PHYS 580 – Computational Physics

Filtering and Interpolation of 2D and 3D Seismic Records

Low Rank Approximation and Reinsertion Algorithm (MSSA)

Gonzalo Rubio

Summary

Dense data with high Signal to noise ratio is fundamental for analyzing seismic data. Data acquisition procedures typically generate and record unwanted signals and are often limited by coverture density. To alleviate this issue, singular spectrum analysis is used as a method to enhance seismic data by reducing the levels of noise in the signal and interpolate seismic traces. This procedure is evaluated on various 2D and 3D sets of synthetic data composed of k, linear events at different levels of signal to noise ratio and sparsity. The quality of the estimated data improves relative to the quality of the input data and an enhanced image is produced. The success of this method is related to the low rank approximation of a Trajectory matrix arising from a predictable signal in frequency-space domain and to the iteratively reinsertion of enhanced data. This procedure is effective and can be applied on a number of spatial dimensions which is beneficial for analyses of non-ideal seismic data.

Introduction

Random or incoherent noise attenuation of seismic records is crucial for improving seismic imaging of local or global scale features, inverting for velocity models, and interpretation for exploration or environmental purposes. More times than not, the signal we wish to analyze is contaminated with noise from many possible sources such as, but not limited to: instrumentation to produce and record the signal, ambient noise, surface waves and multiples. In other words, the recorded signal is a superposition of the signal we want to analyze and unwanted signals or noise.

Seismic data can not only suffer from low signal to noise ratio (SNR) but can also suffer from low coverage. Naturally, one might be interested in creating data sets with a denser coverture to improve imaging and interpretation.

Due to the fundamental nature of these problems, it is no surprise that many methods aiming to alleviate these problems have been developed in the Geophysics community. In the past two decades several methods for denoising seismic records based on the low rank approximation of images have been developed. At the end of the 20th century it was suggested that eigenimage filtering could be a powerful tool for filtering seismic data (Ulrych *et al*, 1999). Years later, eigenimage filtering in frequency-space domain was proposed to enhance seismic records (Trickett, 2003). In a quest to improve the rank reduction based denoising methods, a method which relies on eigenimage filtering and time series arising from dynamical systems (Cadzow, 1998) was proposed to attenuate random noise by exploiting redundancy in space (Trickett, 2008; Sacchi, 2009). In an attempt to simultaneously remove noise and interpolate seismic data, a powerful paper which addressed this issue was published (Oropreza & Sacchi, 2011). The proposed algorithm relies on Cadzow filtering and a reinsertion algorithm like projection onto convex sets (Abma, 2006).

Fast forward to present day, a series of considerable improvements have been developed to deal with higher dimensional data (Jianjun *et al*, 2013), to reduce the computational cost of these algorithms (Browne *et al*, 2008; Xu, 2014).

In this report, emphasis is placed on noise reduction and interpolation of 2D and 3D seismic data by means of Multichannel Singular Spectrum Analysis (MSSA) and a reinsertion algorithm. The discussion begins by addressing the simplest case, considering 2D data only contaminated with noise and continue increasing the difficulty until faced with the problem of reconstruction and denoising of 3D seismic records. A series of quantitative tests evaluating the performance of these methods are explored under varying conditions.

Theory

Noise suppression of digital images (eigenimage filtering)

It has been known that compression and restoration of digital images is one of the many applications of singular value decomposition (Andrews and Patterson, 1976). The SVD of a Matrix can be defined as follows:

$$\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{H} \tag{1}$$

Here, $\boldsymbol{A}$ is a matrix that does not need to be square but for simplicity we will consider it to be of size $n \times n$. As usual, $\boldsymbol{U}$ and $\boldsymbol{V}$ are $n \times n$ unitary matrices whose column vectors are the left and right eigenvectors of $\boldsymbol{A}$ respectively, and Σ is a real diagonal matrix with ordered singular values $\sigma_1 > \sigma_2 > \cdots > \sigma_n > 0$.

It is widely known that one can also represent A as the sum of weighted eigen images:

$$\boldsymbol{A} = \Sigma_{i=1}^{n} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{H} = \Sigma_{i=1}^{n} \boldsymbol{I}_i \tag{2}$$

Since each eigen image is an outer product, each is at most of rank one. One can also define a partial sum of eigen images for $k < n$:

$$\widetilde{\boldsymbol{A}} = \Sigma_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{H} = \Sigma_{i=1}^{k} \boldsymbol{I}_i \tag{3}$$

This is known as the "truncated SVD" or "low rank approximation" of $\boldsymbol{A}$. For digital images, as k approaches n, we obtain a better recovery of the image. However, only a few eigenimages are needed to obtain a reasonably good approximation of the image (Trickett, 2001). This is true since this is the optimal solution in the least squares sense (Hansen, 1987). This method is extremely powerful for noise suppression of digital images since most coherent energy tends to map into the first few eigen images while incoherent energy tends to map onto the remainder.

Noise suppression of 2D seismic records (SSA)

Suppose we have a series of nx equally spaced traces or time series with nt time samples, $d(t,x)$ representing k seismic events arising from some geological structure, in the absence of noise. For instance, a single channel or receiver can be represented as:

$$d\left(t, x_j\right) = [a_{1,j}, a_{2,j}, \dots, a_{nt,j}] \tag{4}$$

Where $a \in \Re$ and it represents the amplitude of the reflectivity at that time sample for that channel. Similarly, all channels at some specific time sample can be represented by:

$$d(t_i, x) = [a_{i,1}, a_{i,2}, \dots, a_{2,nx}] \tag{5}$$

Since the truncated SVD is to be applied on matrices, we must construct a matrix by taking the discrete Fourier transform (DFT) or each seismogram along the time axis. Leaving us with $d(w, x)$, equation (5) now becomes:

$$d(w_i, x) = [c_{i,1}, c_{i,2}, \dots, c_{i,nx}] \tag{6}$$

Where $c \in \Im$ representing the Fourier transform of a. It is possible to prove that there exists a recursive relation between adjacent channels (Sacchi, 2009), meaning that the signal is predictable in space for some constant P.

$$c_{i,n} = Pc_{i,n-1} \tag{7}$$

We can now form a trajectory matrix $\boldsymbol{H}_{w_i}$ composed of lagged vectors (Sacchi, 2009). This matrix is equivalent to the digital image in equation (1). For instance, for 5 adjacent channels at a given frequency we have:

$$\boldsymbol{H}_{w_i} = \begin{matrix} c_1 & c_2 & c_3 \\ c_2 & c_3 & c_4 \\ c_3 & c_4 & c_5 \end{matrix} \tag{8}$$

From the recursive relation in equation (7) we ca rewrite the trajectory matrix as follows:

$$\boldsymbol{H}_{w_i} = \begin{matrix} c_1 & Pc_1 & P^2c_1 \\ c_2 & Pc_2 & P^2c_2 \\ c_3 & Pc_3 & P^2c_3 \end{matrix} \tag{9}$$

The elements are constant along each skew diagonal, giving it a Hankel structure. We can see from equation (9) that in the absence of noise, $rank(\boldsymbol{H}_{w_i}) = k$ since the columns are linear combinations of each other. However, the signal in the presence of random noise implies that $k \leq rank(\boldsymbol{H}_{w_i}) \leq nx$. We can now apply the truncated SVD to such matrix to reduce its rank, hence attenuating any noise that may be present.

$$\widetilde{\boldsymbol{H}_{w_i}} = \Sigma_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^H \tag{10}$$

To obtain an estimate of an ideal observation, that is noise free, we now must average along the antidiagonals of $\widetilde{\boldsymbol{H}}$. In return, for a given frequency we have an estimate of the ideal signal.

$$\tilde{d}(w_i, x) \tag{11}$$

This procedure is repeated for all frequencies in the band of the signal and transformed back to time domain.

Noise suppression of 3D seismic records (MSSA)

Noise attenuation of seismic data that depends on more than one spatial dimension typically outperforms noise attenuation of data that only depends on a single spatial dimension. This is because we can access more data from a more localized area (Trickett, 2008). One of the advantages of extending SSA to more spatial dimensions is that not only do we access more information from a more localized area, but we also increase redundancy of the data by creating Hankel matrices. The steps for this multivariate extension are as discussed.

Suppose we have a grid of nx by ny equally spaced traces sampled nt times, $d(t,x,y)$ contaminated with noise. Then for a given frequency we have the following "slice":

$$d_{w_i} = \begin{matrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{matrix} \tag{12}$$

We can form a block Hankel matrix composed of Hankel matrices:

$$\boldsymbol{M}_{w_i} = \begin{matrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 \\ \boldsymbol{H}_2 & \boldsymbol{H}_3 \end{matrix} \tag{13}$$

Where each block has the same structure as equation (8). Each Hankel matrix will be of size $Crow$ by $Ccol$, where $Crow = int\left(\frac{nx}{2}\right) + 1$ and $Ccol = nx - Crow + 1$. The block Hankel matrix will be of size $Lrow$ by $Lcol$ where $Lrow = int\left(\frac{ny}{2}\right) + 1$ and $Lcol = ny - Lrow + 1$. Like in SSA, we can now find the low rank approximation $\widetilde{\boldsymbol{M}_{w_i}}$and average along the antidiagonals of each block. At this point, it is very convenient to define the MSSA filter operator (Oropreza and Sacchi, 2011).

$$Fssa = F_A F_R F_H \tag{14}$$

Such that:

$$F_{ssa}(\boldsymbol{d}) = F_A\left(F_R\big(F_H(\boldsymbol{d})\big)\right) = \widetilde{\boldsymbol{d}} \tag{15}$$

Where the subscripts indicate averaging, low rank approximation and Hankelization operator respectively. Like in 2D, this process is carried on for all frequencies in the signal and transformed back to time domain for an enhanced data set.

Interpolation of seismic data

As discussed earlier, a reoccurring issue in Geophysics is the lack of complete and dense coverture. Malfunctioning geophones, complex surfaces and the cost associated with deploying dense arrays of receivers makes it almost impossible to avoid interpolation of seismic traces.

We now suppose we have acquired some seismic data somehow. We wish to sort or seismic data by equally spaced traces, and we find out that in the process of doing so, some observations will be missing.

For instance, a time slice of a 3D data set may look as follow:

$$d_{t_i} = \begin{matrix} 0 & 0 & a_{1,3} \\ a_{2,1} & a_{2,2} & 0 \\ a_{3,1} & a_{3,2} & 0 \end{matrix} \tag{16}$$

From our discussion of SSA, we know that the Hankel matrix in equation (9) will be of rank k for an ideal case. Like noise, missing observations also increase the rank of the associated Hankel matrix. This suggest that seismic interpolation can be viewed as a rank reduction problem.

The proposed algorithm is as follows (Oropeza and Sacchi, 2011; Abma and Kabir, 2006):

$$\boldsymbol{d}^0 = \boldsymbol{d}^{obs}$$
$$\boldsymbol{d}^{\boldsymbol{\nu}} = \alpha\, \boldsymbol{d}^{obs} + (\boldsymbol{I} - \alpha F_s) \nu = 1,2,3, \ldots \tag{17}$$

Where $\boldsymbol{S}^{obs}$ is the Hadamard product of a sampling operator F_s and the complete data $\boldsymbol{d}$. The sampling operator is equal to 1 for an existing observation and zero otherwise. $\boldsymbol{I}$ is an operator of the size of the data containing ones and α is a reinsertion parameter which dictates the amount of reinsertion of the current estimate of the data and the original data and it is also associated with the noise level.

The previous discussion can be easily reproduced for data that depends on a single spatial variable.

Assessment of quality of noise removal and data reconstruction

In order to quantify the goodness of data recovery we define:

$$Q = 10 \log \left(\frac{\|\boldsymbol{d}_0\|}{\|\boldsymbol{d}_0 - \widetilde{\boldsymbol{d}}\|} \right) \tag{18}$$

Where $\boldsymbol{d}_0$ represents the ideal data and $\widetilde{\boldsymbol{d}}$ is the estimated data. This metric is useful when comparing different ranks in the denoising process, for comparing different levels of SNR and sparsity of the data.

Examples

2D denoising

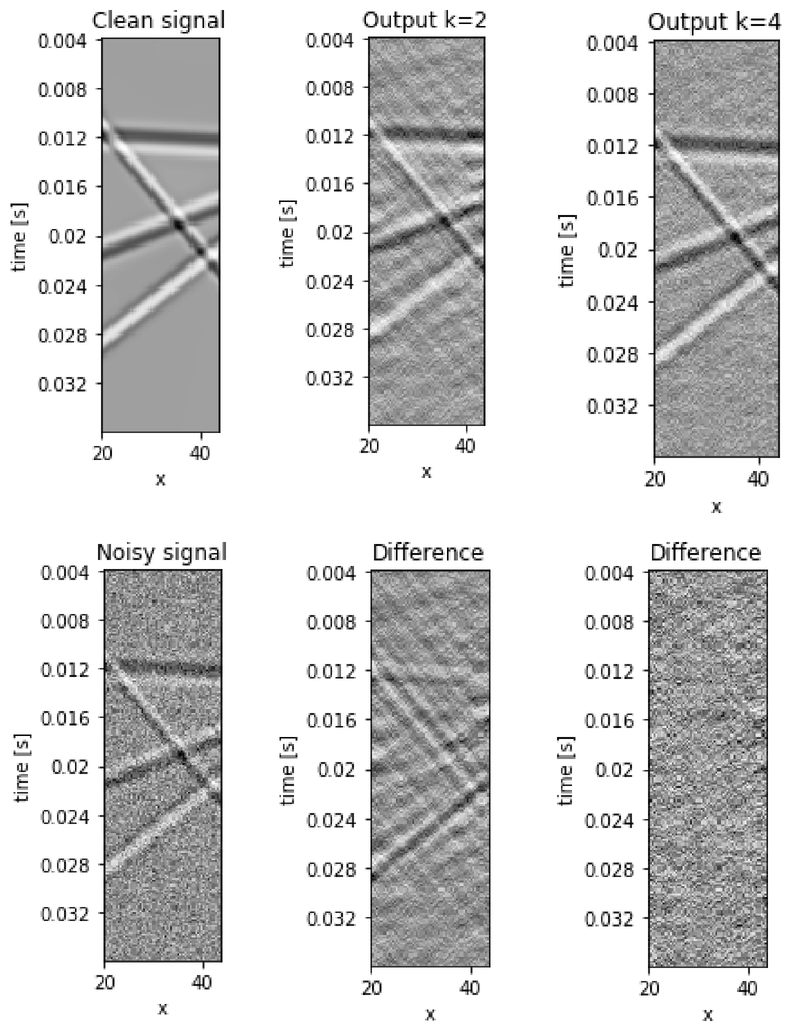

Figure 1. Application of denoising algorithm for a section of synthetic 2D data. Exact number of expected linear events gives the best approximation to the true data while a lower number of desired linear events will not accurately represent the true signal. This is shown by the leakage of signal in the difference panel. In this example a SNR level of 0.5 was used.

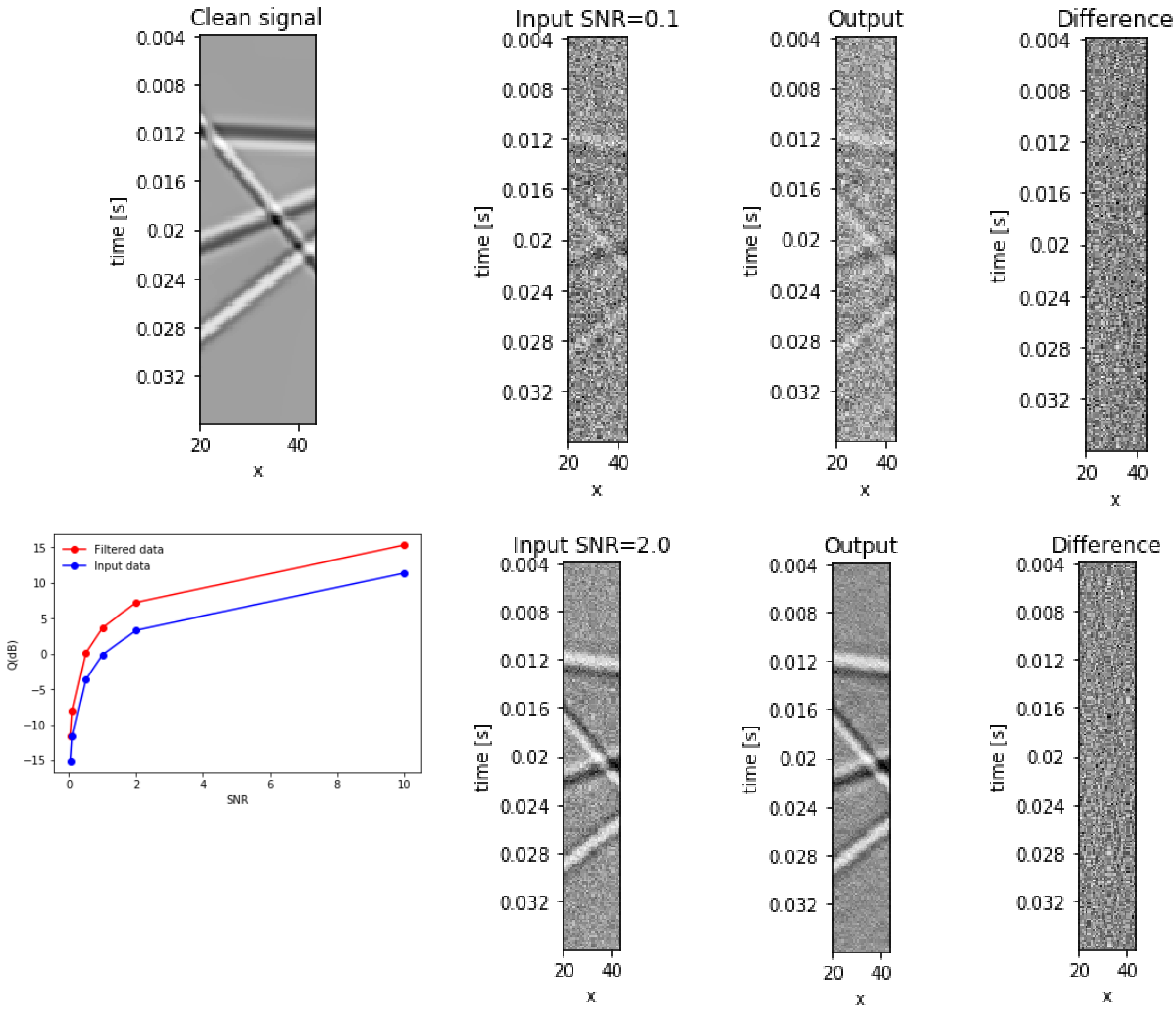

Figure 2. Quality of noise attenuation as a function of input SNR. There is a sharp increase in the quality for lower levels of SNR while for larger values the improvement seems to slow down. This is because high level of input SNR is already desirable while lower levels require harsher denoising. From the difference panels we see that data recovery is very accurate in the sense that there is no observable leakage of signal and only random noise appears in the difference between the estimate and true data.

2D interpolation

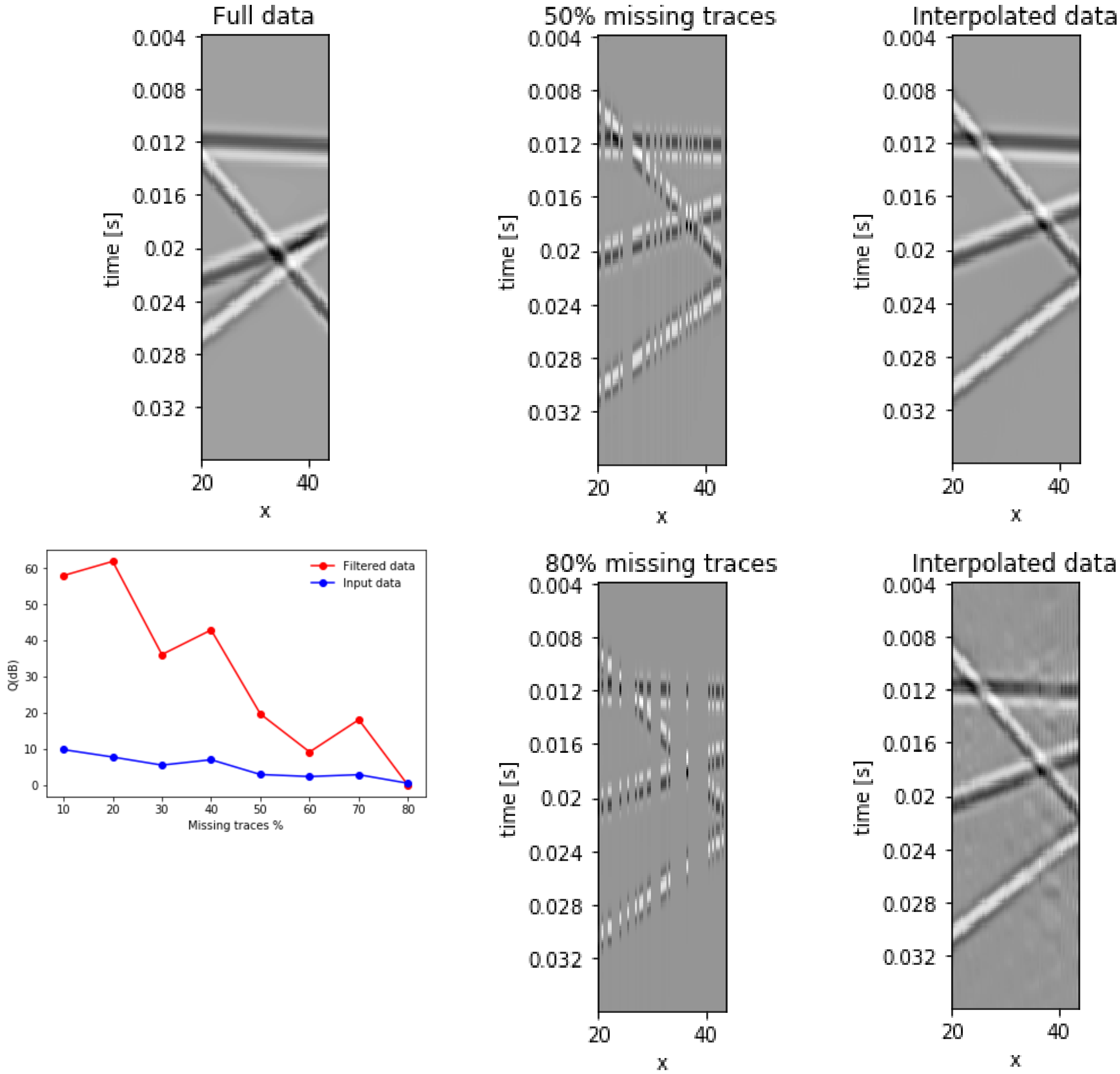

Figure 3. Interpolation of a 2D noise-free section. This is perhaps the area where this method is compromised. A systematic decrease in the quality of data estimation for increasing percentage of missing traces. Although the decrease in quality is eminent, we are still able to obtain a much better estimate for the data when compared to the input data. Although the full amplitude of the traces is not recovered, the timing is. A larger number of iterations will lead to better interpolation at the expense of longer computation time.

2D denoising + interpolation

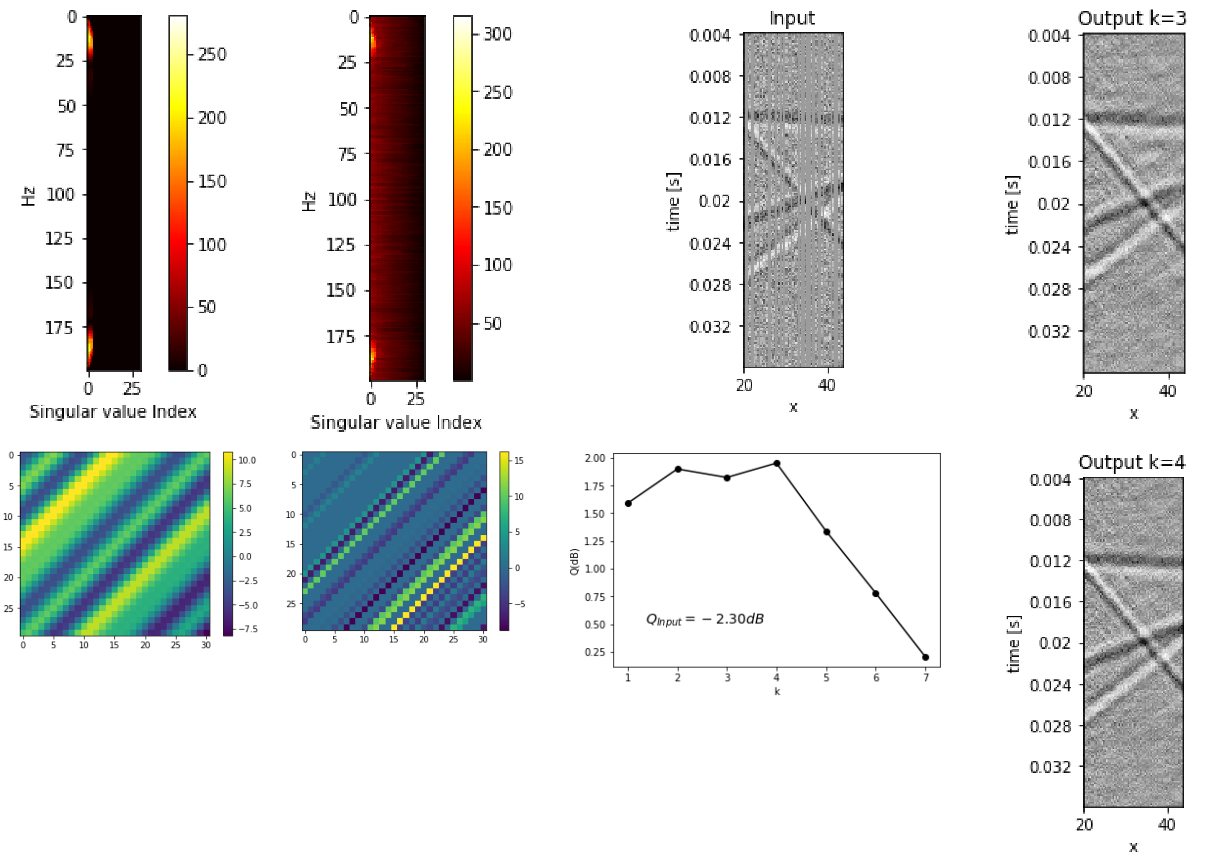

Figure 4. Left two columns: Distribution of singular values for full noise-free synthetic data (Top left), distribution of singular values for incomplete noisy synthetic data (Top right), Hankel matrix at frequency of 20Hz for full, clean data (Bottom left) and Hankel matrix at 20 Hz for noisy and incomplete data (Bottom right). Right two columns: 2D slice of 3D synthetic data decimated at 50% and snr of 0.5 (Top left), output for three and four linear events (Right column) and quality of output data as a function of input rank (Bottom left). Missing samples and random noise will increase the amplitude of the non-zero singular values. The optimal k is four, this is because there are four distinct dips in the window of analysis. A larger value of k means that less noise is being filtered, hence the sharp decrease in quality of reconstruction.

3D denoising + interpolation

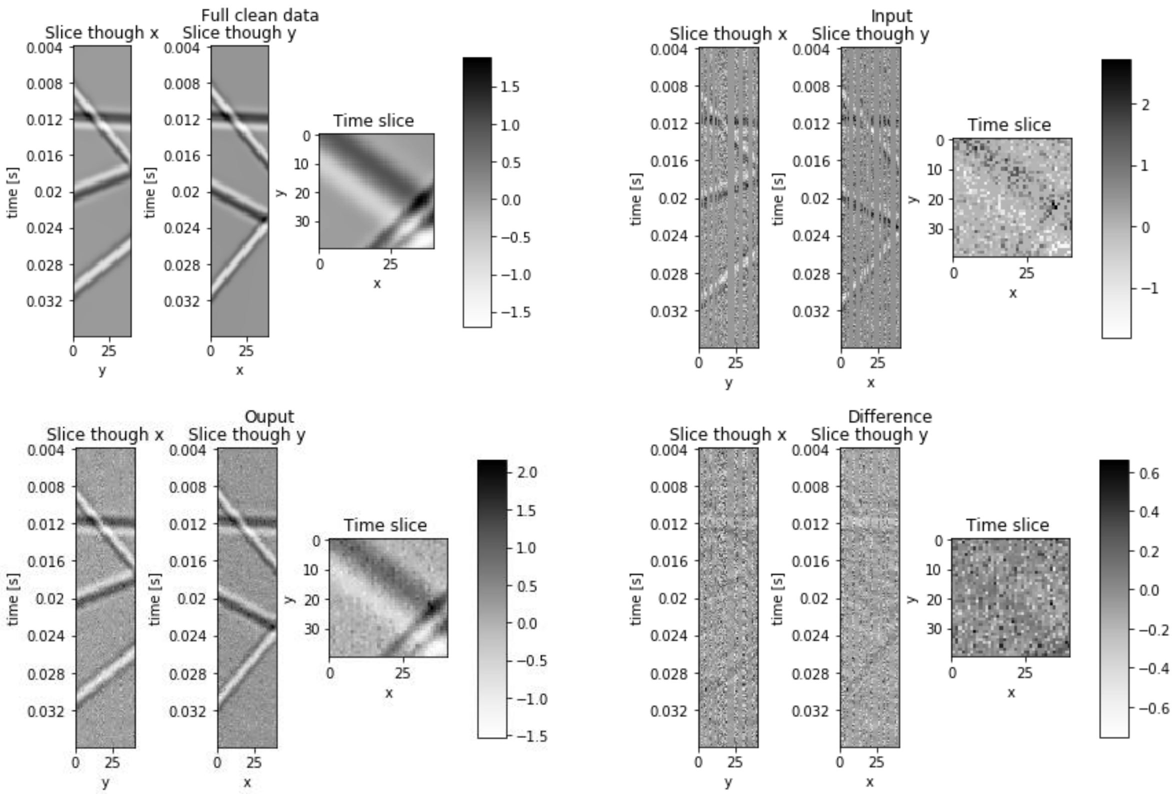

Figure 5. Top left: Vertical slices through x and y, and a time slice of the clean full 3D data set. Top right: Slices of the input data with 50% of missing observations and SNR of 0.5. Bottom left: Output data with $k = 4$, and $\alpha = 0.4$. Bottom right: Difference between output and clean data set. There is a considerable enhancement in all vertical and horizontal slices of the input data. This is reinforced by the difference of these data sets which shows very little data leakage and good amplitude recovery. This example ran for 10 iterations.

Conclusions

It has been demonstrated that SSA and MSSA are a powerful tool for removing random noise from seismic records. Furthermore, interpolation of seismic traces is also achieved when coupled with a reinsertion algorithm. This data enhancing method is tested at different levels of SNR and sparsity, as well as for different number of linear events in the window of analysis and showed significant improvements in the quality of the data set. The effectiveness of this method relies on the predictability of the signal in space and the level of redundancy of information in each frequency slice. Hence the increased quality of reconstruction for increased dimensionality of the data.

The method under examination is dependent on several parameters such as; number of iterations, noise parameter α, and desired rank k. It is demonstrated that a couple shy from a dozen iterations are enough to obtain satisfactory levels of amplitude recovery and filtering. Different values of α are used depending on the level of SNR in the input data, $\alpha = 1$ works particularly well for noise-free data, while $\alpha = 0.4$ worked well for noisy data. As expected, lower values of k will perform harsher levels of noise reduction at the cost of under-representing the data. On the other hand, higher values of k will allow more noise to “leak” to the enhanced data.

Although results are impressive for reasonable levels of noise and sparsity it is important to highlight that there is a systematic decrease in the quality of data recovery for increasing levels of noise and missing observations. When the number of events is known a priori, employment of the algorithm is relatively simple, such is the case for synthetic data. When dealing with real data, the number of distinct linear events is unknown, and one is required though trial and error to suggest an optimal for k. Lastly, a considerable increase in the computational cost comes with the increased amount of information obtained by considering additional spatial coordinates. Thus, creating a trade-off between quality of reconstruction and computational cost and a demand for fast, memory efficient and accurate algorithms.

Finally, it is worth recognizing the potential of rank reduction methods for reconstruction of seismic data. Improvements of this method could have a significant impact on various areas of Geophysics where data enhancement is required as a consequence of the acquisition process.

References

Abma, R., and N. Kabir 2006, 3D interpolation of irregular data with a POCS algorithm: Geophysics, **71**, no. 6, E91-E97.

Andrews, A. C. and Patterson, C. L., 1976, Outer product expansions and their uses in digital image processing: IEEE Transactions on Computers, **25**, 140-148.

Browne, K., Qiao, S., and Wei, Y., 2009, A Lanczos bidiagonalization algorithm for Hankel matrices: Science Direct Linear Algebra and its Applications, **430**, 1531-1543.

Cadzow, J, A., 1988, Signal Enhancement-A composite mapping algorithm: IEEE Transactions on Acoustics, Speech, and Signal Processing, **36**, 49-62.

Gao, J., Sacchi, M. D., and Chen X.,2013, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: Geophysics, **78**, 21-30

Hansen, C. 1987, Truncated *SVD* as a method for regularization: BIT Computer Science and Numerical Mathematics, **27**, 534-553.

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

Oropeza, V., and, Sacchi, M. D., 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, 25-32.

Trickett, S. R., 2001, Eigenimage noise suppression of 3-D seismic in the FXY domain: 71st Ann. Internat. Mtg: Soc. Of. Expl. Geophys., 1981-1984.

Trickett, S. R., 2003, F-xy eigenimage noise suppression: Geophysics, **68**, 751-759.

Trickett, S. R., 2008, F-xy Cadzow noise suppression: 78th Annual International Meeting, SEG, Expanded Abstracts, 2586-2590.

Sacchi, M. D., 2009, FX Singular Spectrum Analysis: CSPG CSEG CWLS Convention, 392-395.

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