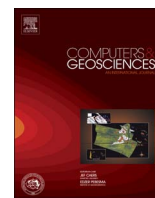




ELSEVIER

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

An open-source Matlab code package for improved rank-reduction 3D seismic data denoising and reconstruction

Yangkang Chen^a, Weilin Huang^b, Dong Zhang^b, Wei Chen^{c,d,*}

^a Bureau of Economic Geology, John A. and Katherine G. Jackson School of Geosciences, The University of Texas at Austin, University Station, Box X, Austin, TX 78713-8924, United States

^b State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Fuxue Road 18th, Beijing 102200, China

^c Key Laboratory of Exploration Technology for Oil and Gas Resources of Ministry of Education, Yangtze University, Wuhan Hubei, 430100, China

^d Hubei Cooperative Innovation Center of Unconventional Oil and Gas, Wuhan Hubei 430100, China

ARTICLE INFO

Article history:

Received 4 March 2016

Received in revised form

8 June 2016

Accepted 23 June 2016

Available online 25 June 2016

Keywords:

Open-source Matlab code

Rank reduction

3D seismic data

Reconstruction

Denoising

ABSTRACT

Simultaneous seismic data denoising and reconstruction is a currently popular research subject in modern reflection seismology. Traditional rank-reduction based 3D seismic data denoising and reconstruction algorithm will cause strong residual noise in the reconstructed data and thus affect the following processing and interpretation tasks. In this paper, we propose an improved rank-reduction method by modifying the truncated singular value decomposition (TSVD) formula used in the traditional method. The proposed approach can help us obtain nearly perfect reconstruction performance even in the case of low signal-to-noise ratio (SNR). The proposed algorithm is tested via one synthetic and field data examples. Considering that seismic data interpolation and denoising source packages are seldom in the public domain, we also provide a program template for the rank-reduction based simultaneous denoising and reconstruction algorithm by providing an open-source Matlab package.

Published by Elsevier Ltd.

1. Introduction

Seismic data interpolation plays a fundamental role in seismic data processing, which provides the regularly sampled seismic data for the following workflows like high-resolution processing, wave-equation migration, multiple suppression, amplitude-versus-offset (AVO) or amplitude-versus-azimuth (AVAZ) analysis, and time-lapse studies (Chiu, 2014; Chen et al., 2015; Wang et al., 2015; Gan et al., 2015a; Zhong et al., 2016; Zhang et al., 2016). Three main types of interpolation approaches have been proposed in the literature for handling with the interpolation problem. The first type of approach is based on prediction (Spitz, 1991; Porsani, 1999). A prediction operator is designed from the low-frequency components of seismic data and is applied to high-frequency components. However, the predictive filtering method can only be applied to regularly sampled seismic data. The second type is a transformed domain method (Candès et al., 2006a; Chen et al., 2014b; Gan et al., 2015b; Liu et al., 2015), which is based on compressive sensing theory (Candès et al., 2006b; Donoho, 2006) to achieve a successful recovery using highly incomplete available

data (Sacchi et al., 1998; Wang, 2003; Chen et al., 2014a). Compressive sensing (CS) is a relatively new paradigm (Candès et al., 2006b; Donoho, 2006; Donoho et al., 2006; Gan et al., 2016; Liu et al., 2016) in signal processing that has recently received a lot of attention. The theory indicates that the signal which is sparse under some basis may still be recovered even though the number of measurements is insufficient as required by the Shannon's criterion. Naghizadeh and Sacchi (2007) propose a multistep autoregressive strategy which combines the first two types of methods to reconstruct irregular seismic data. The third type is based on the wave equation. This type of method utilizes the inherent constraint of the seismic data from wave equation to interpolate seismic data, thus it depends on the known velocity model, which also becomes its limitation (Canning and Gardner, 1996; Fomel, 2003). Recently, an increasing number of researchers have developed algorithms connecting the interpolation and deblending (Berkhout, 2008) problems for the irregular sampled simultaneous-source data (Li et al., 2013), which provides new recipes for conventional seismic interpolation problem.

Rank-reduction based seismic interpolation algorithms become popular (Trickett and Burroughs, 2009; Oropeza and Sacchi, 2011) in the past decade. The rank-reduction methods for seismic data reconstruction can be divided into two main categories. The first category applies rank reduction to multilevel block Hankel/

* Corresponding author.

E-mail addresses: ykchen@utexas.edu (Y. Chen), cup_hwl@126.com (W. Huang), zhangdongconan@163.com (D. Zhang), chenwei2014@yangtzeu.edu.cn (W. Chen).

Toeplitz matrices formed from the entries of the tensor. In other words, multidimensional seismic data are rearranged into a block Hankel or Toeplitz matrix, and a rank reduction algorithm is used to improve the signal-to-noise ratio (SNR) and to reconstruct the data. Such algorithm is usually named as the Cadzow method (Trickett and Burroughs, 2009) or multichannel singular spectrum analysis (MSSA) method (Oropeza and Sacchi, 2011). The other category of rank-reduction methods encompasses techniques that are based on dimensionality reduction of multi-linear arrays or tensors. In this case, the multichannel seismic data is viewed as a multi-linear array, and dimensionality reduction techniques are directly applied to the multi-linear array (Gao et al., 2015). For example, Kreimer and Sacchi (2012) adopt the high-order singular value decomposition (HOSVD) to solve the 5D seismic data reconstruction problem in the frequency-space domain.

In this paper, we focus on simultaneous reconstruction and denoising of 3D seismic data using the MSSA algorithm. MSSA is a data-driven algorithm developed from research on alternative tools for the analysis of multichannel time series, which is based on the truncated singular value decomposition (TSVD) (Golub and Loan, 1996) of the Hankel matrix. MSSA is also an extension of singular spectrum analysis (SSA) (Vautard and Ghil, 1989), which is used to analyze 1D time series. Missing traces and random noise increase the rank of the appropriately constructed Hankel matrix that is composed of seismic data at a given frequency slice. To some extent, missing data in each frequency slice performs like random noise at the first iteration of weighted projection onto convex sets (POCS) like framework. Similar to other simultaneous seismic data reconstruction and denoising approaches, MSSA transforms the noisy data with missing traces into a domain where signal and noise are mapped onto separate subspaces and then removes the noise. Weighted POCS-like method, which is widely used for seismic data reconstruction and first introduced to the community of seismic exploration by Abma and Kabir (2006), is in charge of the reconstruction procedure. Many numerical experiments, however, indicate that the random noise cannot be completely removed using the conventional MSSA algorithm and there still exist some reconstruction errors. The main reason is that the traditional TSVD can only decompose the data into a noise subspace and a signal-plus-noise subspace. Huang et al. (2016) suggest using damped MSSA (DMSSA) algorithm to better decompose the data into the signal subspace and noise subspace for random noise attenuation. In order to overcome the defect mentioned above, we extend the DMSSA algorithm further to simultaneous reconstruction and denoising of 3D seismic data. We first review the theory of traditional MSSA algorithm, then we introduce the proposed DMSSA based reconstruction and denoising framework. Next, we introduce the main components in the Matlab package and correlate the interior functions with the mathematical symbols in the theory sections. Finally, one synthetic and one field data examples are used to demonstrate the performance using the proposed algorithm and the presented open-source Matlab package. The first synthetic example is reproducible from the provided Matlab package. The field data example is, however, not in the public domain, therefore we do not provide the *segy* file of the field data in the package. But the field data example is just a straightforward application of the Matlab package.

2. Rank reduction via MSSA

Consider a block of 3D data $\mathbf{D}_{time}(t, x, y)$ of N_t by N_x by N_y samples ($t = 1 \dots N_t$, $x = 1 \dots N_x$, $y = 1 \dots N_y$). The MSSA (Oropeza and Sacchi, 2011) operates on the data in the following way: first, MSSA transforms $\mathbf{D}_{time}(t, x, y)$ into $\mathbf{D}_{freq}(w, x, y)$ ($w = 1 \dots N_w$) of complex values of the frequency domain. Each frequency slice of

the data, at a given frequency w_0 , can be represented by the following matrix:

$$\mathbf{D}(w_0) = \begin{pmatrix} D(1, 1) & D(1, 2) & \dots & D(1, N_x) \\ D(2, 1) & D(2, 2) & \dots & D(2, N_x) \\ \vdots & \vdots & \ddots & \vdots \\ D(N_y, 1) & D(N_y, 2) & \dots & D(N_y, N_x) \end{pmatrix}. \quad (1)$$

To avoid notational clutter we omit the argument w_0 . Second, MSSA constructs a Hankel matrix for each row of \mathbf{D} ; the Hankel matrix \mathbf{R}_i for row i of \mathbf{D} is as follows:

$$\mathbf{R}_i = \begin{pmatrix} D(i, 1) & D(i, 2) & \dots & D(i, m) \\ D(i, 2) & D(i, 3) & \dots & D(i, m+1) \\ \vdots & \vdots & \ddots & \vdots \\ D(i, N_x - m + 1) & D(i, N_x - m + 2) & \dots & D(i, N_x) \end{pmatrix}. \quad (2)$$

Then MSSA constructs a block Hankel matrix \mathbf{M} for \mathbf{R}_i :

$$\mathbf{M} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \dots & \mathbf{R}_n \\ \mathbf{R}_2 & \mathbf{R}_3 & \dots & \mathbf{R}_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{N_y-n+1} & \mathbf{R}_{N_y-n+2} & \dots & \mathbf{R}_{N_y} \end{pmatrix}. \quad (3)$$

The size of \mathbf{M} is $I \times J$, $I = (N_x - m + 1)(N_y - n + 1)$, $J = mn$. m and n are predefined integers chosen such that the matrix \mathbf{R}_i and the block Hankel matrix \mathbf{M} are close to square matrices. The transformation of the data matrix into a block Hankel matrix can be represented in operator notation as follows:

$$\mathbf{M} = \mathcal{H}\mathbf{D}, \quad (4)$$

where \mathcal{H} denotes the Hankelization operator.

In general, the block Hankel matrix \mathbf{M} can be represented as:

$$\mathbf{M} = \mathbf{S} + \mathbf{N}, \quad (5)$$

where \mathbf{S} and \mathbf{N} denote the block Hankel matrix of signal and of random noise, respectively.

We assume that \mathbf{M} and \mathbf{N} have full rank, $rank(\mathbf{M}) = rank(\mathbf{N}) = J$ and \mathbf{S} has deficient rank, $rank(\mathbf{S}) = N < J$. The singular value decomposition (SVD) of \mathbf{M} can be represented as:

$$\mathbf{M} = [\mathbf{U}_1^M \quad \mathbf{U}_2^M] \begin{bmatrix} \Sigma_1^M & \mathbf{0} \\ \mathbf{0} & \Sigma_2^M \end{bmatrix} \begin{bmatrix} (\mathbf{V}_1^M)^H \\ (\mathbf{V}_2^M)^H \end{bmatrix}, \quad (6)$$

where Σ_1^M ($N \times N$) and Σ_2^M ($(J - N) \times (J - N)$) are diagonal matrices and contain, respectively, larger singular values and smaller singular values. \mathbf{U}_1^M ($I \times N$), \mathbf{U}_2^M ($I \times (I - N)$), \mathbf{V}_1^M ($J \times N$) and \mathbf{V}_2^M ($J \times (J - N)$) denote the associated matrices with singular vectors. The symbol $[\cdot]^H$ denotes the conjugate transpose of a matrix. In general the signal is more energy-concentrated and correlative than the random noise. Thus, the larger singular values and their associated singular vectors represent the signal, while the smaller values and their associated singular vectors represent the random noise. We let Σ_2^M be $\mathbf{0}$ to achieve the goal of attenuating random noise while recovering the missing data during the first iteration in reconstruction process as follows:

$$\tilde{\mathbf{M}} = \mathbf{U}_1^M \Sigma_1^M (\mathbf{V}_1^M)^H. \quad (7)$$

Eq. (7) is referred to as the TSVD, which is used in the conventional MSSA approach.

3. 3D seismic data reconstruction via DMSSA

Nevertheless, $\tilde{\mathbf{M}}$ is actually still contaminated with residual

Download English Version:

<https://daneshyari.com/en/article/506789>

Download Persian Version:

<https://daneshyari.com/article/506789>

[Daneshyari.com](https://daneshyari.com)