Fast dual-domain reduced-rank algorithm for 3D deblending via randomized QR decomposition

Jinkun Cheng and Mauricio D. Sacchi

ABSTRACT

We have developed a fast dual-domain algorithm based on matrix rank reduction for separating simultaneous-source seismic data. Our algorithm operates on 3D common receiver gathers or offset-midpoint gathers. At a given monochromatic frequency slice in the $\omega$-$x$-$y$ domain, the spatial data of the ideal unblended common receiver or offset-midpoint gather could be represented via a low-rank matrix. The interferences from the randomly and closely fired shots increased the rank of the aforementioned matrix. Therefore, we could minimize the misfit between the blended observation and the predicted blended data subject to a low-rank constraint that was applied to the data in the $\omega$-$x$-$y$ domain. The low-rank constraint could be implemented via the classic truncated singular value decomposition (SVD) or via a randomized QR decomposition (rQRd). The rQRd yielded nearly one order of processing time improvement with respect to the truncated SVD. We have also discovered that the rQRd was less stringent on the selection of the rank of the data. The latter was important because we often had no precise knowledge of the optimal rank that was required to represent the data. We adopted a synthetic 3D vertical seismic profile and a real seismic data set acquired at the North Viking Graben to test the performance of the proposed source separation algorithm. The proposed algorithm effectively eliminated the interferences while preserving the desired unblended signal. Especially for the synthetic vertical seismic profile example, experiments were evaluated under different survey time ratios. Our tests indicated that the proposed method could save up to 90% of acquisition time under a self-simultaneous source acquisition scenario.

INTRODUCTION

Simultaneous source acquisition, or blended acquisition, has been proposed to reduce acquisition time by allowing overlapping between different sources (Garotta, 1983; Beasley et al., 1998; Berkhout, 2008; Howe et al., 2008). In the configuration of simultaneous source acquisition, several seismic sources fire at close time intervals. The responses are then recorded by the same set of receivers. We study a special case of marine seismic acquisition in which we assume the detectors are ocean-bottom nodes and seismic shots are fired by one vessel. Because the airgun sources are impulsive, the vessel keeps traveling until it covers the whole survey area without waiting for new records. As a result, only the adjacent sources are blended with a given time delay $\tau$. This acquisition design resembles the self-simultaneous shooting except that we extend the concept to 3D seismic acquisition (Abma et al., 2013). Although our data configuration is not realistic, it allows study of the relationship between fire time delays and the performance of deblending algorithms (Cheng and Sacchi, 2014). A more realistic scenario requires more than one vessel firing at random intervals at different $x$-$y$ locations as described by Moldoveanu et al. (2012).

We assume that the source locations are yielded to a regular grid (Figure 1). The fire time of the $l$th source is defined in the following equation:

$$t_l = t_{l-1} + \tau_l = \sum_{i=1}^{l} \tau_i,$$

where $\tau_i$ is the time delay for the $i$th source. Figure 2 shows a firing schedule for a one-vessel acquisition design (red). In this example, we save 50% of the total acquisition time compared with the conventional seismic acquisition (blue). For a single receiver $j$, we denote the data associated with sources in the spatial positions $x_t, y_t$ by $D_{x_t}(t, x_t, y_t)$. The blended data are represented as follows:
Therefore, incoherent noise re-
acquired blended data. The problem is tackled by adopting a gradient descent iteration plus a step that projects the estimate of the deblended data at any given iteration to a set of low-rank matrices. Because the computational cost of singular value decomposition (SVD) is a major concern for reduced-rank filtering methods, we present a fast and simple rank-reduction algorithm based on random projection (Halko et al., 2011; Chiron et al., 2014). Our synthetic examples simulated with a 3D vertical seismic profile (VSP) data set, and from a real seismic data set, show that the proposed algorithm can suppress the crosstalk generated by simultaneous source acquisition. In addition, we examine the computational cost of rank reduction via truncated SVD (tSVD) and randomized QR decomposition (rQRd). Our research clearly shows the advantage of adopting rQRd as a strategy to improve the computational cost of iterative rank-reduction algorithms.

THEORY

Deblending via rank minimization

We provide a brief review of f-x-y eigenimage filtering that applies matrix rank reduction on each frequency slice of data in the frequency-space domain. The method operates in the offset-midpoint domain, in which the ideal noiseless data can be represented via a superposition of plane waves. For example, only the first $K$ eigenimages are needed to express the desired data for a section contains $K$ dipping events. The latter is referred as the exactness property of eigenimage filtering (Trickett, 2003). Because incoherent noise spreads equally along all the singular values, tSVD that only keeps the $K$ leading singular values is effective for suppressing the incoherent noise (Ulrych et al., 1999).

We apply the eigenimage filtering to common receiver gathers from multidimensional seismic data for separation of simultaneously fired sources. We will remind the reader that the desired data $\mathcal{D}(t, x_l, y_l)$ can be written in terms of its temporal Fourier-domain transform as follows:

$$\mathcal{D}(t, x_l, y_l) = \int \hat{\mathcal{D}}(\omega, x_l, y_l) e^{i \omega t} d\omega, \quad l \in \mathbb{S}. \quad (4)$$

Because we consider a regular distribution of sources in the $x$-$y$ plane, at a given monochromatic frequency $\omega$, one can express $\hat{\mathcal{D}}(\omega, x_l, y_l)$ in terms of a spectral matrix $\mathbf{D}_\omega$ of size $N_x \times N_y$, where the total number of sources is given by $N_S = N_x \times N_y$. The SVD of $\mathbf{D}_\omega$ is given in the following equation:

$$\mathbf{D}_\omega = \mathbf{U} \Sigma \mathbf{V}^H, \quad (5)$$

where $\mathbf{U}$ and $\mathbf{V}$ are matrices with orthonormal columns and $\Sigma$ is a diagonal matrix whose elements $\Sigma_{ii}$ are called the singular values of the matrix (Golub and van Loan, 1996). Figure 3 shows the distribution of singular values in logarithmic scale of the ideal spectral matrix at 20 Hz (blue). We also portray in red the singular value distribution of the spectral matrix from a pseudodeblended common receiver gather at the same frequency. The desired signal that is coherent in the common receiver domain contributes to at most $k$ largest singular values, whereas interferences due to random time delays boost up all the singular values. In other words, the coherence pass constraint for deblending can be implemented by a low-rank constraint applied to the pseudodeblended data. However, in the common-receiver domain, the exactness property of the eigenimage filtering is no longer valid, and the number of singular values $k$ does not equal the number of dipping events. Moreover, because the interferences caused by simultaneous source acquisition are as strong as, or even stronger than, the desired reflections, the one-time rank-reduction strategy fails to remove the crosstalk completely.

To iteratively suppress the simultaneous source interferences, we propose an inversion scheme that minimizes the rank of the aforementioned spectral matrices while honoring the blended acquisition. Combining equations 3 and 4 in the operator form, we have

$$\mathbf{b} = \mathbf{BF}^{-1} \mathcal{D}, \quad (6)$$

![Figure 3. Distribution of singular values of spatial data in (a) normal and (b) logarithmic scale at 20 Hz from the true unblended common receiver gather in the $\omega$-$x$-$y$ domain (red). We also portray the distribution of singular values for data contaminated with source interfaces in a pseudodeblended common receiver gather (blue) and spatial data from the deblended solution (green) at the same frequency.](image-url)
where $F^{-1}$ denotes the inverse temporal Fourier transform that maps data in the frequency domain to the time domain. The forward Fourier transform that transforms data from the time domain to the frequency domain is denoted by $F$. Considering the low-rank constraint for deblending, the desired unblended data $D$ are estimated by

$$\forall \omega: \min \text{rank}(D_{\omega}) \text{ subject to } b = BF^{-1}\hat{D},$$

or by the equivalent form as follows:

$$\min J_1 = \|b - BF^{-1}\hat{D}\|^2_F, \text{ subject to }$$

$$\forall \omega: D_{\omega} \in C(k) = \{D_{\omega}: \text{rank}(D_{\omega}) \leq k\}.$$  

The gradient projection framework

We show that the rank-minimization problem in equation 8 can be tackled via a projected gradient method that is described in Appendix A. The classic gradient-descent algorithm followed by a conventional tSVD is used. The latter entails keeping the largest $k$ singular values of $X_w$, while setting the other singular values to zero. The data $X_w$ are then reconstructed with the new set of singular values.

Algorithm 1. Dual-domain rank-reduction deblending algorithm.

**Inputs**
- Blending operator $B$ and its adjoint $B^*$
- Observed blended trace $b$
- Stopping criterion $\epsilon$
- Step size $\lambda$

**Initialize**
- $D^0 = B^*b; \nu = 1$

**repeat**
- $X = D^{\nu - 1} - \lambda B^*(BD^{\nu - 1} - b)$
- $D^\nu = P_{fr}[X]$ (see Algorithms 2 and 3)
- $\nu = \nu + 1$

**until** $\|b - BD^{\nu}\|^2_F < \epsilon$

$D = D^\nu$

The low-rank constraint is imposed to each frequency slice of the $\omega$-$x$-$y$ data cube. However, the objective function is acquired by minimizing the misfit function between observation and blended estimation in the $t$-$x$-$y$ domain.

A fast low-rank approximation

Finding the low-rank approximation of a given matrix is ubiquitous in the areas of applied mathematics, numerical analysis, and a variety of scientific computing areas (Liberty et al., 2007). The conventional SVD method requires an order of $O(mn^2)$ operations, in which $m$ and $n$ denote the size of a given matrix. Alternative rapid rank-reduction methods, such as Lanczos bidiagonalization and randomized SVD, have recently been applied for seismic data reconstruction and denoising (Gao et al., 2011, 2013; Oropeza and Sacchi, 2011). In this paper, we present a method named rQRd to improve the efficiency of matrix rank reduction.

Algorithm 2. Projection operator for tSVD $P_{fr}[X]$.

**Initialize**
- $\hat{X} \leftarrow X$ (transform to frequency domain)

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

- $[U, \Sigma, V] = \text{svd}(X_w)$
- **if** $i < k$ **then**
  - $\hat{\Sigma}_{i,i} = \Sigma_{i,i}$
  - else
  - $\hat{\Sigma}_{i,i} = 0$
- **end if**
- $\hat{X}(\omega) = U\hat{\Sigma}V^H$

**end for**

$D^\nu \leftarrow \hat{X}$ (transform back to time)
reduction with the power iteration method that leads to a very efficient method for large-scale problems. A comprehensive review of matrix low-rank approximations based on the random projections is given by Halko et al. (2011).

Algorithm 3. Projection operator for rQRd $P_{fr}[\mathbf{X}]$.

**Initialize**

$\tilde{\mathbf{X}} \leftarrow \mathbf{X}$ (transform to frequency domain)

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

$\mathbf{M} = \mathbf{X}_\omega \mathbf{\Omega}$ (random projection)

$[Q, R] = \text{qr}([\mathbf{M}])$

$\hat{\mathbf{X}}_\omega = QQ^H \mathbf{X}_\omega$

**end for**

$\mathcal{P}_f \leftarrow \tilde{\mathbf{X}}$ (transform back to time)

Let us project the spectral matrix $\mathbf{X}_\omega$ by a set of $P$ random normalized vectors given by the columns of the matrix $\mathbf{\Omega}$:

$$
\mathbf{M} = \mathbf{X}_\omega \mathbf{\Omega}.
$$

(10)

Owing to the randomness, the vectors in matrix $\mathbf{M}$ are linearly independent. Because the unblended data in $\mathbf{D}_\omega$ are low rank, only a few $P$ random vectors will be required to span the full range of the desired signal (Halko et al., 2011). As $P \ll N_{Sy}$, the random projection reduces the size of matrix for rank reduction. Then, we compute the orthonormalized basis $Q$ with the economy-size QR decomposition of matrix $\mathbf{M}$ as follows:

$$
Q \mathbf{R} = \mathbf{M}.
$$

(11)

The low-rank approximation is computed via the following expression:

$$
\hat{\mathbf{X}}_\omega = QQ^H \mathbf{X}_\omega.
$$

(12)

The rQRd procedure is repeated for each frequency slice in the $f$-$x$-$y$ common receiver cube, which makes the rQRd projection operator shown in Algorithm 3. Unlike the tSVD method that directly solves for the closest low-rank approximation of a given matrix, the random projection methods do not constrain the rank strongly as the SVD does. In other words, in rQRd, a subset size equal to the exact rank of the given matrix usually cannot ensure that the solution is a rank $k$ approximation. We consider the size of the random subset $P$ as a relaxation of the exact rank $k$ of the matrix (Chiron et al., 2014). This usually allows us to achieve better results when the singular values do not decay dramatically and when the precise rank of the matrix is not known. Figure 4 shows the comparison of processing time for matrix rank reduction using tSVD (blue) and rQRd (red). In this example, we choose $P$ equal to $3k$, and the rQRd algorithm is approximately 10 times faster than the conventional tSVD.

![Figure 4. The processing time of rank reduction versus the size of matrix both in logarithmic scale. The blue curve shows the processing time using the truncated SVD, whereas the red curve shows the processing time of the rQRd method. We choose $P$ equal to $3k$, and the rQRd algorithm is approximately 10 times faster than the conventional tSVD.](image)

Table 1. Comparison of computational efficiency for two reduced-rank projections: tSVD and rQRd. The value $Q_s$ is the quality of deblending defined by equation 14. The rms error denotes the root-mean-square error of the blended estimation and observation in the time domain. ITER and TIME record the number of iterations and total time that is used for the deblending algorithm. The rQRd outperforms tSVD in accuracy and efficiency in terms of deblending.

<table>
<thead>
<tr>
<th>Size</th>
<th>tSVD</th>
<th>rQRd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Q_s$ (dB)</td>
<td>rms error</td>
</tr>
<tr>
<td>25 x 25</td>
<td>19.3</td>
<td>1.52</td>
</tr>
<tr>
<td>50 x 50</td>
<td>22.5</td>
<td>2.06</td>
</tr>
<tr>
<td>100 x 100</td>
<td>24.6</td>
<td>2.71</td>
</tr>
<tr>
<td>150 x 150</td>
<td>25.3</td>
<td>3.18</td>
</tr>
<tr>
<td>200 x 200</td>
<td>26.6</td>
<td>3.28</td>
</tr>
</tbody>
</table>
EXAMPLE

Example with a synthetic 3D VSP data set

We use a synthetic 3D VSP data set to mimic the process of simultaneous source acquisition. The data set contains 205 source lines with 205 source positions on each line (Brien, 2010). The interval of each source position is 16.67 m, and the line spacing is also 16.67 m (Figure 1). In this example, 31 downhole detectors are deployed at depths of 1350–1850 m with a 16.67 m interval in the center of this x-y grid. The highest frequency that contains useful information is approximately 40 Hz because a 15.4 Hz Ricker is used to generate data. The sources are blended corresponding to the one-vessel scenario described in the introduction of this paper. We measure the efficiency of this acquisition in terms of the survey time ratio (STR) (Berkhout, 2008), which is defined as follows:

$$\text{STR} = \frac{\tau_0}{\bar{\tau}},$$  \hspace{1cm} (13)

where $\tau_0$ denotes the regular fire time interval for the conventional seismic acquisition and $\tau$ is the expectation of time delays for all the blended shots. For instance, in Figure 2, the STR equals two, and the acquisition time with blended sources is 50% of the conventional acquisition. We then separate the numerically blended data via the proposed algorithm. The quality of deblending is measured in dB units via the following equation:

$$Q_s = 10 \log \frac{||d_{\text{true}}||_2^2}{||d_{\text{true}} - d_S||_2^2},$$  \hspace{1cm} (14)

where $d_{\text{true}}$ is the true synthetic data from a conventional common receiver gather and $d_S$ stands for the separated common receiver gather via iterative rank reduction. A large $Q_s$ corresponds to fewer simultaneous source interferences in the deblended results.

Performance of the algorithms

We tested the effectiveness of the deblending algorithm under a different rank (tSVD) and a different subset size $P$ in the rQRd algorithm. In this example, we fixed the firing time delay with STR being equal to 10. Figure 5 shows the quality of the deblending versus rank. It can be observed that the tSVD method (blue) presents the highest quality after deblending only when the selected rank is very close to the exact rank $k$ of the data. In other words, prior information would be required to choose the optimal rank. The rQRd method (red), however, exhibits a broad region of high deblending quality. We can achieve reasonable results when the subset size is in the range of $P \in [1.5k, 5k]$. The test provides evidence that the subset size $P$ in the rQRd algorithm is a relaxation of the desired rank $k$. Table 1 shows the comparison of computational performance between rQRd and tSVD for matrices of different sizes. We tuned the parameters to ensure that the algorithm converges to a similar point. In our case of deblending, the rQRd method outperforms the conventional tSVD in accuracy and efficiency.

Effect of rank and STR

It is important to stress that unlike the Cadzow deblending method and other rank-reduction deblending methods that operate in an offset-midpoint domain (Maraschini et al., 2012; Wason et al., 2014), in which the rank is equivalent to the number of plane waves, in the common receiver domain there is no theoretical justification for selecting an optimal rank. We choose to use the rQRd method to test the effects of different rank and the STR on the proposed deblending algorithm.

In this experiment, we used an STR up to 25 and then tested a wide selection of subset sizes from 1 to 120. For each specific rank and STR, we generated 50 realizations with the same distribution for setting the firing schedule. For each trial, we adopted a relatively small step size to ensure the convergence of the algorithm. When the quality factor after separation $Q_s$ reaches a number higher than 20 dB, we consider the proposed algorithm successfully separated
Figure 7. Results of simultaneous source separation in the common-receiver domain when the STR equals two: (a) the real unblended time slice at 1.2 s, (b) pseudodeblended time slice, (c) deblended time slice after 25 iterations of the proposed algorithm, and (d) differences between panels (a and c). In this example, the quality factor after separation is 36.5 dB.

Figure 8. Results of simultaneous source separation in the common-receiver domain when the STR equals two: (a) the real unblended common receiver gather (center receiver), (b) pseudodeblended common receiver gather, (c) deblended common receiver gather after 25 iterations, and (d) differences between panels (a and c).

Figure 9. Results of simultaneous source separation in the common-shot domain when the STR equals two: (a) the real unblended common shot gather (center shot), (b) pseudodeblended shot record, (c) deblended shot record after 25 iterations, and (d) differences between panels (a and c).
Figure 10. Results of simultaneous source separation in the common-receiver domain when the STR equals 10: (a) the real unblended time slice at 1.2 s, (b) pseudodeblended time slice, (c) deblended time slice via the proposed algorithm, and (d) differences between panels (a and c). In this example, the quality factor after separation is 28.4 dB.

Figure 11. Results of simultaneous source separation in the common-receiver domain when the STR equals 10: (a) the real unblended common receiver gather (center receiver), (b) pseudodeblended common receiver gather, (c) deblended common receiver gather via the proposed algorithm, and (d) differences between panels (a and c).

Figure 12. Results of simultaneous source separation in the common-shot domain when the STR equals 10: (a) the real unblended common shot gather (center shot), (b) pseudodeblended shot record, (c) deblended shot record via the proposed algorithm, and (d) differences between panels (a and c).
Figure 13. Results of simultaneous source separation in the common-receiver domain when the STR equals 20: (a) the real unblended time slice at 1.2 s, (b) pseudodeblended time slice, (c) deblended time slice via the proposed algorithm, and (d) differences between panels (a and c). In this example, the quality factor after separation is 20.2 dB.

Figure 14. Results of simultaneous source separation in the common-receiver domain when the STR equals 20: (a) the real unblended common receiver gather (center receiver), (b) pseudodeblended common receiver gather, (c) deblended common receiver gather via the proposed algorithm, (d) differences between panels (a and c).

Figure 15. Results of simultaneous source separation in the common-shot domain when the STR equals 20: (a) the real unblended common shot gather (center shot), (b) pseudodeblended shot record, (c) deblended shot record via the proposed algorithm, and (d) differences between panels (a and c).
the responses from the blended sources. The definition of success can also be interpreted as the percentage error as follows:

$$\frac{\|d_{\text{true}}\|_2^2}{\|d_{\text{true}} - d_S\|_2^2} \leq 0.01.$$  \hspace{1cm} (15)

In other words, 1% relative mean-squared-error is our threshold for a successful run. Figure 6, which is very similar to the Tanner-Donoho plot (Donoho and Tanner, 2009), shows the percentage of successful runs in terms of rank and STRs. At a given STR and rank, the white color means all 50 runs of the algorithm successfully removed the interferences. In contrast, a point in the dark area indicates a combination of rank and STR for which all the trials of the deblending algorithm failed to improve the data quality to 20 dB. It is clear from the figure that for a relatively small STR, a broad range of rank could be adopted to ensure successful separations. However, as the STR grows, we need to select the optimal rank for simultaneous source separation. In addition, there always exists a limitation of the blended acquisition in terms of the firing scheme. For this specific model and acquisition design, the algorithm fails to separate sources when the STR is greater than 21.

We show the deblending results for the rQRd method as well. In this case, we used different STRs and values of the parameter $P$. The proposed algorithm has eliminated the noise and reestablished a distribution of singular values similar to the distribution of singular values of the ideal data (Figure 3). In the first example (Figures 7–9), the STR equals two. Figure 7 shows a time slice of the results after 25 iterations. Figure 8 shows the deblending result for the center receiver, and Figure 9 shows the result for the center shot. The interferences from simultaneously fired shots are effectively suppressed. We improve the quality factor of the pseudodeblended data set to 36.5 dB. As a result, the unblended solution becomes comparable with the true shot record. Then, the STR continues to increase to a factor of 10. Figures 10–12 show the deblending results in different gathers. The deblending algorithm is still robust with the quality of deblending improved to 28.4 dB. When the STR reaches a factor of 20 (Figures 13–15),

Figure 16. Results of simultaneous source separation of a 2D marine line: (a) the real unblended CMP gather, (b) CMP gather sorted from pseudodeblended data, (c) deblended results via the proposed algorithm, and (d) differences between panels (a and c). In this example, the quality factor after separation is 14.2 dB.
we start to see the leave-out energy in the difference panels and the quality of deblending reduces to 20.2 dB.

Example simulated from a real data set

We adopted a 2D marine seismic data set from the North Viking Graben, North Sea to test the efficacy of the proposed source separation method. The sources and receivers are sampled with an interval of 25 m (Keys and Foster, 1998). The data are then numerically blended assuming the receivers are ocean-bottom nodes. The sources are fired according to the self-simultaneous source acquisition design, and the STR equals two. We then applied the dual-domain rank-reduction algorithm to separate the blended data. The blending and pseudo-deblending operators are applied in the source-receiver domain, whereas rank reduction is applied in offset-midpoint domain (Appendix B). Figure 16 shows the deblending results in the CMP domain. The interferences are effectively suppressed by the proposed algorithm. Figure 17 exhibits a near-offset gather. The quality of data $Q_S$ has been improved to a factor of 14.2 dB.

CONCLUSIONS

This paper illustrates an inversion scheme that can be used for separation of simultaneous source data. The method relies on the randomization of fire time delays and operates in the common receiver domain of a multidimensional seismic data set. The cost function is defined by the source blending system. As the interferences perturbed by firing-time delays increase the rank of each frequency slice of a common receiver cube, a low-rank constraint is imposed to enforce the coherence of solution. The gradient projection algorithm has been adopted for solving the rank-constrained inverse problem. We presented a fast rank-reduction algorithm based on random projection and rQRd to improve processing speed.

Through a test with a synthetic 3D VSP data set, we showed that the algorithm effectively separated the responses from simultaneous sources. We also tested the performance of the algorithm versus rank and STR. For low-STR values, a broad range of rank would ensure the success of the deblending algorithms. At high-STR values, prior knowledge of the model would be required for selecting the optimal rank.

ACKNOWLEDGMENTS

The authors wish to thank the sponsors of the Signal Analysis and Imaging Group at the University of Alberta. This research was also supported by The Natural Sciences and Engineering Research Council of Canada via a grant to M. D. Sacchi. The authors acknowledge M. O’Brien and the Reservoir Characterization Project for providing the 3D-9C VSP data. The authors also thank G. Drijkoningen, G. Blacquiere, R. van Borselen, and two anonymous reviewers for their constructive suggestions that improved this paper.

APPENDIX A

DERIVATIONS OF DUAL-DOMAIN RANK-REDUCTION ALGORITHM

We show that the rank-minimization problem in equation 8 can be tackled via the gradient projection method. We successively update a current estimate $\tilde{D}^\nu$ to minimize the objective function. This is typically performed by modifying $\tilde{D}^\nu$ in the opposite direction of the gradient as follows:

$$\tilde{D}^{\nu+1} = \tilde{D}^\nu - \lambda \partial J_1,$$

where the gradient can be calculated as follows:

$$\partial J_1 = F B^\nu (B F^{-1} \tilde{D}^{-1} - b),$$
where $\lambda$ denotes the step size that is selected to ensure a decrease of the misfit in each equation. The optimal solution to equation 8 is also the optimal solution to the following cost function:

$$J_2 = \| \mathbf{D} - \mathbf{D}' \|^2_F \text{ subject to }$$

$$\forall \omega \in \mathbf{D}_w \subseteq C(k) = \{ \mathbf{D}_w : \text{rank}(\mathbf{D}_w) \leq k \}.$$  

(A-3)

only if $\mathbf{D}'$ converges to the optimal solution of $J_1$ (Ji and Ye, 2009; Cai et al., 2010; Ma et al., 2011). At a given temporal frequency $\omega$, equation A-3 reduces to:

$$J(\omega) = \| \mathbf{D}_w - \mathbf{D}'_w \|^2_F \text{ subject to }$$

$$\forall \omega \in \mathbf{D}_w \subseteq C(k) = \{ \mathbf{D}_w : \text{rank}(\mathbf{D}_w) \leq k \}.$$  

(A-4)

Equation A-4 entails finding a low-rank approximation of $\mathbf{D}_w$ at a given frequency $\omega$. The classic solution is the well-studied tSVD. If we use $\mathcal{P}_r$ to demonstrate a projection operator that projects each frequency slice of data to a low-rank matrix, then the solution for separating the simultaneous sources can be expressed as follows:

$$\mathbf{D}' = \mathcal{P}_r[\mathbf{D}^{-1} - \lambda \mathcal{F}^* (\mathcal{F} \mathbf{F}^{-1} \mathbf{D}^{-1} - \mathbf{b})]. \quad \text{ (A-5)}$$

This gradient projection algorithm entails searching for a solution in the gradient descent direction. The solution is then projected to a set of low-rank matrices. In practice, it is more convenient to adapt the above algorithm to the following form:

$$\mathbf{X} = \mathbf{D}'^{-1} - \lambda \mathcal{F}^* (\mathcal{F} \mathbf{F}^{-1} \mathbf{D}'^{-1} - \mathbf{b})$$

$$\mathbf{D}' = \mathbf{F}^{-1} \mathcal{P}_r \mathcal{F} \mathbf{X} = \mathcal{P}_r \mathbf{X}. \quad \text{ (A-6)}$$

We remind the reader that $\mathbf{D}'$ and, therefore, $\mathbf{X}$ are deblended data at iteration $\nu$ in the t-x-y domain. However, the rank-reduction constraint $\mathcal{P}_r$ must be applied in the f-x-y domain.

APPENDIX B

DEBLENDING IN OFFSET-MIDPOINT DOMAIN

The proposed iterative rank-reduction method can also be applied to deblending 2D seismic data sets. However, for 2D seismic data sets, the frequency slices of the common receiver gathers are vectors. Matrix rank-reduction methods can no longer be applied directly to the data. The key is to sort data from the source-receiver domain to the offset-midpoint domain before applying rank reduction in each iteration. This is because in common-offset and CMP domains, the interferences from simultaneously fired shots are perturbed by random fire times. The coherence constraint for deblending holds in a 3D offset-midpoint gather. In the meantime, as discussed in the theory section, f-x-y eigenimage filtering is an effective tool for suppressing incoherent noise in the offset-midpoint domain (Trickett, 2003). The projected gradient step that operates on 2D seismic data sets for simultaneous source separation is as follows:

$$\mathbf{D}' = \mathcal{P}_r[\mathbf{D}^{-1} - \lambda \mathbf{S}^* \mathcal{F}^* (\mathcal{F} \mathbf{F}^{-1} \mathbf{S} \mathbf{D}'^{-1} - \mathbf{b})], \quad \text{ (B-1)}$$

where $\mathbf{S}$ entails sorting data from the source-receiver domain to the offset-midpoint domain. The $\mathbf{S}^*$ is the adjoint operator that sorts data back from offset-midpoint gathers to source-receiver domain. It is more convenient to move the Fourier transform and sorting operator to the projection operator as follows:

$$\mathbf{X} = \mathbf{D}'^{-1} - \lambda \mathbf{S}^* (\mathcal{F} \mathbf{D}'^{-1} - \mathbf{b})$$

$$\mathbf{D}' = \mathbf{S}^* \mathbf{F}^{-1} \mathcal{P}_r \mathcal{F} \mathbf{X} = \mathcal{P}_n \mathbf{X}. \quad \text{ (B-2)}$$

In each iteration, the blending and pseudo-deblending operators are applied to data in the source-receiver domain, whereas the f-x-y eigenimage filtering operates on each frequency slice of the offset-midpoint data cube.

REFERENCES


