) to suppress erratic noise in seismic records

Jinkun Cheng, Ke Chen and Mauricio D. Sacchi, University of Alberta*

SUMMARY

Seismic data are always contaminated with noise. Therefore, signal-to-noise ratio enhancement plays an important role in seismic data processing. This paper illustrates a robust principal component analysis (RPCA) method to suppress erratic noise that contaminates seismic data. The method operates in the frequency-space domain and relies on a robust low-rank approximation of the seismic data volume. We adopt a nuclear norm constraint that yields the low-rank approximation of the desired data while using an ℓ_1 norm constraint to properly estimate the erratic (sparse) noise. The problem is then tackled via the first-order gradient iteration method with two steps of soft-thresholding. We illustrate the effectiveness of this method via synthetic examples.

INTRODUCTION

Principal component analysis (PCA) is an important tool for multivariate analysis in statistics. The idea is to reduce the dimensionality of a data set while preserving as much variability of data variables as possible (Jolliffe, 2010). Let us consider to recover a low-rank matrix $\mathbf{L}$ from the observed data

$$\mathbf{D} = \mathbf{L} + \mathbf{E}, \tag{1}$$

where $\mathbf{E}$ is a matrix representing the additive error. If we assume $\mathbf{E}$ is composed by small random perturbations, an optimal estimate of $\mathbf{L}$ can be acquired via the following optimization problem

$$\begin{aligned} &\min \ \|\mathbf{E}\|_F^2 \\ &s.t. \quad \text{rank}(\mathbf{L}) = k, \quad \mathbf{D} = \mathbf{L} + \mathbf{E}. \end{aligned} \tag{2}$$

The problem can be efficiently solved via singular value decomposition (SVD) (Golub and van Loan, 1996). The observed data $\mathbf{D}$ can be decomposed into a group of eigen-images via the SVD. The low-rank component $\mathbf{L}$ can be described with a few eigen-images that are associated to the largest singular values. The error $\mathbf{E}$, however, will have energy spread over all the eigen-images (Trickett, 2003).

A variety of methods based on PCA have been developed in seismic data processing. For instance, Ulrych et al. (1999) introduced a time domain matrix rank reduction method to eliminate incoherent noise from seismic records. A related family of methods, the Karhunen-Loeve transform, has also been introduced for the enhancement of the signal-to-noise ratio of prestack gathers (Al-Yahya, 1991). Dipping events can be handled by applying eigen-image filters in the $f-x-y$ domain (Trickett, 2003). Similarly, Cadzow de-noising (Trickett and Burroughs, 2009), or singular spectrum analysis (SSA) (Sacchi, 2009; Oropeza and Sacchi, 2011) also operates in the frequency-space domain and are capable of preserving coherent dips while attenuating random noise. Cadzow and SSA de-noising, unlike traditional eignen-images methods (Ulrych et al., 1999), do not operate on the spatial data themselves but on Hankel matrices that are formed from data in the $f-x$ or $f-x-y$ domain.

Although the aforedescribed rank reduction methods are very effective techniques for attenuating random Gaussian noise, their applications to real data problems are limited due to lack of robustness to erratic noise. In seismic data processing, erratic noise includes swell noise, power line noise and artifacts caused by glitches in recording instruments. Outliers tend to manifest as high-amplitude isolated signals that do not obey the Gaussian distribution. Therefore, the conventional least-squares error criterion utilized by PCA will perform poorly (Golub and van Loan, 1996; Trickett et al., 2012; Chen and Sacchi, 2014). In this article, we replace the least-squares constraint by an ℓ_1 norm constraint where we have assumed that the erratic noise is a sparse signal. Finding a low-rank approximation matrix subject to an l_1 misfit constraint is a non-convex optimization problem (Wright et al., 2009). A practical algorithm can be developed by simply replacing the non-convex optimization problem by a convex one where a gradient based optimization method is used. The resulting algorithm is named robust principal component analysis (RPCA) (Candès et al., 2011). We will show that that RPCA is capable of suppressing erratic noise. We tested the proposed algorithm with synthetic examples to de-noise erratic noise present in seismic records.

ROBUST PRINCIPAL COMPONENT ANALYSIS

We assume that the erratic noise can be represented using a sparse matrix $\mathbf{S}$. Only a few entries of $\mathbf{S}$ are non-zero elements and can be arbitrarily large in amplitude (Zhou et al., 2010). Robust principal component analysis suggests the following optimization problem:

$$\begin{aligned} &\min \ \|\mathbf{S}\|_0 \\ &s.t. \quad \text{rank}(\mathbf{L}) = k, \quad \mathbf{D} = \mathbf{L} + \mathbf{S}, \end{aligned} \tag{3}$$

where $||\mathbf{S}||_0$ denotes the ℓ_0 norm of $\mathbf{S}$, which means the number of non-zero elements in $\mathbf{S}$. Equation (3) is a NP-hard problem. To make the problem tractable, we use the ℓ_1 norm, which is defined by the summation of absolute values of the elements of the matrix $\mathbf{S}$, to replace the ℓ_0 norm. In the meantime, we consider to replace the the low-rank constraint by the nuclear norm of $\mathbf{L}$, which is defined as the sum of all singular values of the matrix $\mathbf{L}$. One can show that the ℓ_1 norm is the tightest convex relaxation of the ℓ_0 norm (Donoho, 2006). Similarly, the nuclear norm is the tightest convex relaxation to the low-rank constraint (Fazel, 2002). We also introduce a *Frobenius* norm constraint, $||\mathbf{D}-\mathbf{L}-\mathbf{S}||_F^2$, to tolerate the inclusion of Gaussian noise. The resulting cost function can be written as follows

$$\min J = \frac{1}{2\mu}||\mathbf{D}-\mathbf{L}-\mathbf{S}||_F^2 + \lambda||\mathbf{S}||_1 + ||\mathbf{L}||_*. \tag{4}$$

DOI http://dx.doi.org/10.1190/segam2015-5869427.1

λ is a trade-off parameter that balances the sparsity and low rank constraints. The scalar μ is a small constant that controls the inclusion of Gaussian noise.

We consider to minimize Equation (4) via an iterative scheme to estimate the low-rank data $\mathbf{L}$ as well as the sparse erratic noise $\mathbf{S}$. We split the cost function into two sub-problems based on the sub-gradient method. The solution of Equation (4) is equivalent to the solution of the following system of equations

$$\min J_S = \lambda||\mathbf{S}||_1 + ||\mathbf{S} - \hat{\mathbf{S}}^k||_F^2 \quad (5a)$$

$$\min J_L = ||\mathbf{L}||_* + ||\mathbf{L} - \hat{\mathbf{L}}^k||_F^2 \quad (5b)$$

only if $\hat{\mathbf{L}}^k$ and $\hat{\mathbf{S}}^k$ converge to the solution of

$$\min J_0 = \frac{1}{2\mu}\|\mathbf{D} - \mathbf{L} - \mathbf{S}\|_F^2 . \quad (6)$$

Therefore, we can calculate $\hat{\mathbf{L}}^k$ and $\hat{\mathbf{S}}^k$ by updating a current estimation in the opposite direction of the gradient of J_0

$$\begin{aligned} \hat{\mathbf{S}}^k &= \mathbf{S}^k - \frac{1}{2\mu}(\mathbf{L}^k + \mathbf{S}^k - \mathbf{D}) \\ \hat{\mathbf{L}}^k &= \mathbf{L}^k - \frac{1}{2\mu}(\mathbf{L}^k + \mathbf{S}^k - \mathbf{D}) . \end{aligned} \quad (7)$$

Equation (5a) are commonly seen in the field of compressive sensing. It leads to a soft-thresholding step to all the entries of the updated solution $\hat{\mathbf{S}}^k$ (Beck and Teboulle, 2009). The solution to Equation (5b) can also be found in recent developments of matrix completion. In this case, instead of applying soft-thresholding directly to the entries, a soft-thresholding step are performed to the singular values of the matrix $\hat{\mathbf{L}}^k$ (Cai et al., 2010; Recht et al., 2010) (see algorithm (1)).

Algorithm 1 RPCA

Inputs:

Observed data $\mathbf{D}$, trade-off parameter λ and stopping criterion ε

Initialize:

$\mathbf{L}^0 = 0$; $\mathbf{S}^0 = 0$; $k = 1$

repeat

$\hat{\mathbf{S}}^k = \mathbf{S}^k - \frac{1}{2\mu}(\mathbf{L}^k + \mathbf{S}^k - \mathbf{D})$

$\hat{\mathbf{L}}^k = \mathbf{L}^k - \frac{1}{2\mu}(\mathbf{L}^k + \mathbf{S}^k - \mathbf{D})$

$\mathbf{S}^{k+1} = \max(|\hat{S}_k(i,j)| - \frac{\lambda\mu}{2}, 0)$

$[\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \text{svd}[\hat{\mathbf{L}}^k]$

$\hat{\mathbf{\Sigma}} = \max(|\Sigma(i,i)| - \frac{\lambda\mu}{2}, 0)$

$\mathbf{L}^{k+1} = \mathbf{U}\hat{\mathbf{\Sigma}}\mathbf{V}^*$

$k = k+1$

until $\|\mathbf{D} - \mathbf{L}^k - \mathbf{S}^k\|_F^2 < \varepsilon$

Outputs:

Low-rank estimation $\mathbf{L}^{k+1}$;
Estimated erratic noise $\mathbf{S}^{k+1}$.

The RPCA method is summarized in Algorithm (1). In each iteration, we modify the current estimate of the low-rank data and erratic noise in the opposite direction to the gradient of the quadratic term. Then, we apply two steps of soft-thresholding to the modified estimators. Zhou et al. (2010) have proved that the selection of $\lambda = 1/\max(m,n)$ can guarantee high quality recovery of the matrix $\mathbf{L}$, where m and n are the size of the data matrix $\mathbf{D}$. The tuning parameter μ can be chosen according to $\mu = 0.1\sqrt{\max(m,n) + 8\sigma\sqrt{\max(m,n)}}$ (Tao and Yuan, 2011) where σ is an estimator of the standard error of the additive noise in the data.

RPCA SEISMIC DATA NOISE ATTENUATION

The Robust Principal Component Analysis can be adopted to suppress erratic noise present in the seismic data. We consider to apply the RPCA algorithm to a 3D volume $\mathcal{D}(t,x,y)$ extracted from a multi-dimensional seismic data set. We remind the reader that the Fourier transform that maps data from $t-x-y$ domain to $f-x-y$ domain can be expressed by

$$\tilde{\mathcal{D}}(\omega,x,y) = \int \mathcal{D}(t,x,y)e^{i\omega t}\,d\omega . \quad (8)$$

At a given frequency ω, the spatial data can be denoted via a spectral matrix $\mathbf{D}_\omega$. The desired signal is coherent along the two spatial directions. Therefore, $\mathbf{D}_\omega$ is a low-rank matrix (Trickett, 2003; Cheng and Sacchi, 2014). The de-noising method, which applies robust matrix rank reduction to each frequency slice of the $f-x-y$ data cube, is shown in Algorithm (2).

Algorithm 2 RPCA de-noising

Inputs:

3D seismic volume $\mathcal{D}$
trade-off parameter λ

$\tilde{\mathcal{D}}(\omega,x,y) \leftarrow \mathcal{D}(t,x,y)$ (FFT)

for $\omega = \omega_{min} : \omega_{max}$ **do**

$\mathbf{D}_\omega \leftarrow \tilde{\mathcal{D}}(\omega,:,:)$

$[\mathbf{L},\mathbf{S}] = \text{RPCA}[\mathbf{D}_\omega]$ (Algorithm 1)

$\hat{\tilde{\mathcal{D}}}(\omega,:,:) \leftarrow \mathbf{L}$

end for

$\hat{\mathcal{D}}(t,x,y) \leftarrow \hat{\tilde{\mathcal{D}}}(\omega,x,y)$ (IFFT)

EXAMPLES

Erratic noise elimination

A synthetic data set is utilized to test the robustness of the proposed de-noising method. Figure 1 (a) shows the 3D $t-x-y$ data set, which has 30×30 traces and a total time of 1 second. As is shown in Figure 1 (b), we added Gaussian noise with signal-to-noise ratio equals to 3. We also added isolated traces with erratic noise with an amplitude that is about 3 times of the maximum amplitude of desired signal. The processing

DOI http://dx.doi.org/10.1190/segam2015-5869427.1

frequency band ranges from 1 to 40 Hz. The results of the $f-x-y$ eigen-image filtering and the RPCA de-noising were compared. We evaluate the performance of the algorithm via a quality factor

$$Q = \frac{\|\hat{\mathcal{D}} - \mathcal{D}\|_F^2}{\|\mathcal{D}\|_F^2},$$

where $\mathcal{D}$ is the noise-free data and $\hat{\mathcal{D}}$ denotes the de-noised data. Figure 1 (c) shows the result of $f-x-y$ eigen-image filtering. The rank was set to 3. Figure 1 (d) shows the difference between the filtered data and the true data.The erratic noise was not properly removed and the estimated data shows noticeable artifacts. Figure 1 (e) is the result of the RPCA de-noising. Both the Gaussian noise and the erratic noise are successfully suppressed. The proposed method effectively suppressed the incoherent noise. Figure 1 (f) shows the error panel corresponding to RPCA de-nosing. We improve the quality of the data from $Q = -6.7$ dB to $Q = 11.9$ dB.

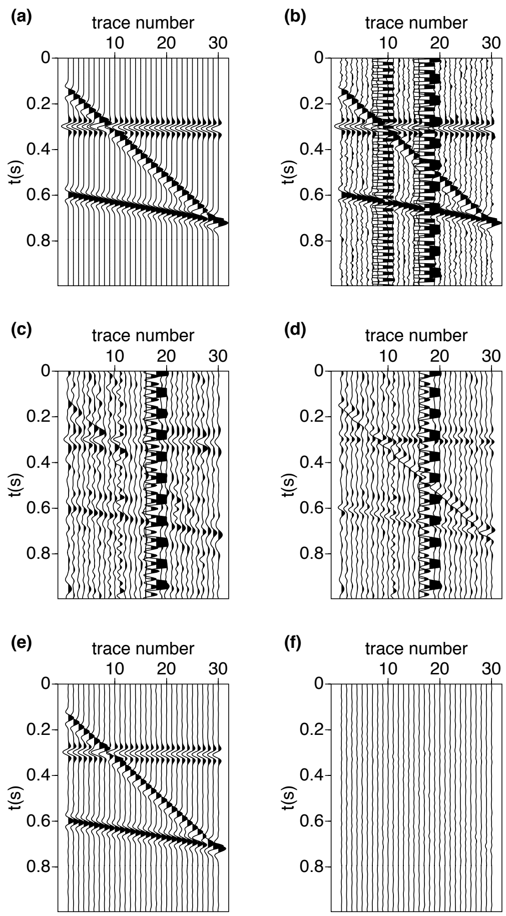

Figure 1: Results of Robust de-noising for the synthetic data set. (a) The ideal CMP gather of the 3D data cube. (b) The CMP gather contaminated with Gaussian and erratic noise. (c) De-noised gather with $f-x-y$ eigen-image filtering. (d) Differences between (a) and (c). (e) CMP gather after RPCA De-noising. (f) Difference between (a) and (e).

We tested the RPCA de-noising method with a real data set. Figure (2) shows a small patch (50×50 traces) from an offset-midpoint gather of a post-stack seismic volume. The total time length of the patch is 1 second. We added erratic noise to isolated traces with cosine signals that mimic power line noise. Figure (2) shows that the RPCA de-noising method has effectively suppressed the erratic noise while preserving the desired signal. In this example, the quality factor of data was improved from $Q-4.7$ dB to $Q = 9$ dB.

Simultaneous source noise suppression

We also applied the proposed algorithm to solve the deblending problem in common receiver domain. We adopted a synthetic 3D vertical seismic profile to mimic a simultaneous source acquisition. The data set contains 205 source lines with 205 source positions on each line. The interval of each source position is 16.67m and the line spacing is also 16.67m (O'Brien, 2010). One common receiver gather has been extracted at depth 1600m. We assume that the data are blended with the self-simultaneous shooting technique. Only one vessel fires with small random time delays generated via a uniform distribution. The expected time delays are 70% of one conventional shot record length. Figure 3 (b) shows the center shot line of the pseudo-deblended common gather which contains the blending noise. In this example, we assume that 30 % of the acquisition time is saved via the proposed acquisition design. Figure 3 (c) shows the de-noising result with the proposed method. The RPCA algorithm is able to remove the interferences and yields a value $Q = 10.2$. This example portrays a deblending technique based on robust de-noising (Huo et al., 2009; Ibrahim and Sacchi, 2014). However, it is important to mention that deblending methods that are based on rank constrained inversion could also be adopted (Cheng and Sacchi, 2014).

CONCLUSIONS

We presented a robust principal component analysis method for suppressing erratic noise that is often present seismic data. We assume the ideal data can be represented via low-rank matrices in the frequency-space domain and that the erratic noise can be represented via a sparse matrix. A nuclear norm constraint as well as an l_1 norm constraint are used to simultaneously recover the data and the erratic noise. We tackled the problem via first order gradient method. Through tests with synthetic examples, the proposed RPCA de-noising method is shown to be able to remove both Gaussian and erratic noise. The method has the potential to be adapted for simultaneous source seismic data de-noising and reconstruction. Furthermore, higher dimensional version of the algorithm could be developed by interchanging matrices by multilinear arrays (tensors) to represent multi-dimensional spatial data at a given temporal frequency.

ACKNOWLEDGMENTS

The authors would like to thank the sponsors of Signal Analysis and Imaging Group (SAIG) at the University of Alberta.

DOI http://dx.doi.org/10.1190/segam2015-5869427.1

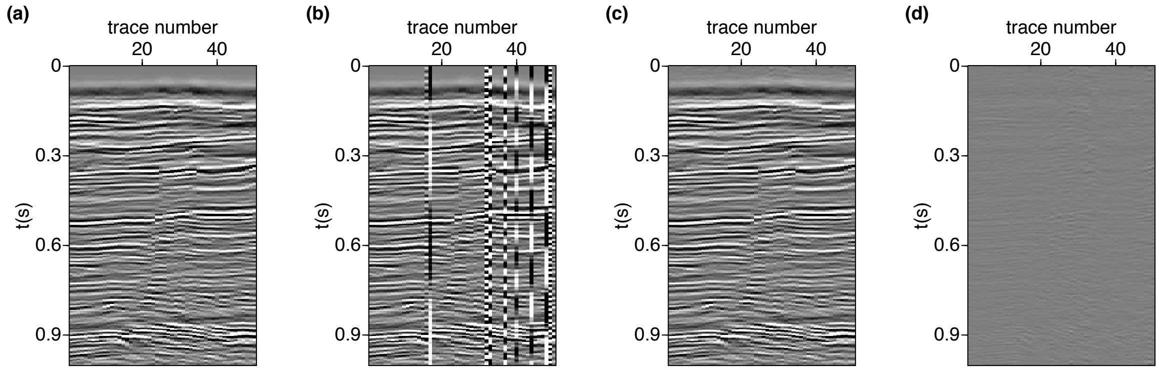

Figure 2: Results of Robust de-noising for a real 3D data set (post-stack). (a) An ideal CMP gather of the 3D data cube. (b) The CMP gather contaminated with erratic noise. (c) CMP gather after RPCA de-noising. (d) Differences between (a) and (c).

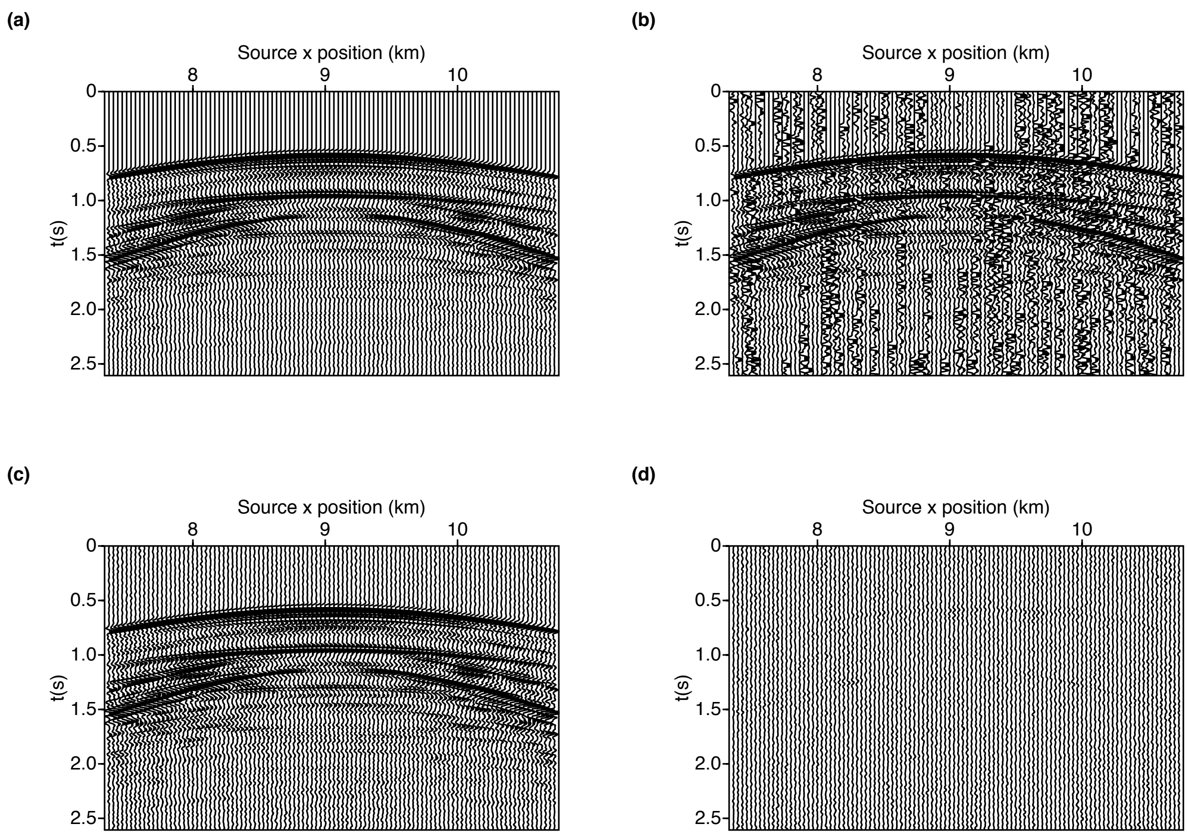

Figure 3: (a) The center shot line of the ideal unblended common receiver gather. (b) The center shot line of a pseudo-deblended common receiver gather, in this example, we save 30 % of acquisition time. (c) The center shot line of the deblended common receiver gather after RPCA de-noising. (d) The difference between section (a) and (c).

DOI http://dx.doi.org/10.1190/segam2015-5869427.1

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2015 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Al-Yahya, K., 1991, Application of the partial Karhunen-Loève transform to suppress random noise in seismic sections: Geophysical Prospecting, **39**, no. 1, 77–93, http://dx.doi.org/10.1111/j.1365-2478.1991.tb00302.x.

Beck, A., and M. Teboulle, 2009, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: SIAM Journal on Imaging Sciences, **2**, no. 1, 183–202, http://dx.doi.org/10.1137/080716542.

Cai, J.-F., E. J. Candès, and Z. Shen, 2010, A singular value thresholding algorithm for matrix completion: SIAM Journal on Optimization, **20**, no. 4, 1956–1982, http://dx.doi.org/10.1137/080738970.

Candès, E. J., X. Li, Y. Ma, and J. Wright, 2011, Robust principal component analysis?: Journal of the ACM (Association for Computing Machinery), **58**, no. 3, Article 11, doi:10.1145/1970392.1970395.

Chen, K., and M. D. Sacchi, 2015, Robust reduced-rank filtering for erratic seismic noise attenuation: Geophysics, **80**, no. 1, V1–V11, http://dx.doi.org/10.1190/geo2014-0116.1.

Cheng, J., and M. D. Sacchi, 2014, A dual domain algorithm for minimum nuclear norm deblending: 84th Annual International Meeting, SEG, Expanded Abstracts, 99–104.

Donoho, D., 2006, For most large underdetermined systems of linear equations the minimum l1-norm solution is also the sparsest solution: Communications on Pure and Applied Mathematics, **59**, no. 6, 797–829, http://dx.doi.org/10.1002/cpa.20132.

Fazel, M., 2002, Matrix rank minimization with applications: Master's thesis, Stanford University.

Golub, G. H., and C. F. van Loan, 1996, Matrix computations, 3rd ed.: The Johns Hopkins University Press.

Huo, S., Y. Luo, and P. Kelamis, 2009, Simultaneous sources separation via multi-directional vector-median filter: 79th Annual International Meeting, SEG, Expanded Abstracts, 31–35.

Ibrahim, A., and M. D. Sacchi, 2014, Simultaneous source separation using a robust Radon transform: Geophysics, **79**, no. 1, V1–V11, http://dx.doi.org/10.1190/geo2013-0168.1.

Jolliffe, I. T., 2010, Principal component analysis, 2nd ed.: Springer.

O'Brien, M. J., 2010, Mounting an offense against poor-quality shear data: 80th Annual International Meeting, SEG, Expanded Abstracts, 177–181.

Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, no. 3, V25–V32, http://dx.doi.org/10.1190/1.3552706.

Recht, B., M. Fazel, and P. Parrilo, 2010, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization: SIAM Review, **52**, no. 3, 471–501, http://dx.doi.org/10.1137/070697835.

Sacchi, M. D., 2009, FX singular spectrum analysis: Presented at the CSPG/CSEG/CWLS Convention, 392–395.

Tao, M., and X. Yuan, 2011, Recovering low-rank and sparse components of matrices from incomplete and noisy observations: SIAM Journal on Optimization, **21**, no. 1, 57–81, http://dx.doi.org/10.1137/100781894.

Trickett, S. R., 2003, *F-xy* eigenimage noise suppression: Geophysics, **68**, 751–759, http://dx.doi.org/10.1190/1.1567245.

Trickett, S. R., and L. Burroughs, 2009, Prestack rank-reducing noise suppression: 79th Annual International Meeting, SEG, Expanded Abstracts, 3332–3336.

Trickett, S., L. Burroughs, and A. Milton, 2012, Robust rank-reduction filtering for erratic noise: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi:10.1190/segam2012-0129.1.

Ulrych, T. J., M. D. Sacchi, and S. L. M. Freire, 1999, Eigenimage processing of seismic sections, *in* R. L. Kirlin and W. J. Done, eds., Covariance analysis for seismic signal processing: SEG, 241–274.

Wright, J., A. Ganesh, S. Rao, Y. Peng, and Y. Ma, 2009, Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization: Proceedings of the 23rd Annual Conference on Advances in Neural Information Processing Systems (NIPS), **22**, 2080–2088, http://papers.nips.cc/paper/3704-robust-principal-component-analysis-exact-recovery-of-corrupted-low-rank-matrices-via-convex-optimization.

Zhou, Z., X. Li, J. Wright, E. Candès, and Y. Ma, 2010, Stable principal component pursuit: Proceedings of the International Symposium on Information Theory (ISIT), IEEE, 1518–1522.

DOI http://dx.doi.org/10.1190/segam2015-5869427.1