Journal of Geophysics and Engineering

PAPER

A robust polynomial principal component analysis for seismic noise attenuation

To cite this article: Yuchen Wang *et al* 2016 *J. Geophys. Eng.* **13** 1002

IOP Publishing
Journal of Geophysics and Engineering
J. Geophys. Eng. **13** (2016) 1002–1009
doi:10.1088/1742-2132/13/6/1002

A robust polynomial principal component analysis for seismic noise attenuation

Yuchen Wang, Wenkai Lu, Benfeng Wang and Lei Liu

Easysignal group, Department of Automation, State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, People's Republic of China

E-mail: lwkmf@mail.tsinghua.edu.cn

Received 14 March 2016, revised 6 September 2016
Accepted for publication 11 October 2016
Published 28 November 2016

Abstract
Random and coherent noise attenuation is a significant aspect of seismic data processing, especially for pre-stack seismic data flattened by normal moveout correction or migration. Signal extraction is widely used for pre-stack seismic noise attenuation. Principle component analysis (PCA), one of the multi-channel filters, is a common tool to extract seismic signals, which can be realized by singular value decomposition (SVD). However, when applying the traditional PCA filter to seismic signal extraction, the result is unsatisfactory with some artifacts when the seismic data is contaminated by random and coherent noise. In order to directly extract the desired signal and fix those artifacts at the same time, we take into consideration the amplitude variation with offset (AVO) property and thus propose a robust polynomial PCA algorithm. In this algorithm, a polynomial constraint is used to optimize the coefficient matrix. In order to simplify this complicated problem, a series of sub-optimal problems are designed and solved iteratively. After that, the random and coherent noise can be effectively attenuated simultaneously. Applications on synthetic and real data sets note that our proposed algorithm can better suppress random and coherent noise and have a better performance on protecting the desired signals, compared with the local polynomial fitting, conventional PCA and a L1-norm based PCA method.

Keywords: noise attenuation, principle component analysis, L1-norm, AVO

(Some figures may appear in colour only in the online journal)

Introduction

In seismic data processing, the separation of signal and noise is a significant issue, because the random and coherent noise widely exists in seismic data. To attenuate these two kinds of noise, a number of methods have been proposed in previousdecades.

For random noise attenuation, several methods exist. Based on signal predictability, the prediction filter in the frequency-space ($f - x$) domain was proposed and became a widely used technique (Canales 1984). Based on it, Wang (1999) proposed a random noise attenuation method, which overcomes the transient-error inherent and minimizes prediction residues. The polynomial fitting (PF) (e.g. Yu *et al* 1988) technique is another useful tool for random noise suppression. Although it can remove random noise effectively by considering the continuity of the desired signals, it can also damage them. Then, the edge-preserving smoothing (EPS) method (Luo *et al* 2002) was proposed to preserve desired signals as well as attenuate random noise. Since then, the EPS method was extended to 3D cases (AlBinHassan *et al* 2006). Lu and Lu (2009) proposed the edge-preserving polynomial fitting (EPPF) method, which had several merits compared with EPS and PF methods.

The attenuation of coherent noise is another main issue in seismic signal enhancement. Coherent noise suppression methods can be classified into two types: global filtering and local filtering. Among global filtering methods, the Radon transform filtering is widely used, with which Turner (1990) proposed a coherent noise suppression method. However, a common drawback of global filters is that it may cause wormy appearances in the results. Among local filtering methods, many different methods are proposed in recent years. Lu (2001) proposed a localized 2D filter in the Fourier projection domain (FPF). The methods in the radial trace domain were applied to both prestack and poststack seismic data by Brown and Claerbout (2000) and Henley (2003). Polynomial fitting, as a random noise suppression method, can also be used in coherent noise attenuation. Lu *et al* (2006) proposed a local polynomial approximation (LPA) based coherent noise attenuation method, which can estimate locally linear coherent noise (LLCN) effectively, proving that a LPA filter can gain a brilliant performance in signal enhancement. Even though, polynomial fitting based filters have a common drawback of low robustness to outliers, and they are also sensitive to the size of the analysis window.

Principle component analysis (PCA) is another useful tool in signal enhancement, especially for the pre-stack gathers flattened by normal moveout (NMO) correction or migration. Ulrych *et al* (1999) applied a singular value decomposition (SVD) filter to seismic image enhancement. Although a conventional SVD filter based on the L2-norm can separate signal and noise subspaces, it shows low robustness to outliers. To alleviate this drawback, much work has been done (Baccini *et al* 1996, Aanas *et al* 2002, Ke and Kanade 2003, 2005, De La Torre *et al* 2003). Ding *et al* (2006) proposed the R1-PCA algorithm, which keeps the merits of L2-PCA and at the same time reduces the effects of outliers. Besides, it owns the advantage of rotational invariance compared to L1-PCA. Kwak (2008) proposed the PCA-L1 algorithm, which maximizes the L1-norm in the feature space instead of maximizing variance, in which the proposed L1-norm optimization algorithm is fast, simple and intuitive using a greedy algorithm.

For seismic signal enhancement, many PCA based algorithms can achieve good performance, but also with some artifacts. A main cause of these artifacts is the strong coherent noise. To overcome this shortcoming, many researchers try to estimate and remove the coherent events from the original data first when using PCA-based multichannel filters. For instance, Lu (2006) used a dip scanning method to estimate the coherent events and extract them before attenuating random noise with a SVD filter. However, the performance can deteriorate if there are multiple coherent noises with different slopes, and the computational cost will increase at the same time. Besides, L1-norm based robust PCA methods were introduced, instead of the conventional PCA method. This solution can obtain a slight improvement to the results according to the signal-to-noise ratio (SNR), but the artifacts still exist to some extent. In this paper, we focus on the noise attenuation for flattened prestack gathers, where the amplitude variation with offset (AVO) property exists. We find that the influence of the coherent noise can be eliminated if we use the AVO property to constrain the coefficient vectors during the PCA. Thus, we propose a robust polynomial PCA (RPPCA) algorithm by optimizing the coefficient vectors with a polynomial constraint and apply it to seismic random and coherent noise attenuation. The proposed method can separate signal and noise subspaces directly without estimating the coherent events and be robust to outliers. To ameliorate the problem and make sure that the subspaces are orthogonal with each other, we simplify the problem to a sub-optimal form and solve it with a greedy algorithm (Kwak 2008). In practice, we subtract a certain percentage of the energy of the original signal in each processing window. Thus, the number of maintained principle components (PCs) can be chosen adaptively in each calculation, which can insure that signal with higher SNR can have more PCs maintained.

This paper is organized as follows. In Theory, the procedure of our algorithm is described and some details are discussed. Some synthetic and real data examples are shown to illustrate the performance of the proposed algorithm in Examples. Finally, we give our conclusions in Conclusions.

Theory

Review of PCA-L1 algorithm (Kwak 2008)

Assume that $\mathbf{D} = [\mathbf{d}_1, \cdots, \mathbf{d}_n] \in \mathbf{R}^{m\times n}$ is the original flattened prestack seismic data, where m denotes the number of samples in the time direction and n indicates the number of traces. To separate signal and noise subspaces with conventional PCA, a projection matrix $\mathbf{W} \in \mathbf{R}^{m\times k}$ should be established and its corresponding coefficient matrix $\mathbf{V} \in \mathbf{R}^{k\times n}$ which can satisfy the following minimization problem:

$$\min_{\mathbf{W},\mathbf{V}} \|\mathbf{D} - \mathbf{W}\mathbf{V}\|_2^2, \text{ s.t. } \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}_k, \tag{1}$$

where $\|\cdot\|_2^2$ is defined as the energy of a matrix in this paper. We can get the global minimum of equation (1) by the SVD algorithm.

It is well known that L2-norm optimization shows low robustness to outliers, which make an SVD filter introduce some artifacts when applied to signal and image enhancement. To improve its robustness, many researchers drew their attention to the PCA methods based on L1-norm optimization (Baccini *et al* 1996, Aanas *et al* 2002, Ke and Kanade 2003, 2005, Torre *et al* 2003, Ding *et al* 2006, Kwak 2008), which can be expressed as:

$$\min_{\mathbf{W},\mathbf{V}} \|\mathbf{D} - \mathbf{W}\mathbf{V}\|_1, \text{ s.t. } \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}_k. \tag{2}$$

Since the exact optimum of equation (2) is difficult to get and is not invariant to rotations, Kwak (2008) solved the following maximization problem instead:

$$\max_{\mathbf{W}} \|\mathbf{W}^{\mathrm{T}}\mathbf{D}\|_1 = \max_{\mathbf{W}} \sum_{i=1}^{n}\sum_{j=1}^{k}\left|\sum_{l=1}^{m} w_{lj}d_{li}\right|, \text{ s.t. } \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}_k. \tag{3}$$

The projection matrix obtained by solving equation (3) is invariant to rotations and robust to outliers. However, the global optimum of equation (3) for $k > 1$ is difficult to find. To ameliorate the problem, Kwak (2008) used a greedy search method to get a local optimum, in which $k = 1$ is assumed in each iteration, thus equation (3) becomes a series of following problems:

$$\max_{\mathbf{w}} \|\mathbf{w}^{\mathrm{T}}\mathbf{D}\|_1 = \max_{\mathbf{w}} \sum_{i=1}^{n} \left|\mathbf{w}^{\mathrm{T}}\mathbf{d}_i\right|, \text{s.t.} \ \|\mathbf{w}\|_2^2 = 1. \tag{4}$$

With equation (4) solved iteratively, the signal subspace is separated with higher robustness to outliers.

RPPCA algorithm

Although the PCA-L1 algorithm is more robust to outliers than the conventional PCA, it still causes some wormy areas and artifacts in the result because of the strong coherent noise. In our research, we find that the influence of strong coherent noise can be eliminated if we take the AVO of seismic events into account.

In essence, the prestack seismic data can be approximated by:

$$d(i,j) = \sum_{k=1}^{l} a_k(j) s_k(i), \tag{5}$$

where l is the number of flat events in the selected seismic data, $s_k(i)$ is the seismic wavelet corresponding to the kth event, and $a_k(j)$ is the AVO curves of the kth event. It can be proved that if we use SVD to separate the signal and noise subspaces, every coefficient vector will vary following a polynomial. Furthermore, a seismic data consisting of several n-order polynomial reflection coefficients can be decomposed into at most $n + 1$ components. The related proof can be found in the appendix.

Therefore, in order to eliminate the artifacts, we can pose a polynomial constraint to the coefficient vectors when enhancing the seismic signal with PCA filters. Thus, different from the SVD filter, we want to find a projection matrix $\mathbf{W}$ and the compatible coefficient matrix $\mathbf{V}$ to satisfy the following optimization problem:

$$\min_{\mathbf{W},\mathbf{V}} \|\mathbf{D} - \mathbf{W}\mathbf{V}\|_2^2 + \lambda\|\mathbf{V} - \mathbf{V}_0\|_2^2, \text{s.t.} \ \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}_n, \tag{6}$$

where every row of $\mathbf{V}_0$ is a polynomial vector and λ is a weight factor.

However, as mentioned above, SVD-based filters show low robustness to outliers, then a L1, L2-norm optimization is adopted instead of equation (6). The new optimization problem is expressed as below:

$$\min_{\mathbf{W},\mathbf{V}} \|\mathbf{D} - \mathbf{W}\mathbf{V}\|_1 + \lambda\|\mathbf{V} - \mathbf{V}_0\|_2^2, \text{s.t.} \ \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}_n. \tag{7}$$

Similar to Kwak (2008), finding a global optimum of equation (7) for $k > 1$ is difficult. To ameliorate the challenge and ensure the orthogonality at the same time, we replace equation (7) with a sub-optimal problem, which simplifies and approximates equation (7) into a series of vector optimization problem as follows:

$$\min_{\mathbf{v}} \|\mathbf{D}\mathbf{v} - \widetilde{\mathbf{w}}\|_1 + \lambda\left\|\mathbf{v} - \mathbf{C}_{\mathrm{r}}(\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\mathbf{C}_{\mathrm{r}})^{-1}\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\tilde{\mathbf{v}}\right\|_2^2, \tag{8}$$

where $\widetilde{\mathbf{w}}$ is the projection vector obtained by the PCA-L1 algorithm, $\tilde{\mathbf{v}}$ is the corresponding coefficient vector,
$$\mathbf{C}_{\mathbf{r}} = \begin{bmatrix} 1 & 1-\left[\frac{L}{2}\right] & \cdots & \left(1-\left[\frac{L}{2}\right]\right)^{\mathrm{r}} \\ 1 & 2-\left[\frac{L}{2}\right] & \cdots & \left(2-\left[\frac{L}{2}\right]\right)^{\mathrm{r}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & L-\left[\frac{L}{2}\right] & \cdots & \left(L-\left[\frac{L}{2}\right]\right)^{\mathrm{r}} \end{bmatrix},$$
L is the length of $\tilde{\mathbf{v}}$ and $\mathbf{C}_{\mathrm{r}}(\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\mathbf{C}_{\mathrm{r}})^{-1}\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\tilde{\mathbf{v}}$ is the r-rank polynomial approximation of $\tilde{\mathbf{v}}$.

These optimization problems are solved with a greedy search method, and a L1–L2 optimization problem for each dimension should be solved. Here, we adopt the iterative reweighted least square (IRLS) algorithm to solve the above mentioned optimization problem. IRLS is a fast L1 estimation algorithm and has been applied into many seismic data processing problems (Guitton *et al* 2004).

For the vector $\boldsymbol{\beta} = \mathbf{D}\mathbf{v} - \widetilde{\mathbf{w}} \in \mathbf{R}^k$, the L1-norm is defined as

$$f(\boldsymbol{\beta}) = \|\boldsymbol{\beta}\|_1 = \sum_{i=1}^{k} |\beta_i|. \tag{9}$$

In the IRLS algorithm, $f(\boldsymbol{\beta})$ in each iteration is written as

$$f^{(t)}(\boldsymbol{\beta}) = \sum_{i=1}^{k} (\beta_i^{(t)})^2 / \left(\left|\beta_i^{(t-1)}\right| + \gamma\right), \tag{10}$$

where t is the number of iterations and γ is a regularization factor.

Thus, the objective function value in each iteration should be written as

$$f^{(t)}(\mathbf{v}) = \sum_{i=1}^{n} \left(\mathbf{d}_i(\mathbf{v}_i^{(t-1)})^{\mathrm{T}} - \widetilde{w}_i\right)^2 / \left(\left|\mathbf{d}_i(\mathbf{v}_i^{(t-1)})^{\mathrm{T}} - \widetilde{w}_i\right| + \gamma\right) + \lambda\left\|\mathbf{v}^{(t-1)} - \mathbf{C}_{\mathrm{r}}(\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\mathbf{C}_{\mathrm{r}})^{-1}\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\tilde{\mathbf{v}}\right\|_2^2. \tag{11}$$

In practice, we set an energy-proportion threshold to end the iterations, which is determined by the SNR of the processed signal. In this paper, we estimate the SNR as follows:

$$\mathrm{SNR} = \frac{\lambda_1'}{\sum_{i=1}^{n} \lambda_i}, \tag{12}$$

where λ_1 is the largest eigenvalue and $\sum_{i=1}^{n} \lambda_i$ is the sum of all eigenvalues. Thus, the energy-proportion can be set as

$$\eta = 1 - \left(1 - \frac{\lambda_1}{\sum_{i=1}^{n} \lambda_i}\right)^2. \tag{13}$$

By doing so, our algorithm can adaptively choose the number of maintained PCs. Then, the procedures of the proposed algorithm can be summarized as follows:

(1) Initialization: Pick any $\mathbf{w}(0)$. Set $\mathbf{w}(0) \leftarrow \mathbf{w}(0)/\|\mathbf{w}(0)\|_2$ and $t = 0$. Set λ and energy-proportion η to suitable values. Set $\widetilde{\eta} = 0$.

(2) For all $i \in \{1, \ \cdots, \ n\}$, if $\mathbf{w}^{\mathrm{T}}(t)\mathbf{d}_i < 0$, $p_i(t) = -1$, otherwise $p_i(t) = 1$.

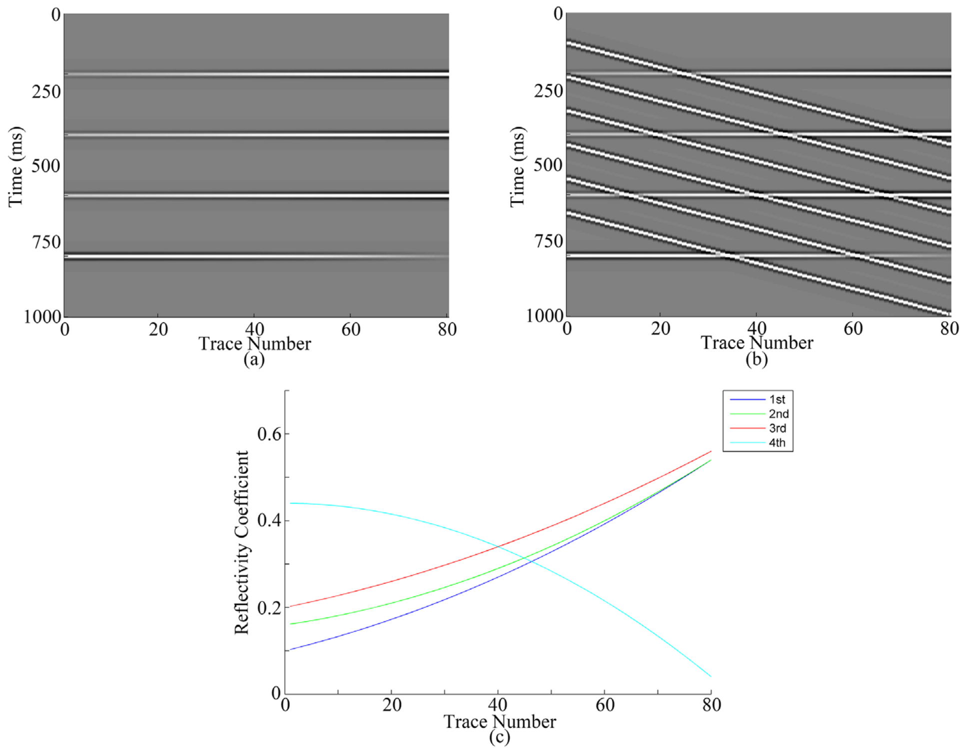

Figure 1. Synthetic data example I. (a) Desired signal; (b) desired signal with six coherent events; (c) the AVO curves of the desired events.

(3) Set $t \leftarrow t+1$ and $\mathbf{w}(t)=\sum_{i=1}^{n} p_i(t-1)\mathbf{d}_i$. Set $\mathbf{w}(t) \leftarrow \mathbf{w}(t)/\|\mathbf{w}(t)\|_2$.

(4) Converge check: (When Convergent, Stop.)

a. If $\mathbf{w}(t) \neq \mathbf{w}(t-1)$, go to Step 2.

b. Else if there exists i such that $\mathbf{w}^{\mathrm{T}}(t)\mathbf{d}_i = 0$, set $\mathbf{w}(t) \leftarrow (\mathbf{w}(t)+\Delta\mathbf{w})/\|\mathbf{w}(t)+\Delta\mathbf{w}\|_2$ and go to Step 2. Here, Δw is a small nonzero random vector.

c. Otherwise, set $\widetilde{\mathbf{w}}=\mathbf{w}(t)$.

(5) Set $\tilde{\mathbf{v}}=\mathbf{D}^{\mathrm{T}}\widetilde{\mathbf{w}}$.

(6) Solve $\min_{\mathbf{v}}\|\mathbf{D}\mathbf{v}-\widetilde{\mathbf{w}}\|_1+\lambda\left\|\mathbf{v}-\mathbf{C}_{\mathrm{r}}(\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\mathbf{C}_{\mathrm{r}})^{-1}\mathbf{C}_{\mathrm{r}}^{\mathrm{T}}\tilde{\mathbf{v}}\right\|_2^2$ with IRLS.

(7) Set $\mathbf{w}=\mathbf{D}\cdot\mathbf{v}/\|\mathbf{v}\|_2$.

(8) Set $\tilde{\eta}=\tilde{\eta}+\|\mathbf{w}\mathbf{v}^{\mathrm{T}}\|_2^2/\|\mathbf{D}\|_2^2$,

a. If $\tilde{\eta}<\eta$, Set $\mathbf{D}=\mathbf{D}-\mathbf{w}\mathbf{v}^{\mathrm{T}}$. Go to Step 1.

b. Else end loop.

In summary, the proposed method separates the signal and noise subspaces and maintains the AVO curve by solving an L1–L2 problem iteratively. The proposed algorithm can attenuate random noise and coherent noise effectively when applied to horizontally aligned seismic data and protect the desired signal at the same time. The performance of the proposed method is demonstrated in the next part.

Examples

Synthetic data example I

To illustrate our assertion in the Theory part and the artifacts brought by the coherent noises, we simulate a dataset and use the SVD filter, the PCA-L1 filter and the RPPCA filter to extract the desired signal. The size of the dataset is 80 samples in the lateral direction and 1000 ms in the time direction. The

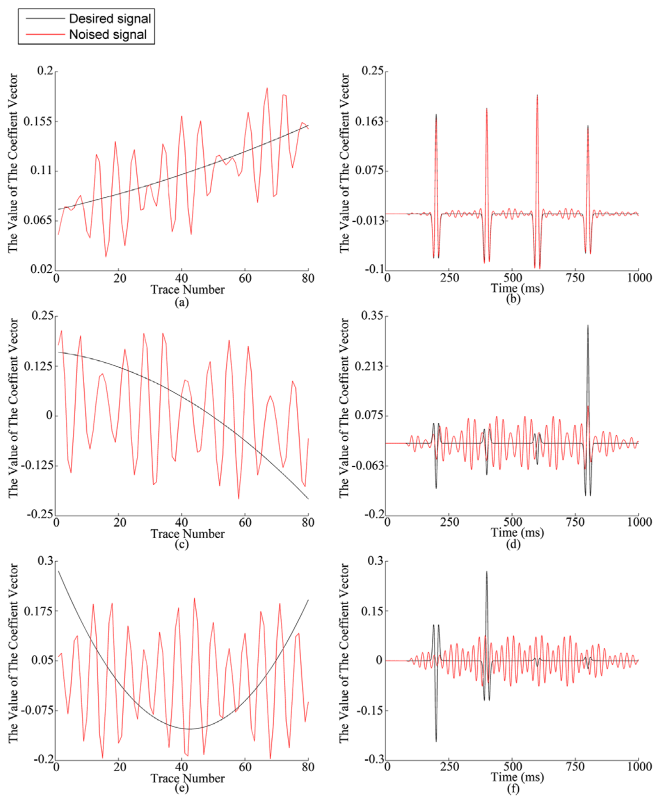

Figure 2. Synthetic data example I. (a) The first coefficient vector obtained by SVD; (b) the first projection vector obtained by SVD; (c) the second coefficient vector obtained by SVD; (d) the second projection vector obtained by SVD; (e) the third coefficient vector obtained by SVD; (f) the third projection vector obtained by SVD.

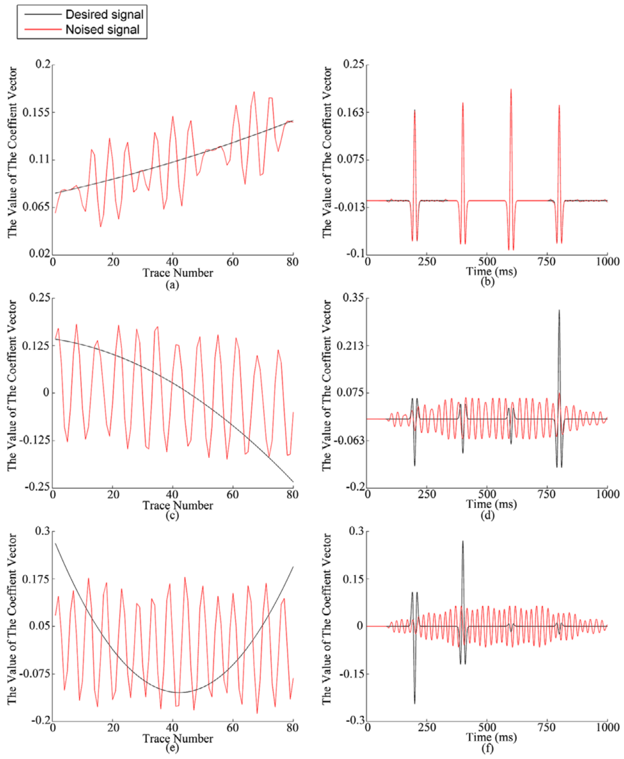

Figure 3. Synthetic data example I. (a) The first coefficient vector obtained by PCA-L1; (b) the first projection vector obtained by PCA-L1; (c) the second coefficient vector obtained by PCA-L1; (d) the second projection vector obtained by PCA-L1; (e) the third coefficient vector obtained by PCA-L1; (f) the third projection vector obtained by PCA-L1.

time interval is 1 ms. As shown in figure 1(a), the dataset consists of four desired events with different AVO curves, which are quadratic polynomial, shown in figures 1(c) and sixcoherent events with constant reflection coefficient are inserted in figure 1(b). We use the SVD, PCA-L1 and RPPCA filters to obtain the first three principles of the desired signal and the noisy signal. The coefficient and projection vectors obtained by SVD are shown in figures 2(a)–(f), and those obtained by PCA-L1 and RPPCA are shown in figures 3(a)–(f) and 4(a)–(f), respectively. The vectors extracted from the desired signal and the noisy signal are denoted by black lines and red lines, respectively. In this example, the size of the analysis window is equal to the data size.

It can be noted from the black lines of figures 2 and 3, that the coefficient vectors follow different quadratic polynomials. In constrast, the red lines note that the coherent events bring strong fluctuations to the projection vectors and the coefficient vectors which make them deviate from the true values. Compared with figure 2, figure 3 shows that PCA-L1 is more robust to outliers than SVD and it alleviates the fluctuations on the projection vectors and the coefficient vectors. However, the artifacts still exist to some extent.

In contrast, comparing the red and black lines in figure 4, we can note that the result obtained from the noisy signal is quite similar to the true value, especially in the first two components, which indicates that the influence of the coherent

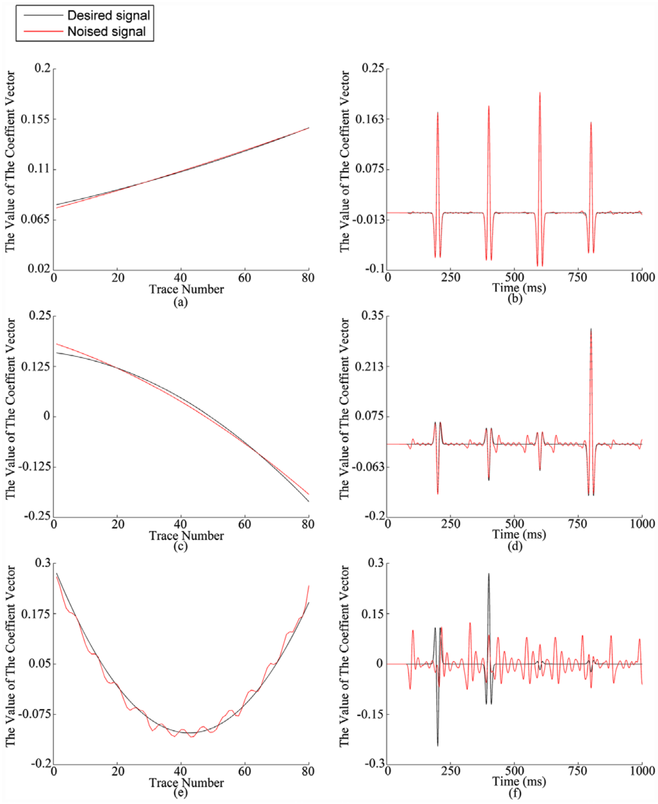

Figure 4. Synthetic data example I. (a) THE first coefficient vector obtained by RPPCA; (b) the first projection vector obtained by RPPCA; (c) the second coefficient vector obtained by RPPCA; (d) the second projection vector obtained by RPPCA; (e) the third coefficient vector obtained by RPPCA; (f) the third projection vector obtained by RPPCA.

events is eliminated effectively and the robustness to outliers is enhanced with the proposed method.

Synthetic data example II

In this example, synthetic seismic data shown in figure 5 consists of eight flat seismic events with different AVO curves, eight coherent noises with different dips, and -5 dB additive Gaussian random noise. The desired signal is shown in figure 5(a) and the corresponding AVO curves are shown in figure 5(c). Figure 5(b) shows the noisy signal. Results obtained by LPF, SVD, PCA-L1 and RPPCA are shown in figures 6(a)–(d), respectively. The corresponding noise removed by each method is shown in figures 7(a)–(d), respectively. In this example, the size of the processing window used for the SVD filter, PCA-L1 filter and RPPCA filter is 200 samples in the time direction and 120 samples in the lateral direction, and the size of the processing window used for LPF is 80. The number of PC we use in the three algorithms is two, and the rank of polynomial fitting used in LPF and RPPCA method is two. The reason that we do not use the third PC is that the SNR of it is too low to use. Figures 6(b) and (c) show that SVD and PCA-L1 have good performance in attenuating random noise and coherent noise, but with strong artifacts and damage the desired signal. By contrast, RPPCA and

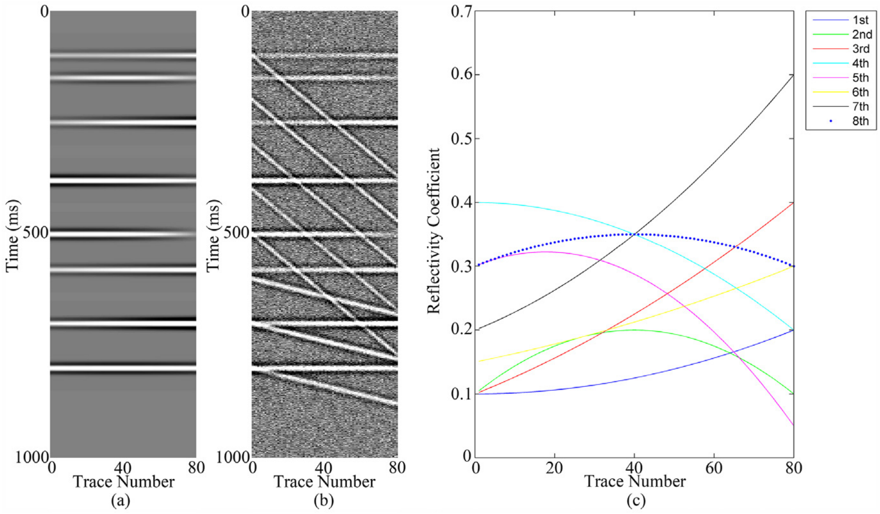

Figure 5. Synthetic data example II. (a) Original signal; (b) noisy signal; (c) the AVO curves of the eight desired events.

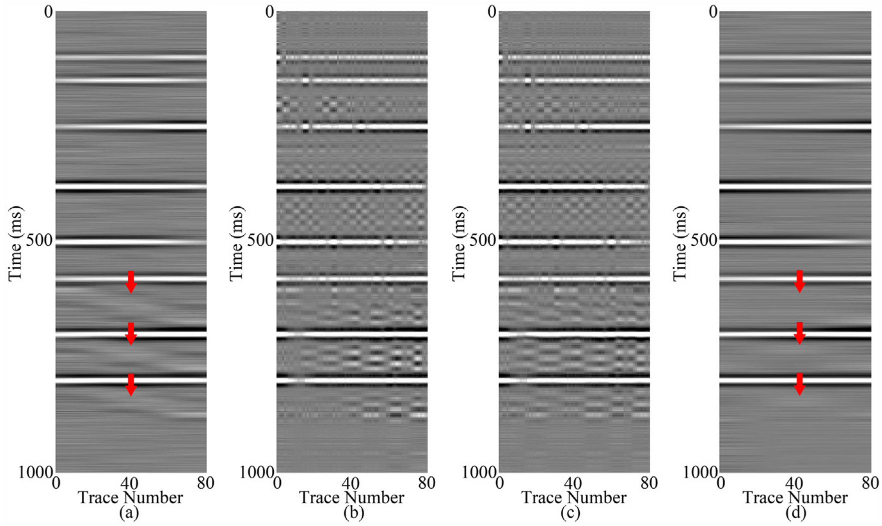

Figure 6. Synthetic data example II. (a) Result obtained by LPF filter; (b) result obtained by SVD filter; (c) result obtained by PCA-L1 filter; (d) result obtained by RPPCA filter.

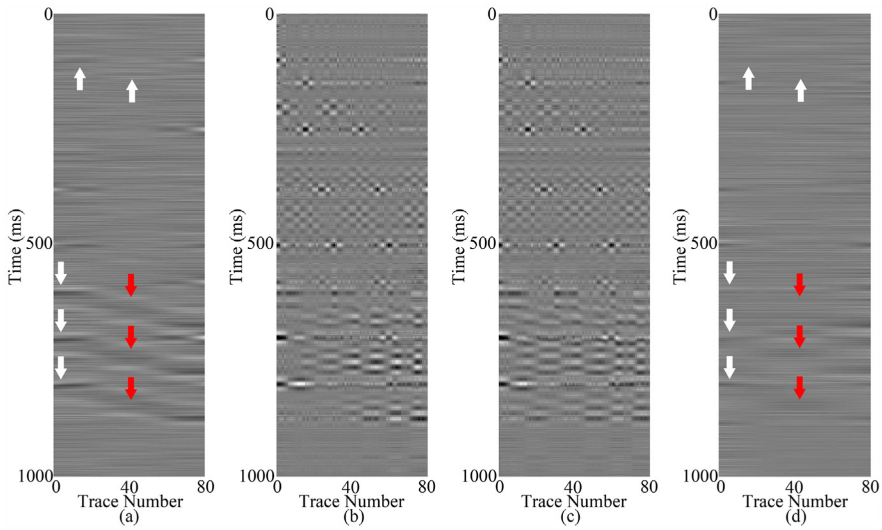

Figure 7. Synthetic data example II. (a) Removed noise by LPF filter; (b) removed noise by SVD filter; (c) removed noise by PCA-L1 filter; (d) removed noise by RPPCA filter.

Table 1. SNR of the four results obtained by the LPF, SVD, PCA-L1 and RPPCA methods in synthetic data example II.

Method	LPF	SVD	PCA-L1	RPPCA
SNR (dB)	16.77	13.89	14.82	18.96

LPF have better performance in protecting desired signal and eliminating artifacts. Comparing figures 6(a) and (d), we can also find that RPPCA removes noise better than LPF, especially in the areas marked by the red arrows. When the dip of the coherent noise is small, the LPF filter may contain some residue. In contrast, the result obtained by the RPPCA filter is cleaner in those areas. Figures 7(a)–(d) indicate the noise removed by the four filters. Comparing figures 7(a)–(d), we can conclude that RPPCA has a better performance in protecting valid signal. Furthermore, the result obtained by the RPPCA method has less edge effect compared with LPF. Table 1 shows the signal-noise-ratio of the results obtained by the four different methods. This is a statistics result of 100 times experiments.

Real data

To further demonstrate the performance of the proposed method, we conduct an experiment on a field data set acquired in a desert in northwest China, which is arranged in common imaging point gathers (CIG). The whole data set has 830 gathers, and there are 23 traces in each gather. The length of each trace is 1100 ms, and the sampling time interval is 2 ms. The size of analysis window is 23 samples in the lateral direction and 100 samples in the time direction. The polynomial rank used in LPF and RPPCA is five. Figure 8(a) shows one gather of original data, which notes that the coherent noise and the random noise are complicated. Firstly, the LPF filter is used to attenuate the noise, and the denoised result and the removed noise are shown in figures 8(b) and 9(a), respectively. Although much noise is suppressed, a little coherent noise remained. The denoised result obtained by the SVD filter is shown in figure 8(c), and the corresponding removed noise is shown in figure 9(b). The SVD filter suppresses most coherent noise and random noise, but there exist some wormy and fluctuant appearance. In contrast, the PCA-L1 filter is a little more robust than the SVD filter, but, the denoised result is still distorted in a certain degree. The denoised result obtained by the PCA-L1 filter and the corresponding removed noise are shown in figures 8(d) and 9(c), respectively. Figures 8(e) and 9(d) indicate the denoised result obtained by the proposed RPPCA method and the removed noise, respectively, which note that the coherent noise and the random noise are suppressed effectively, and the distortion of the desired signal is also reduced. In this example, the size of the processing window used in figures 8(c)–(e) is equal within each CIG gather, and the rank of polynomial fitting used in figures 8(b) and (e) is three. In the experiments of the SVD, L1-PCA and RPPCA algorithms, we subtract 55% of the total energy for every processing window, thus the algorithm can choose the number of PCs adaptively.

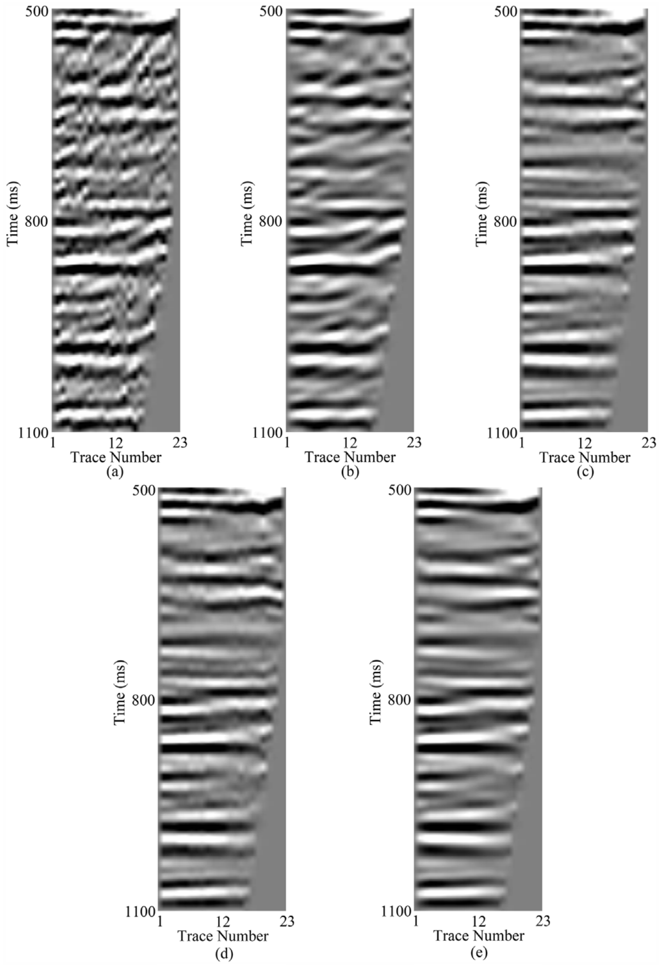

Figure 8. Real data example with a common imaging point gather. (a) Original data; (b) result obtained by LPF filter; (c) result obtained by SVD filter; (d) result obtained by PCA-L1 filter; (e) result obtained by RPPCA filter.

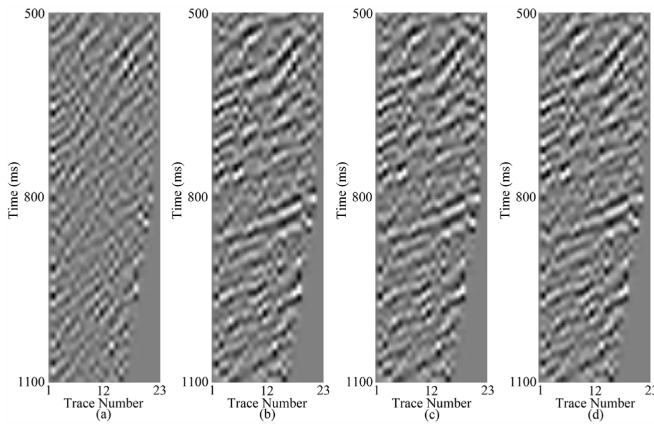

Figure 9. Real data example with a common imaging point gather. (a) Removed noise by LPF filter; (b) removed noise by SVD filter; (c) removed noise by PCA-L1 filter; (d) removed noise by RPPCA filter.

Conclusion

We propose a robust polynomial principle component analysis method for the suppression of random noise and coherent noise contaminated in flattened prestack seismic data. By posing polynomial constraints to the coefficient vectors, the proposed method can better extract the desired signal when it is polynomial. Because of the AVO of seismic events, when applied to seismic noise attenuation, the proposed method can reduce random noise and coherent noise effectively and overcome the drawback of causing artifacts on desired signal. Furthermore, the proposed method can attenuate the two kinds of noise simultaneously instead of estimating the coherent events at first. Applications on synthetic data and real data show that the proposed algorithm can achieve a good performance on enhancement and protection of desired signals. By setting an energy-proportion threshold, we make our algorithm subtract a certain percentage of the energy of the original signal in each window. Thus, our algorithm can choose the number of maintained PCs adaptively. The example on a real data set shows that our proposed method can be effective on CIGs, which can lay the foundation for several other seismic processing or interpretation works, like AVO analysis.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (41604096, 41674116) and the National Science and Technology of Major Projects of China under Grant (2016ZX05024001-005).

Appendix

Since an AVO curve follows a polynomial, the prestack seismic data can be expressed as:

$$d(i,j) = \sum_{k=1}^{l} s_k(i) \sum_{h=0}^{r} p_{kh} j^h, \tag{A.1}$$

where r is the rank of the polynomial, and $\mathbf{p}_k = [p_{k0} \ , \ \cdots \ , \ p_{kr}]^{\mathrm{T}}$ are the polynomial coefficients.

Rewriting equation (A.1), we have,

$$d(i,j) = \sum_{k=1}^{l} s_k(i) \sum_{h=0}^{r} p_{kh} j^h = \sum_{h=0}^{r} \sum_{k=1}^{l} (s_k(i) p_{kh}) j^h = \sum_{h=0}^{r} q_h(i) j^h. \tag{A.2}$$

Equation (A.2) notes that the amplitude variation with offset follows a polynomial for the same sampling time, the coefficients of which are defined by $\mathbf{q}(i) = [q_0(i) \ , \ \cdots \ , \ q_r(i)]^{\mathrm{T}}$.

Let $\mathbf{w}$ and $\mathbf{v}$ be the first principle component and its corresponding coefficient vector obtained by SVD, then $\mathbf{v}$ should satisfy the following equation:

$$\mathbf{v} = \underset{\mathbf{w},\mathbf{v}}{\arg\min} \sum_{i=1}^{m} \| w(i)\mathbf{v} - \mathbf{C}_{\mathrm{r}} \mathbf{q}(i) \|_2^2, \tag{A.3}$$

where $\mathbf{C}_{\mathrm{r}} = \begin{bmatrix} 1 & 1-\left[\frac{L}{2}\right] & \cdots & \left(1-\left[\frac{L}{2}\right]\right)^{\mathrm{r}} \\ 1 & 2-\left[\frac{L}{2}\right] & \cdots & \left(2-\left[\frac{L}{2}\right]\right)^{\mathrm{r}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & L-\left[\frac{L}{2}\right] & \cdots & \left(L-\left[\frac{L}{2}\right]\right)^{\mathrm{r}} \end{bmatrix}$, L is the length of $\mathbf{v}$.

For each i, $w(i)$ is a constant. Thus, the coefficient vector $\mathbf{v}$ is a linear combination of a series of polynomial, which is obviously a polynomial.

Let $\mathbf{D}' = \mathbf{D} - \sum_{i=1}^{p} \mathbf{w}_i \mathbf{v}_i$, where $\mathbf{w}_i$ and $\mathbf{v}_i$ are the ith principle component and its corresponding coefficient vector, respectively. Suppose that for all $i < p$, $\mathbf{v}_{\mathbf{i}}$ follows a polynomial. Thus, the first coefficient vector of $\mathbf{D}'$, which is the $(i+1)$th coefficient vector of $\mathbf{D}$, also follows a polynomial, which means that every coefficient vector of $\mathbf{D}$ is a polynomial.

References

Aanas H, Fisker R, Astrom K and Carstensen J 2002 Robust factorization *IEEE Trans. Pattern Anal. Mach. Intell.* **24** 1215–25

AlBinHassan N M, Luo Y and Al-Faraj M N 2006 3D edge-preserving smoothing and applications *Geophysics* **71** P5–11

Baccini A, Besse P and Falguerolles A D 1996 A L1-norm PCA and a heuristic approach *Ordinal and Symbolic Data Analysis* eds E Diday, Y Lechevalier and P Opitz (Berlin: Springer) pp 359–68

Brown M and Claerbout J 2000 Ground roll and radial trace transform—revisited *Stanford Exploration Project Report* SEP-103 pp 205–31

Canales L L 1984 Random noise reduction *54th Annual Int. Meeting, SEG, Expanded Abstracts* pp 525–7

De La Torre F and Black M J 2003 A frame work for robust subspace learning *Int. J. Comput. Version* **54** 117–42

Ding C, Zhou D, He X and Zha H 2006 R1-PCA: rotational invariant L1-norm principal component analysis for robust subspace factorization *Proc. 23rd Int. Conf. Machine Learning* pp 281–8

Guitton A and Verschuur D J 2004 Adaptive subtraction of multiples using theL1-norm *Geophys. Prospect.* **52** 27–38

Henley D C 2003 Coherent noise attenuation in the radial trace domain *Geophysics* **68** 1408–16

Ke Q and Kanade T 2003 Robust subspace computation using L1 norm *Carnegie Mellon University Technical Report* CMU-CS-03-172

Ke Q and Kanade T 2005 Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming *Proc. IEEE Conf. Computer Vision and Pattern Recognition (June 2005)*

Kwak N 2008 Principal component analysis based on L1-Norm maximization *IEEE Trans. Pattern Anal. Mach. Intell.* **30** 1672–80

Lu W 2001 Localized 2D filter based linear coherent noise attenuation *IEEE Trans. Image Process.* **10** 1379–82

Lu W 2006 Adaptive noise attenuation of seismic images based on singular value decomposition and texture direction detection *J. Geophys. Eng.* **3** 28–34

Lu W, Zhang W and Liu D 2006 Local linear coherent noise attenuation based on local polynomial approximation *Geophysics* **71** V163–9

Lu Y and Lu W 2009 Edge-preserving polynomial fitting method to suppress random seismic noise *Geophysics* **74** V69–73

Luo Y, Marhoon M, Al-Dossary S and Alfaraj M 2002 Edge-preserving smoothing and applications *Leading Edge* **21** 136–58

Turner G 1990 Aliasing in the tau-p transform and the removal of spatially aliased coherent noise *Geophysics* **55** 1496–503

Ulrych T J, Sacchi M D and Graul J M 1999 Signal and noise separation: art and science *Geophysics* **64** 1648–56

Wang Y 1999 Random noise attenuation using forward-backward linear prediction *J. Sseismic Explor.* **8** 133–42

Yu S P, Cai X and Su Y 1988 Improvement of signal-to-noise ratio of stack section using polynomial fitting of seismic signals (in Chinese) *Oil Geophys. Prospect.* **23** 131–9