Considerations for effective rank based noise attenuation

Aaron Stanton and Jeff Deere

Key Seismic Solutions Ltd.

Summary

Incoherent noise attenuation remains a challenging problem in seismic data processing. While many tools are successful at removing noise, this often comes at the expense of some signal. We seek the ideal noise attenuator– one that is effective, practical, and above all safe. Multichannel Singular Spectrum Analysis (MSSA) is a step in this direction. MSSA exploits the spatial redundancy of data by expressing a frequency slice as a structured matrix. This matrix is then projected onto its low rank subspace, thereby removing incoherent energy. The maximum rank of a frequency slice increases as a product of the spatial dimension half-lengths, making the rank constraint more powerful in higher dimensions; although this power comes at a significant increase in computational cost. In this article we address several considerations that make MSSA an effective noise attenuation algorithm. We first introduce the algorithm and outline an efficient matrix-free implementation of MSSA that exploits the redundancy of multi-level Hankel structures. Finally, we discuss the unique ability of MSSA to automatically estimate the signal to noise ratio (SNR) as a function of time, space and frequency and illustrate this ability on a dataset from northeast BC.

Introduction

Recent advances have made rank reduction a powerful and practical approach to attenuate noise in seismic data. Early efforts to attenuate coherent and random noise via rank reduction were applied to data directly in the time-space domain (Jones, 1985; Freire and Ulrych, 1988). Trickett (2002) greatly advanced rank based noise attenuation by considering 2D data in the frequency-space domain. Here a constant frequency slice is assumed to consist of a finite number of plane waves and is therefore of limited rank. This research also introduced the concept of reshaping a 1D frequency slice of seismic data into a 2D Hankel matrix. Reshaping, otherwise known as hankelization, prior to rank reduction, originates from the study of dynamical systems (Broomhead and King, 1986) and has been variously referred to as MSSA, the Caterpillar Method, and Cadzow Filtering. Trickett (2008), and later Oropeza and Sacchi (2011), extended hankelization to the multidimensional case by reshaping a long vector into a multi-level block Hankel matrix. Since the introduction of the MSSA algorithm there has been much research into various aspects of its implementation (see Trickett (2016) for an overview). In this article we address several considerations that make MSSA an effective noise attenuation algorithm.

MSSA

Multichannel Singular Spectrum Analysis (MSSA) is based on a multidimensional plane wave model of the data. By reshaping a 1D frequency slice, $\boldsymbol{x}$, into a Hankel matrix, $\boldsymbol{X}$, we are able to express this matrix as a weighted sum of k rank-1 matrices, where k corresponds to the number of distinct dips in the data. This low rank property is readily extended to the multidimensional case by expressing $\boldsymbol{X}$ as a multi-level block Hankel matrix. When data are corrupted with random noise, imposing a low rank constraint can be used to extract the underlying signal. A straightforward method to impose a low rank constraint is the truncated singular value decomposition. The matrix $\boldsymbol{X}$ is decomposed into its singular values and vectors, and the rank of $\boldsymbol{X}$ is reduced to k by the low rank projection

$$\boldsymbol{X}^{'} = \boldsymbol{U}_k \boldsymbol{U}_k^H \boldsymbol{X}, \tag{1}$$

where the columns of $\boldsymbol{U}_k$ are the first k singular vectors of $\boldsymbol{X}$. As a final step $\boldsymbol{X}^{'}$ must be reshaped into frequency slice $\boldsymbol{x}^{'}$ via anti-diagonal averaging.

Fast memory efficient Lanczos bidiagonalization

To extract the low rank structure of the data we require only a select number of singular values and vectors. Lanczos bidiagonalization is an efficient method to extract these elements. Similar to conjugate gradients, the Lanczos algorithm is a Krylov subspace method that uses forward and adjoint matrix-vector multiplications to iteratively decompose a matrix into a small, real-valued, bidiagonal matrix that can be quickly factorized. Lanczos bidiagonalization is incorporated into the MSSA workflow following algorithm 1.

Algorithm 1: Fast MSSA via Lanczos bidiagonalization

1. hankelize frequency slice: $\boldsymbol{x} \rightarrow \boldsymbol{X}$
2. factorize $\boldsymbol{X} \rightarrow \boldsymbol{L}\boldsymbol{B}\boldsymbol{R}^H$ via Lanczos bidiagonalization
3. factorize $\boldsymbol{B} \rightarrow \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ via the SVD
4. project the matrix onto a rank "k" subspace: $\boldsymbol{U}_k \boldsymbol{U}_k^T \boldsymbol{B} \rightarrow \boldsymbol{B}'$
5. multiply $\boldsymbol{L}\boldsymbol{B}'\boldsymbol{R}^H \rightarrow \boldsymbol{X}'$
6. average along anti-diagonals: $\boldsymbol{X}' \rightarrow \boldsymbol{x}'$

The speed of this algorithm results from the fact that steps 2 to 4 only consider a small number of singular values, making the bidiagonal matrix $\boldsymbol{B}$ small and real valued. The cost of bidiagonalization in step 2 is dominated by matrix-vector multiplication. Because the data are embedded in a Hankel matrix it is possible to exploit this structure to write matrix-vector multiplication as element-to-element multiplication in the Fourier domain (Xu and Qiao, 2008). This method accelerates computation greatly by making use of the Fast Fourier Transform (FFT). Recently, Gao et al. (2013) extended this method to multiple spatial dimensions via block Toeplitz matrices, although this required the use of multidimensional (ND) FFTs which are less efficient for short axes. Recently Lu et al. (2015) introduced an improved version of this method that allows for efficient 1D FFTs to be used. The steps for the multiplication of $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ are outlined in algorithm 2.

Algorithm 2: Fast Hankel matrix-vector multipliciation

1. express a frequency slice as a regularly sampled multidimensional long vector, apply the 1D FFT and store for all future Lanczos iterations. This vector, $\tilde{\boldsymbol{a}}$, represents the information contained in Hankel matrix $\boldsymbol{A}$
2. prior to each matrix-vector multiplication embed the vector $\boldsymbol{x}$ into a zero-padded, structured long vector and apply the 1D FFT, obtaining vector $\tilde{\boldsymbol{x}}$ (Lu et al., 2015). This step ensures that $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ mimics the multiplication of a block-circulant matrix with a vector, allowing for convolution to be expressed as element-to-element multiplication in the Fourier domain
3. compute the element-to-element multiplication $\tilde{\boldsymbol{b}} = \tilde{\boldsymbol{a}} \odot \tilde{\boldsymbol{x}}$
4. take the inverse 1D FFT of $\tilde{\boldsymbol{b}}$ and reverse the embedding applied in step 2 to obtain $\boldsymbol{b}$

Two important features of this algorithm are that it greatly improves the efficiency of the multiplication and it acts directly on the original frequency slice, providing a savings of computation as well as storage. Unfortunately, the final steps of algorithm 1 require building a Hankel matrix as the factored matrices are multiplied and anti-diagonal averaging is applied to recover the denoised frequency slice. Recently Cheng and Sacchi (2016) introduced a fast method to multiply factored matrices from QR decomposition and to perform anti-diagonal averaging in a single step via the ND FFT. We have incorporated this improvement into algorithm 1; adapting it to the Lanczos algorithm, and using the 1D FFT rather than the ND FFT. The result is a fast, matrix-free implementation of MSSA that does not require building, storing or averaging Hankel structures.

Domains and dimensions

The above modifications make the algorithm flexible enough to apply in a variety of domains (including common source gathers, common receiver gathers, post stack volumes, common offset vectors, and cross-spreads); and efficient enough to apply in 5D. While other noise attenuation methods must be applied to well sampled input volumes, MSSA is able to overcome this problem by iteratively compensating for missing input traces via

$$\boldsymbol{x}_j = \boldsymbol{S}\boldsymbol{x}_0 + (\boldsymbol{I}-\boldsymbol{S})\boldsymbol{P}\{\boldsymbol{x}_{j-1}\}, \tag{2}$$

where $\boldsymbol{S}$ is a diagonal sampling operator, $\boldsymbol{x}_0$ is an input frequency slice, and $\boldsymbol{P}\{\boldsymbol{x}_{j-1}\}$ projects a frequency slice from iteration $j-1$ onto its low rank subspace. After N iterations the denoised data is obtained from $\boldsymbol{P}\{\boldsymbol{x}_N\}$.

Automatic discrimination of signal and noise

The amplitude of a singular value in MSSA represents the magnitude of coherency along its dip (given by the singular vector). It is straightforward to visually discriminate singular values associated with signal from those associated with noise; by definition, random noise is incoherent. To ensure signal preservation for all positions, times, and frequencies in a survey it is necessary to automate the discrimination between signal and noise in the MSSA workflow. Gavish and Donoho (2014) provide a simple estimator of the cutoff between signal and noise as $2.858 s_{med}$, where s_{med} is the median singular value. It is more difficult to estimate this cutoff when only a limited number of singular values are known, although knowing the norm of the data as well as the norm of the available singular values allows for a reasonable estimate to be made (Trickett, 2015).

Field data example

We demonstrate the algorithm on a 3D dataset acquired in northeastern BC. The data are contaminated with a mixture of random noise with intermittent regions of high amplitude, high frequency scattered noise. The stacked data are shown in figure 1. The pre-stack data were organized into overlapping subsurface 5D patches with CMP dimensions 25x25, and inline and crossline offset dimensions 5x5, with a time gate of 200ms. Under this geometry, before prestack interpolation has been applied, the input volumes consist of approximately 80% missing traces. Ten iterations of MSSA were performed with the rank automatically determined by the distribution of singular values. Figure 2 shows the results of applying MSSA and stacking the data, with the difference panel shown in figure 3. MSSA has clearly improved the continuity of reflections while preserving weak diffracted energy. The average SNR estimated by the MSSA algorithm for frequencies between 50-120Hz (applied to the input stack data) is shown in figure 4. The algorithm has identified the region to the right to have a SNR less than 1, meaning the energy of the noise is greater than the energy of the signal in this bandwidth. Analyzing the results on prestack data (figure 5), we see an even more obvious improvement to the data quality; several weak events that are difficult to discern on the input CMP are easily identified after MSSA.

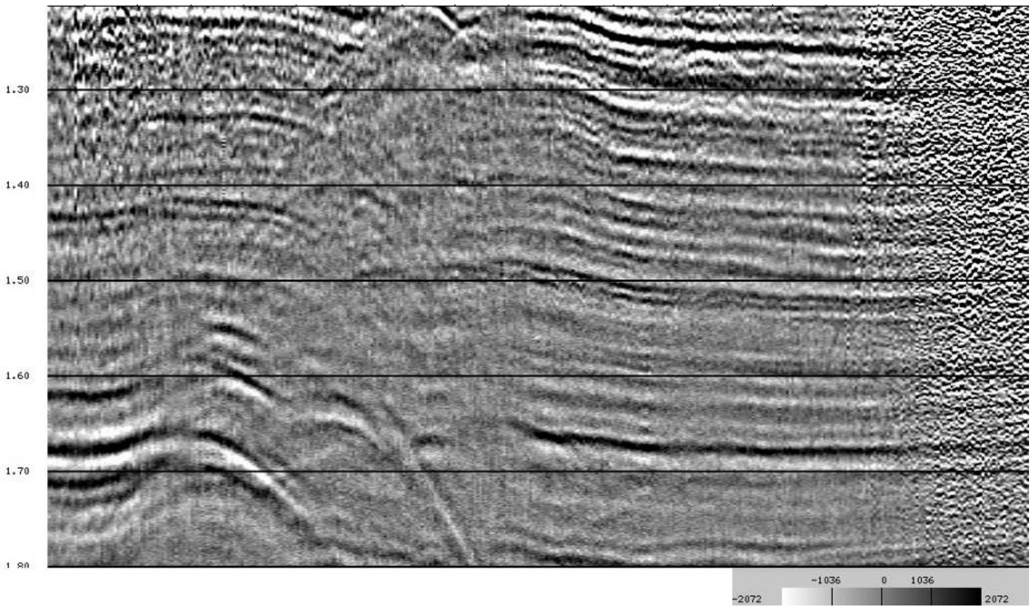

Figure 1: Stacked input data.

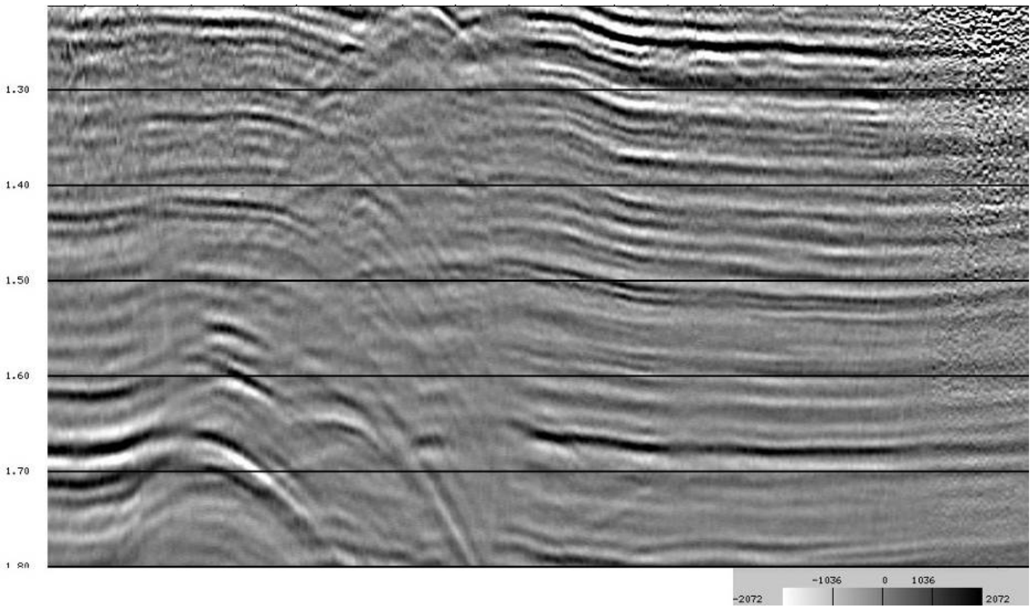

Figure 2: Stacked data after applying 5D MSSA.

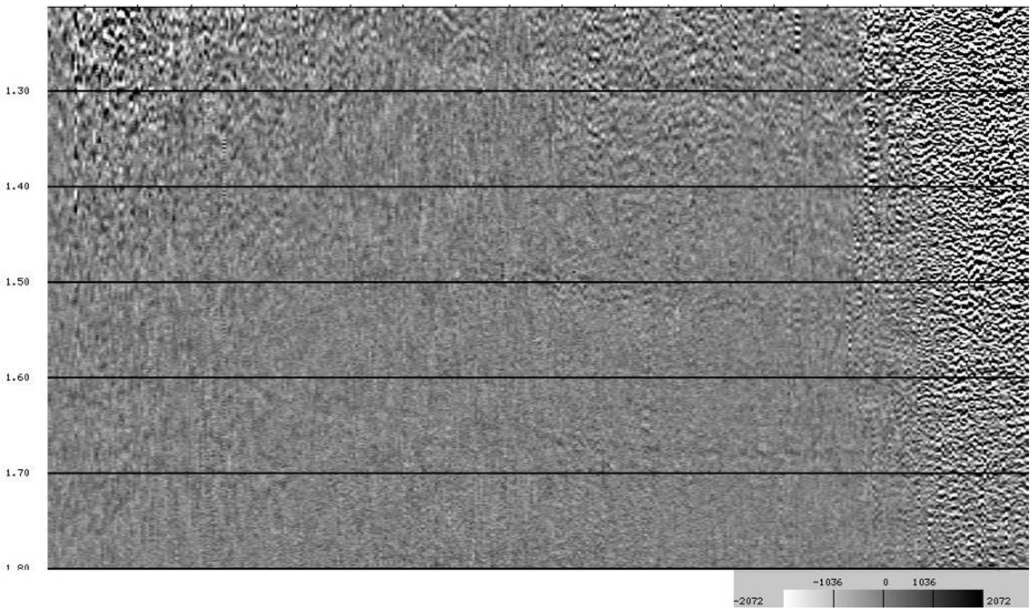

Figure 3: Difference between input and 5D MSSA stacks.

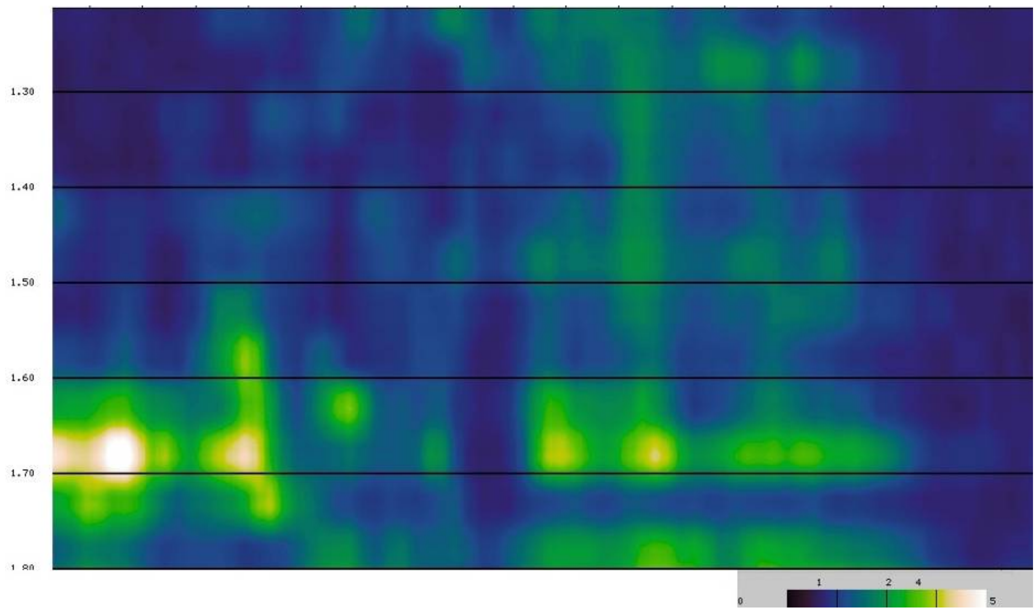

Figure 4: Avergage Signal to Noise Ratio estimated from the distribution of singular values for frequencies between 50-120 Hz, computed using the post-stack data in figure 1.

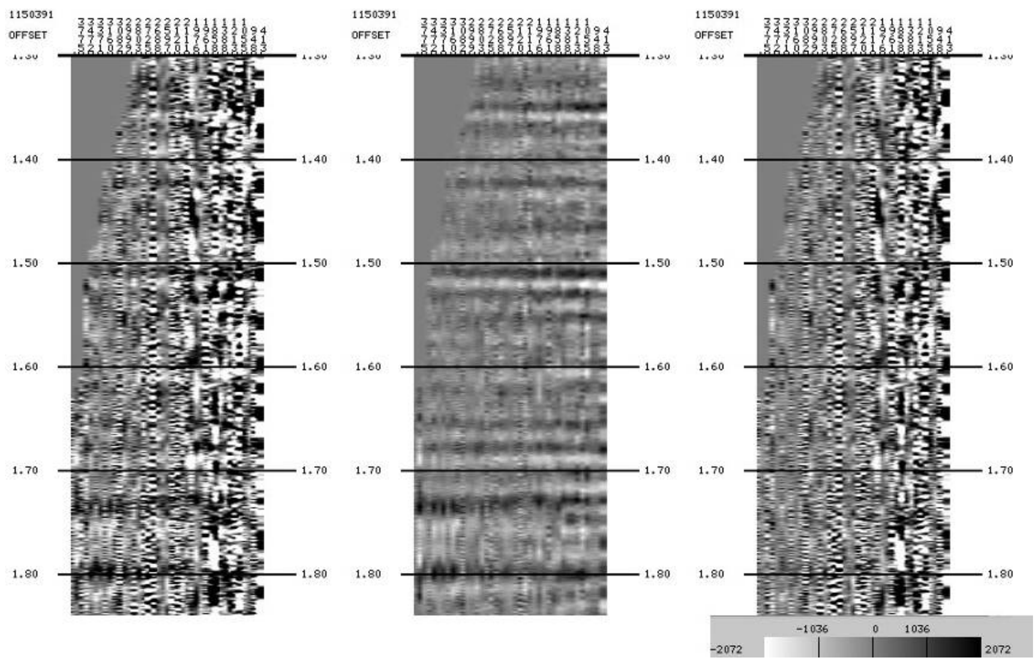

Figure 5: A CMP gather before MSSA (left), after MSSA (middle), and the estimated noise (right).

Conclusions

We have outlined several considerations that make MSSA an effective noise attenuation algorithm. We reviewed two algorithmic efficiencies that make it practical to extend MSSA to higher dimensions. Next we introduced an iterative strategy that allows poorly sampled volumes to be denoised effectively, and finally we reviewed the ability of MSSA to discriminate between signal and noise as a function of time, space and frequency. A field data example from northeast BC demonstrated the ability of MSSA to attenuate noise while preserving weak diffracted energy.

Acknowledgements

We are grateful to Seitel Canada Ltd. for supplying the seismic data used in the example. We also thank our colleagues Richard Bale and Mark Jiang for their advice and technical support and Key Seismic Solutions for permission to publish this work.

References

Broomhead, D. S., and G. P. King, 1986, Extracting qualitative dynamics from experimental data: Physica D: Nonlinear Phenomena, **20**, 217–236.

Cheng, J., and M. Sacchi, 2016, Fast and memory-efficient singular spectrum analysis for seismic data reconstruction and denoising, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists, 4064–4068.

Freire, S. L., and T. J. Ulrych, 1988, Application of singular value decomposition to vertical seismic profiling: Geophysics, **53**, 778–785.

Gao, J., M. D. Sacchi, and X. Chen, 2013, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: Geophysics, **78**, V21–V30.

Gavish, M., and D. L. Donoho, 2014, The optimal hard threshold for singular values is: IEEE Transactions on Information Theory, **60**, 5040–5053.

Jones, I. F., 1985, Applications of the Karhunen-Loéve transform in reflection seismology: PhD thesis, University of British Columbia.

Lu, L., W. Xu, and S. Qiao, 2015, A fast svd for multilevel block hankel matrices with minimal memory storage: Numerical Algorithms, **69**, 875–891.

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

Trickett, S. R., 2002, Fx eigenimage noise suppression: Presented at the 2002 SEG Annual Meeting, Society of Exploration Geophysicists.

———, 2008, F-xy cadzow noise suppression: Presented at the 2008 SEG Annual Meeting, Society of Exploration Geophysicists.

———, 2015, Preserving signal: Automatic rank determination for noise suppression: Presented at the 2015 SEG Annual Meeting, Society of Exploration Geophysicists.

———, 2016, Frequency-domain rank reduction in seismic processing–an overview: Presented at the CSEG conference and exhibition.

Xu, W., and S. Qiao, 2008, A fast symmetric svd algorithm for square hankel matrices: Linear Algebra and its Applications, **428**, 550–563.