

Robust Singular Spectrum Analysis for Erratic Noise Attenuation

Ke Chen*, University of Alberta, Edmonton, Canada

ke7@ualberta.ca

and

Mauricio D. Sacchi, University of Alberta, Edmonton, Canada

msacchi@ualberta.ca

Summary

The Singular Spectrum Analysis (SSA) method, also known as Cadzow filtering, adopts the truncated singular value decomposition (TSVD) or fast approximations to TSVD for rank-reduction. SSA is efficient for attenuating Gaussian noise but it cannot eliminate erratic noise (non-Gaussian). We propose a robust SSA method for simultaneously removing Gaussian and non-Gaussian noise. A robust low rank approximation is used in the newly proposed method. Iteratively reweighted least squares (IRLS) is adopted to estimate the approximated robust rank reduction that is required by the SSA method. Synthetic and real data examples are used to illustrate the performance of the proposed method.

Introduction

Recently, several reduced-rank filtering techniques have been developed for random seismic noise suppression, e.g., f - xy eigenimage analysis (Trickett, 2003) and singular spectrum analysis method (Cadzow filtering) (Sacchi, 2009; Trickett, 2008). Rank-reduction methods have also been developed for simultaneous data completion and random noise attenuation (Oropeza and Sacchi, 2011; Trickett, 2010; Kreimer and Sacchi, 2012; Gao et al., 2013). These rank-reduction methods have two main advantages: first, they are easy and natural to be applied on multidimensional data; second, they preserve the signal. In the SSA method, the seismic data consisting of a superposition of plane waves is transformed to the frequency-space domain. SSA embeds each frequency slice into a Hankel matrix. The rank of this matrix should be equal to the number of distinct dips in the data. Additive incoherent noise in the data will increase the rank of the Hankel matrix. Thus, the denoising problem is posed as a matrix rank-reduction problem. Then, the anti-diagonal elements of the rank-reduced matrix are averaged to recover the signal in frequency domain. In general the TSVD or fast approximations to TSVD are applied in the SSA method. However, the TSVD approximates a matrix by one of a lower rank in a least squares sense. The latter leads to a suboptimal performance of the SSA method when the data are contaminated with erratic (non-Gaussian) noise. Erratic noise is often contained in seismic data in the form of noise bursts, incoherent signals arising from improper geophone coupling and source generated noise. Trickett (2012) proposed a robust rank-reduction filtering method by iteratively applying Cadzow filtering on the reweighted combination of observed and reconstructed data. In this abstract, we use an M-estimator (Huber, 1981) to compute the reduced rank approximation of the noisy Hankel matrix.

Theory

The SSA method (Trickett, 2008; Sacchi, 2009) can be summarized as follows:

$$\hat{\mathbf{D}}_{\omega} = \mathcal{A}R\mathcal{H}\mathbf{D}_{\omega}, \quad (1)$$

where \mathbf{D}_ω is the data slice at frequency ω , \mathcal{H} is the Hankel operator, \mathcal{R} is the TVSD operator and \mathcal{A} is the anti-diagonal averaging operator. Instead of TSVD, we use a robust low rank approximation (Torre and Black, 2003) for the operator \mathcal{R} . The embedded Hankel matrix \mathbf{M} is approximated by the product of two lower dimensional factor matrices \mathbf{U} and \mathbf{V} . The bisquare function (Beaton and Tukey, 1974) is used as error criterion. The cost function is:

$$E(\mathbf{U}, \mathbf{V}) = \|\mathbf{M} - \mathbf{U}\mathbf{V}^H\|_\rho = \sigma^2 \sum_{i=1}^m \sum_{j=1}^n \rho \left(\frac{m_{ij} - \sum_{q=1}^K u_{iq} v_{jq}^*}{\sigma} \right), \quad (2)$$

where σ is a scale parameter to make the estimate scale equivalent, K is the rank of the reconstructed matrix. In general, there is no closed-form solution to this problem, iteratively reweighted least squares (IRLS) is used to find the approximate optimal solution. The weighted Frobenius norm low rank approximation problem is to find the low rank matrix by minimize the following cost function:

$$E_W(\mathbf{U}, \mathbf{V}) = \|\mathbf{W}^{\frac{1}{2}} \odot (\mathbf{M} - \mathbf{U}\mathbf{V}^H)\|_F^2 = \sum_{i=1}^m \sum_{j=1}^n w_{ij} |m_{ij} - \sum_{q=1}^K u_{iq} v_{jq}^*|^2, \quad (3)$$

where \mathbf{W} is the weighting matrix calculated by the residuals of previous IRLS iteration. \odot is the Hadamard product (element-wise). $\frac{1}{2}$ on the upper right corner of matrix denotes the element-wise square root operator. The complex domain bisquare function and its weighting function are given as follows:

$$\rho(u) = \begin{cases} \frac{1}{6}\alpha^2 \left\{ 1 - \left[1 - \left(\frac{|u|}{\alpha} \right)^2 \right]^3 \right\} & |u| \leq \alpha \\ \frac{1}{6}\alpha^2 & |u| > \alpha \end{cases}, \quad w(u) = \begin{cases} \left[1 - \left(\frac{|u|}{\alpha} \right)^2 \right]^2 & |u| \leq \alpha \\ 0 & |u| > \alpha \end{cases}, \quad (4)$$

where α is the tuning constant. We chose $\alpha = 4.685$ to get 95% asymptotic efficiency at standard normal distribution (Holland and Welsch, 1977). The scale parameter σ is chosen to be normalized MAD: $\sigma = 1.4826 \text{ MAD} = 1.4826 \text{ med } |\mathbf{r} - \text{med } |\mathbf{r}||$, \mathbf{r} is the residual vector got by reshaping the residual matrix. Actually, $\alpha\sigma$ is the threshold to distinct outliers and inliers. In each IRLS iteration, problem (2) is reduced to a weighted Frobenius norm low rank approximation problem (equation (3)). Problem (3) can be solved by alternating minimization, which is also named criss-cross regression (Gabriel and Zamir, 1979). The method alternately optimizes one of the two matrices \mathbf{U} and \mathbf{V} with the other one fixed. In this way, the bilinear problem is reduced to alternately solved linear least squares problems. The alternating update of \mathbf{U} and \mathbf{V} is got by solving the following *weighted* normal equations:

$$\text{for } i = 1, 2, \dots, m \quad \mathbf{V}^H \mathbf{W}_i \mathbf{V} \mathbf{u}_i = \mathbf{V}^H \mathbf{W}_i \mathbf{m}_i, \quad (5a)$$

$$\text{for } j = 1, 2, \dots, n \quad \mathbf{U}^H \mathbf{W}^j \mathbf{U} \mathbf{v}_j = \mathbf{U}^H \mathbf{W}^j \mathbf{m}^j, \quad (5b)$$

where \mathbf{m}_i is the conjugate transpose of the i th row of \mathbf{M} , \mathbf{m}^j is the j th column of \mathbf{M} . Similar denotation is used for matrices \mathbf{U} and \mathbf{V} . \mathbf{W}_i is a diagonal matrix with i th row of \mathbf{W} on the diagonal. Here, we use Conjugate Gradient least squares (CGLS) (Paige and Saunders, 1982) to solve the weighted normal equations. The robust low rank approximation algorithm is summarized as follows:

- (1) Initialize models \mathbf{U} and \mathbf{V} .
- (2) Calculate weighting matrix \mathbf{W} using weighting function in equation (4).
- (3) Update factor matrix \mathbf{U} by solving normal equations (5a).
- (4) Update factor matrix \mathbf{V} by solving normal equations (5b).
- (5) Iterate steps (3)-(4) until convergence or achieve a maximum iteration number.
- (6) Iterate steps (2)-(5) until convergence or achieve a maximum iteration number.

Examples

A synthetic example (Figure 1) is used to evaluate the performance of the robust SSA method. Its result is compared with the results of least squares SSA and f - x deconvolution. The data consists of three linear events, erratic noise (with amplitude 2 and 3 times the max amplitude of wavelet) and smoothed Gaussian noise with SNR equals to 1. U and V are started with random matrices and the rank K is selected to be 3. We can see that the results of f - x deconvolution and SSA are not acceptable. However, robust SSA method can resist the erratic noise. Also, it's efficient for attenuating Gaussian noise. Besides, it inherits the merit of the traditional SSA that it preserves the signal.

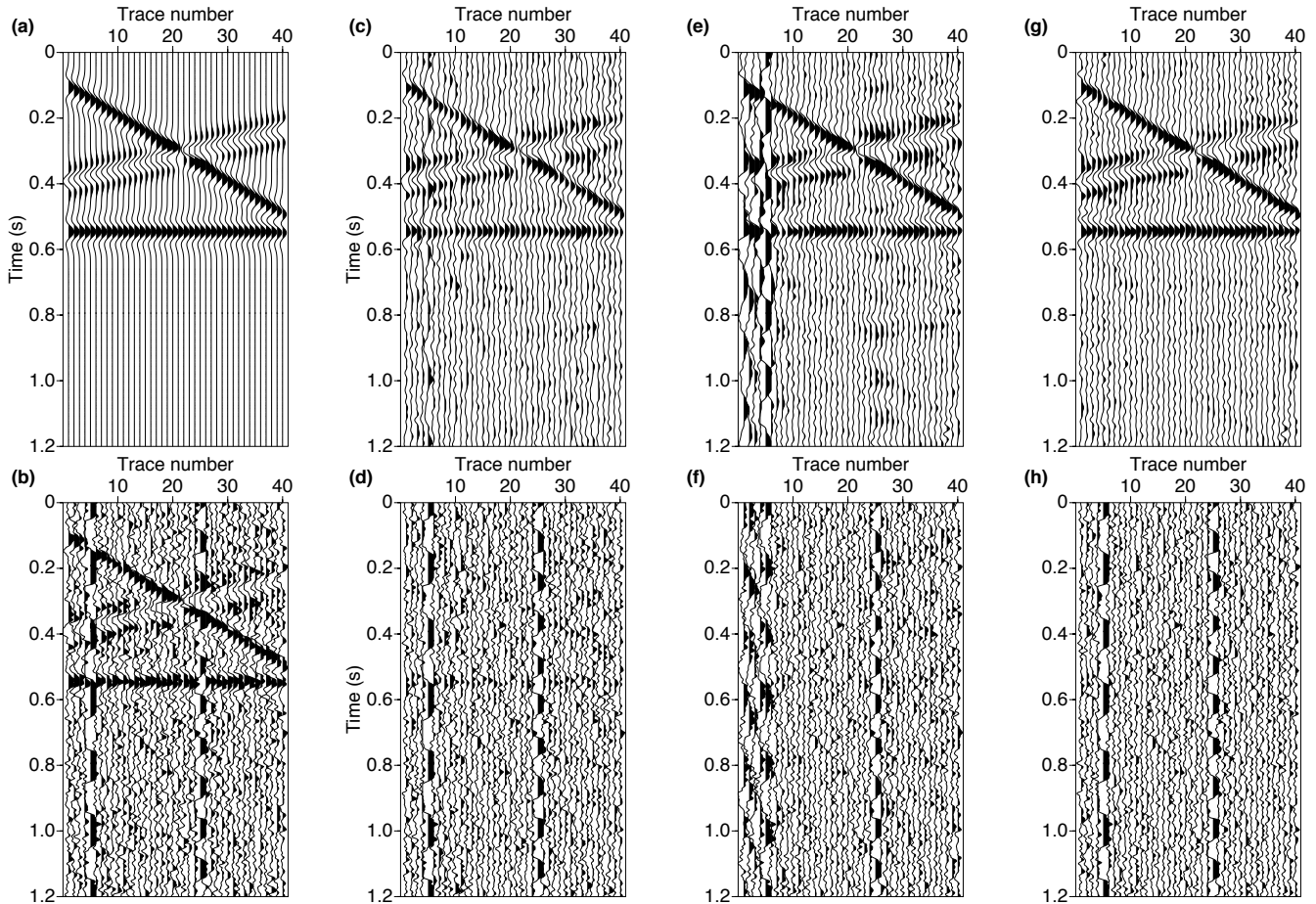


Figure 1: (a) Noise free data. (b) Data corrupted with erratic noise and Gaussian noise (SNR=1), (clipped, clip=1). (c) Result of f - x deconvolution. (d) Difference between (b) and (c). (e) Result of SSA. (f) Difference between (b) and (e). (g) Result of the robust SSA. (h) Difference between (b) and (g).

The newly proposed method is also tested on a real data set. The seismic section contains with 800 traces. The data are divided in overlapping windows. The rank for reconstruction is selected to be 2. Figure 2 (a) shows a part of the data. Figure 2 (b) shows the result of f - x deconvolution. Figure 2 (c) shows the result of least squares SSA. Figure 2 (d) shows the result of robust SSA. We can find that both f - x deconvolution and SSA method have problems in handling the erratic noise. In contrast, robust SSA performs very well in the presence of the erratic noise.

Conclusions

In this abstract, we propose a robust singular spectrum analysis method which can remove Gaussian and non-Gaussian noise. A robust low rank approximation is used in the new method instead of the traditional truncated SVD. The robust low rank approximation is obtained by considering the representation of the original Hankel matrix in the SSA method in terms of a low rank approximation

under the bisquare norm. Synthetic and real data examples demonstrate the efficiency of the new proposed algorithm.

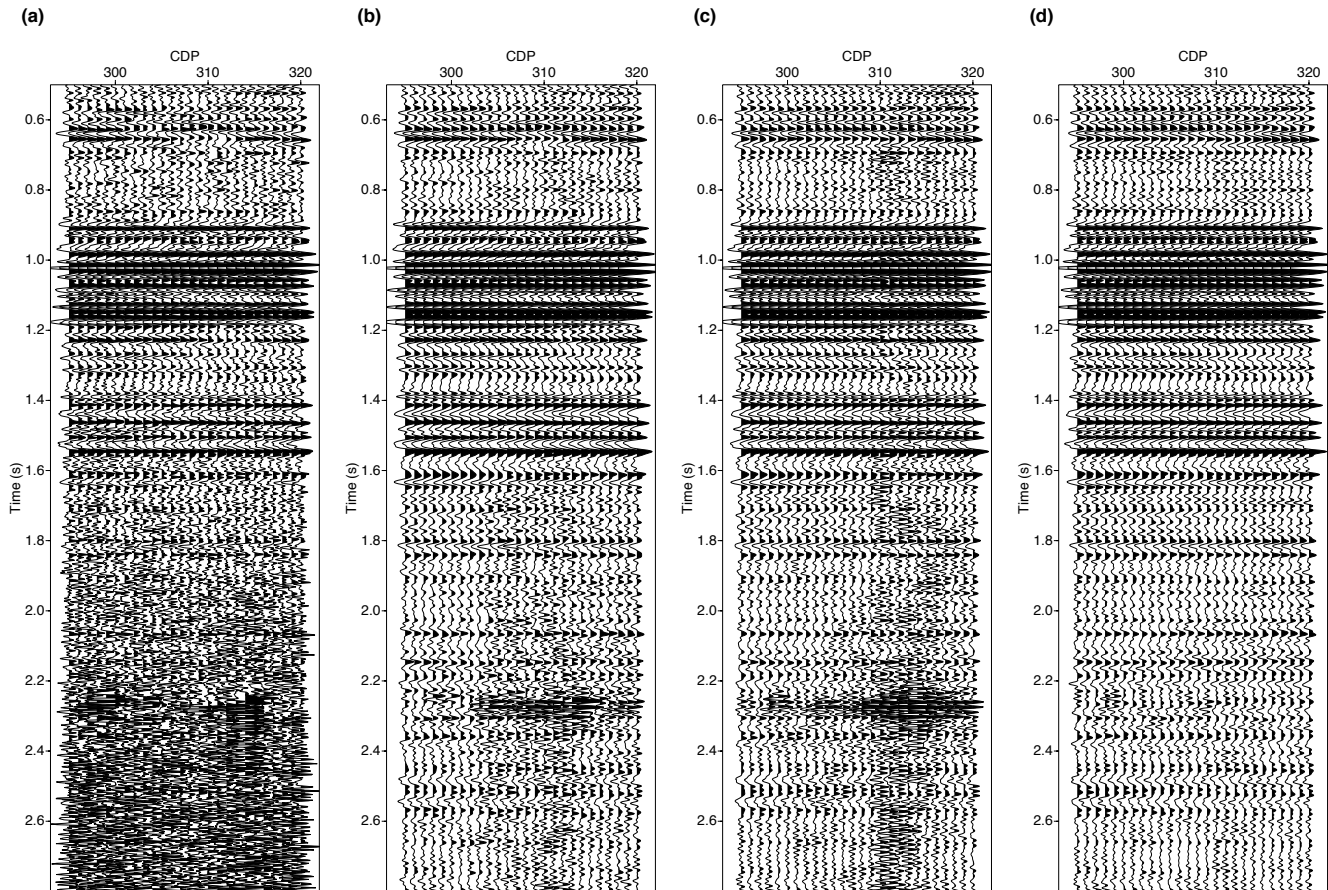


Figure 2: (a) A portion of a real data set contaminated with erratic noise. (b) Result of f - x Deconvolution. (c) Result of SSA. (d) Result of robust SSA.

Acknowledgements

We wish to thank the sponsors of the Signal Analysis and Imaging Group (SAIG) at the Department of Physics of the University of Alberta for supporting this research.

References

- Beaton, A. E., and J. W. Tukey, 1974, The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data: *Technometrics*, 16, pp. 147–185.
- De la Torre, F., and M. J. Black, 2003, A framework for robust subspace learning: *International Journal of Computer Vision*, 54, no. 1, 117–142.
- Gabriel, K. R., and S. Zamir, 1979, Lower rank approximation of matrices by least squares with any choice of weights: *Technometrics*, 21, pp. 489–498.
- Gao, J., M. D. Sacchi, 2013, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: *Geophysics*, 78, V21–V30.
- Holland, P. W., and R. E. Welsch, 1977, Robust regression using iteratively reweighted least-squares: *Communications in Statistics - Theory and Methods*, 6, 813–827.
- Huber, P.J., 1981, *Robust Statistics*. Wiley and Sons.
- Kreimer, N. and M. D. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: *Geophysics*, 77, V113–V122.
- Oropeza, V. and M. D. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multi-channel singular spectrum analysis: *Geophysics*, 76, V25–V32.
- Paige, C. C., and M. A. Saunders, 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: *ACM Trans. Math. Softw.*, 8, 43–71.

- Sacchi, M. D., 2009, FX Singular Spectrum Analysis: CSPG CSEG CWLS CONVENTION, 392-395.
- Trickett, S., 2003, F-xy eigenimage noise suppression: *Geophysics*, 68, 751–759.
- Trickett, S., 2008, F-xy Cadzow noise suppression: SEG Technical Program Expanded Abstracts, 27, 2586–2590.
- Trickett, S., L. Burroughs, A. Milton, L. Walton, and R. Dack, 2010, Rank-reduction-based trace interpolation: 80st Annual International Meeting, SEG, Expanded Abstracts, 1989–1992.
- Trickett, S., L. Burroughs and A. Milton, 2012, Robust rank-reduction filtering for erratic noise : 82st Annual International Meeting, SEG, Expanded Abstracts. 3829-3834.