Journal of Applied Geophysics 111 (2014) 95–101

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Factorized Hankel optimal singular spectral approach for erratic and noisy seismic signal denoising

R.K. Tiwari, R. Rajesh *

CSIR-NGRI, Hyderabad, India
AcSIR-NGRI, Hyderabad, India

ARTICLE INFO

Article history:
Received 12 April 2014
Accepted 29 September 2014
Available online 5 October 2014

Keywords:
Seismic reflection
Noise
Hankel matrix
Coal bed

ABSTRACT

De-noising erratic and noisy seismic records is one of the crucial issues in seismic data processing. The presence of colored noise significantly distorts the real signal amplitude leading to misleading results. We present here, a new algorithm based on the Factorized Hankel Optimization method, which is robust in dealing with the above problem. The method involves following two steps: (i) factorization of Hankel matrix and (ii) transformation of trajectory matrix to optimize the noise in singular spectral domain. Initially, we tested the performance of the new algorithm on complex synthetic data and then applied to the real seismic reflection data recorded from the Singareni coalfield, India. The underlying scheme is fast and effective in denoising the complex erratic–noisy seismic signals. We have also compared the robustness of factorization and optimization techniques with respect to conventional singular spectral analysis (SSA). The coal beds and faulted structures identified in the actual seismic reflection data using the new algorithm correlate well with the regional geological structures.

1. Introduction

Denoising field seismic records is critical to seismic data processing. Suppression of random and coherent seismic noise in frequency or wave number domain has been practiced inquisitively for the long time. The dynamics of the pure signal can be properly identified from the random noise background using Eigen estimates (vectors and values) of the data. However, the colored noise in the data significantly deviates the variance of Eigen components (Allen and Smith, 1997).

Singular spectral analysis (SSA) is one of the efficient time series analyzing tools, which has invariably been employed to various kinds of complex geophysical and astronomical problems (Golyandina et al., 2001; Vautard et al., 1992). The proficiency of the methods has been vigorously exploited for principle component analysis (Serita et al., 2005; Tiwari and Rajesh, 2014, Tiwari et al., 2014), reconstruction of missing data and noise suppression in frequency domain (Oropeza and Sacchi, 2011) and prediction of evolutionary time series. The underlying method uses the singular value decomposition scheme for computing eigenvalues and eigenvectors from the trajectory or Hankel matrix of the data. The computational cost of the singular value decomposition (SVD) scheme for lengthy data sets in the form of large Hankel matrix has become one of the major problems that limit the applicability of SSA technique.

Recently some researchers (Rahim and Zokaei, 2011; Xu and Qiao, 2008) have devised a method to reduce the floating point operations (FLOPS) in SVD computation of square Hankel matrix and thereby attempted to reduce the computational cost (compared to the classical $O(n^3)$ method). We developed here an alternative "skilled algorithm" based on factorization of Hankel matrix along the diagonal to reduce the floating point operations in SVD process by exploiting the symmetry property of Hankel matrix. In the proposed scheme the underlying process reduces the FLOPS from $O\ (m \times n^2)$ to $d \times O\ ((m/d) \times ((n + d)/d)^{\ 2})$. Here $m \times n$ is the size of the Hankel matrix ($m < n$) and d is the number of factor segments.

2. Methodology

In geophysical field observation, the noise is an additive component to pure signal. More generally the field data can be represented as follows.

$$\text{Data}\,(D) = \text{Signal}\,(S) + \text{Noise}\,(N). \tag{1}$$

In a mathematical notation the expectation value of the data can be written as follows

$$E(\mathbf{T_D}) = E(\mathbf{T_S}) + E(\mathbf{T_N}) \tag{2}$$

where, $\mathbf{T_D}$, $\mathbf{T_S}$ and $\mathbf{T_N}$ represent respectively trajectory matrix of the observed data, pure signal and noise.

* Corresponding author at: Room number 125, CSIR-NGRI, Uppal Road, Hyderabad 500007, Andhra Pradesh, India. Tel.: +91 9912545974.
E-mail address: rekapalli@gmail.com (R. Rajesh).

http://dx.doi.org/10.1016/j.jappgeo.2014.09.019
0926-9851/

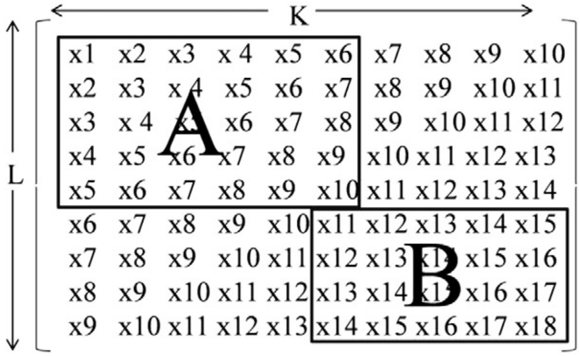

Fig. 1. An example of two step factorization of Hankel matrix of size 9 × 10 into factor matrices A (size 5 × 6) and B (size 4 × 5).

For nonrandom noise process,

$$E(\mathbf{T_N}) \neq \sigma^2 I. \tag{3}$$

The random noise perturbs the real expectation values of the signal only along the diagonal which can be annulled by scaling. However, scaling doesn't work for the estimation of the expectation values of pure signal in the presence of colored noise. The estimate of the Eigen vectors and Eigen values of the observed data matrix (with colored noise) significantly differs from that of the pure signal. The transformation of the observed data trajectory matrix using the Eigen values and Eigen vectors of noise trajectory matrix would allow us to distinguish the actual signal (Allen and Smith, 1997). A fast and robust algorithm based on factorization Hankel matrix and optimization process would help effectively to reduce the erratic and colored noise from seismic reflection data.

Our methodology comprises the following five steps namely (i) formulation and factorization of trajectory matrix (ii) transformation of each factor using the Eigen vectors and values of estimated noise trajectory matrix (iii) SVD, grouping and reconstruction of each factor of trajectory matrix (iv) inverse transformation of all the reconstructed trajectory matrices and (v) diagonal averaging of each trajectory matrix and reconstruction of denoised data series. These steps are explained mathematically in detail as follows:

Step (i) We begin with the formulation of trajectory matrix of size $L \times (N - L + 1)$ from the observed data series $Y(x) = \{y(x_1), y(x_2), y(x_N)\}$ using an appropriate window length (L) using

$$T_{L\times K} = [X_{1:}\ldots : X_K] \tag{4}$$

where X_i indicates the vector of length L, where $K = N - L + 1$, and N represents the length of the series. The transformed trajectory Hankel matrix is systematically factorized into number of square or rectangular matrices along the diagonal as shown in Fig. 1.

Step (ii) Following Allen and Smith, 1997, we define signal to noise ratio or more precisely signal to noise variance ratio (ρ) in terms of data and noise trajectory matrices as $\rho \equiv \frac{\mathbf{e^T}\cdot\mathbf{T_D}\cdot\mathbf{e}}{\mathbf{e^T}\cdot\mathbf{T_N}\cdot\mathbf{e}}$, where $\boldsymbol{e}$ is the vector and define the state space direction and $\boldsymbol{e^T}$ is its transpose. In the data space, the dynamical behavior of the system can be understood from the eigenvectors which are also known as state space vector and define the direction of the data space. From a predefined estimate of ρ, the noise trajectory matrix $\mathbf{T_N}$ can be obtained using the eigenvectors of the observed data trajectory matrix. Next using the Eigen vector and Eigen values of the noise trajectory matrix, the observed data can be transformed using the following equation

$$[\mathbf{U}, \boldsymbol{\lambda}, \mathbf{V}] = \mathrm{SVD}(\mathrm{COV}(\mathbf{T_N})) \tag{5}$$

where **U** and **V** are respectively the left and right eigenvectors and **λ** is the eigenvalue matrix of the covariance (COV) trajectory matrix. The transformed trajectory matrix, say $\mathbf{T_D}'$ can be given by

$$\mathbf{T_D'} = \boldsymbol{\lambda}^{1/2}\cdot\mathbf{U^T}\cdot\mathbf{T_D}\cdot\mathbf{U}\cdot\boldsymbol{\lambda}^{-1/2}. \tag{6}$$

Step (iii) Then we apply SVD, grouping and reconstruction processes on each factor of the transformed trajectory matrix.

a) Singular value decomposition (SVD): Here, we decompose the factors of the transformed trajectory matrix ($\mathbf{T_D}'$) into eigenvectors and eigenvalues as follows

$$\mathrm{SVD}(\mathbf{T_D'}) = \left[\mathbf{U}_i\sqrt{\boldsymbol{\lambda}_i}\mathbf{V}_i\right]. \tag{7}$$

The group of ith eigenvectors and eigenvalue i.e., ($\sqrt{\boldsymbol{\lambda}_i}$, $\mathbf{U_i}$, $\mathbf{V}_i$) is called the eigentriple.

Hence p eigentriples for non-zero eigenvalues, the transformed factor trajectory matrix $\mathbf{T_D}'$ can be given by (Golyandina and Zhigljavsky, 2013)

$$\mathbf{T_D'} = \sum_{j=1}^{p}\sqrt{\boldsymbol{\lambda}_{ij}}\mathbf{U}_{ij}\mathbf{V}_{ij}^{T}. \tag{8}$$

b) Grouping: Accordingly the significant eigentriples are identified on the basis of periodicity and variance

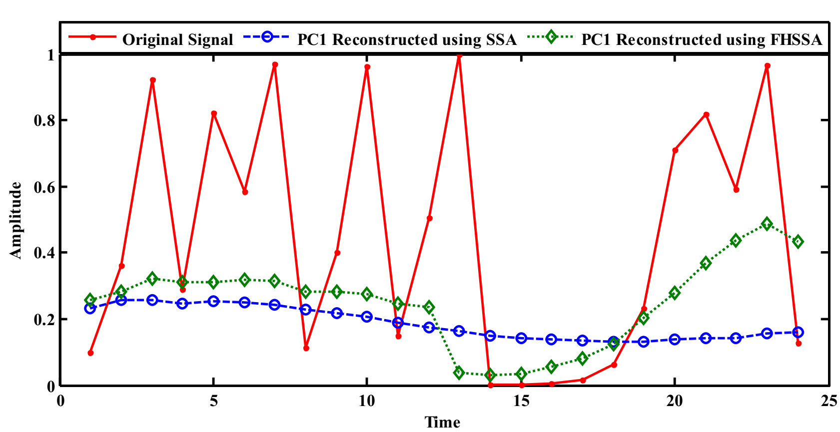

Fig. 2. Complex chaotic signal (red color solid line) along with its first principle component reconstructed using classical SSA (blue color dotted line with circle marker) and FHSSA (green color dotted line with diamond marker).