Journal of Applied Geophysics 115 (2015) 140–144

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

A fast algorithm for coherency estimation in seismic data based on information divergence

Tao Yang [a,c], Bing Zhang [a,b,c], Jinghuai Gao [a,b,c,*]

[a] Institute of Wave and Information, Xi'an Jiaotong University, China
[b] Center for Computational Geosciences, Xi'an Jiaotong University, China
[c] National Engineering Laboratory for Offshore Oil Exploration, Xi'an Jiaotong University, China

ARTICLE INFO

Article history:
Received 30 June 2014
Accepted 4 November 2014
Available online 13 November 2014

Keywords:
Coherency attribute
Information divergence
Discontinuities measure
Fast algorithm

ABSTRACT

Coherency measurements have proven to be an effective method for representing geological discontinuities such as faults and stratigraphic features in 3-D seismic data volumes. Unfortunately, application of the algorithm suffers from the limitation of computation cost. This paper describes a new fast method for efficient and robust coherency estimation in 3-D seismic data. The method is characterized by greatly increasing the computational efficiency based on information divergence, which uses a recursion method and defines a new criterion by information divergence to calculate the coherency, to avoid directly computing the eigenvalues of the covariance matrix. In contrast to other algorithms, this method possesses higher computational efficiency and better anti-noise ability than commonly used methods. We demonstrate the advantage of this method using real seismic data examples.

1. Introduction

The coherency technology is a powerful tool for seismic data interpretation in detection and imaging of geological discontinuities (Chopra and Marfurt, 2007). The first algorithm was proposed by Bahorich and Farmer (1995), which cross correlated each trace with its in-line and cross-line neighbors in the two adjacent axis directions of one quadrant and combined the two results after normalizing by the energy (Bahorich and Farmer, 1995). Since this approach deals with only three traces, it is computationally very efficient, but it may lack robustness, especially when the data is noisy. The second generation of coherency detection algorithm used semblance as a coherence estimate (Marfurt et al., 1998). Using more traces in coherency computation results in better stability in presence of noises. The third generation coherency algorithm, C3, is eigen-structure based, the coherency measures the ratio of dominant eigenvalue and the trace of the covariance matrix which is constructed with a small subvolume within an analyzing window (Gersztenkorn and Marfurt, 1999). It is more robust to measure coherency than the previous algorithm does; however, its application was affected by a large computational cost to construct the covariance matrix and calculate the dominant eigenvalue. In order to reduce the computational cost in this new algorithm, Cohen and Coifman (2002) proposed to use the local structural entropy (LSE) as a coherency measure. Li et al. (2006) and Lu et al. (2005) both used a supertrace technique to reduce the dimensions of the covariance matrix to 4×4. However, all these earlier works have not discussed the proper implementation to maintain the computational efficiency, especially for calculating the dominant eigenvalue of the covariance matrix (Wang et al., 2012).

Wang et al. (2012) proposed an efficient implementation of eigenstructure-based coherency algorithm using recursion strategies, which can greatly reduce the calculation of coherency estimation and improve the computational efficiency. However, as an iterative algorithm using a power method, certain calculation is still needed, especially when all traces within an analysis window are inconsistent, where all of the eigenvalue of the covariance matrix are almost the same, a large number of iterations are required to calculate the dominant eigenvalue. In real applications, the number of iterations is usually predetermined, thus, in the case that all of the eigenvalue of the covariance matrix are nearly indistinguishable, the accuracy will be hard to manage. In this paper, we propose an accelerated algorithm for coherency measurement called information divergence coherency algorithm (IDCA). This new algorithm mainly uses normalized information divergence as a coherency measurement and aims to avoid directly compute the dominant eigenvalue. By combining the recursion strategies proposed by Wang et al. (2012) with information divergence, the method can be rather fast with a high computing precision.

The paper is organized as follows. In Section 2, we briefly introduce the C3 algorithm. The proposed method based on information divergence is introduced in Section 3. In Section 4, a real data example is provided to demonstrate the effectiveness of our method. Finally, we give the conclusion in Section 5.

* Corresponding author at: Institute of Wave and Information, Xi'an Jiaotong University, 710049 Xi'an, China. Tel.: +86 29 82665060.
E-mail address: jhgao@mail.xjtu.edu.cn (J. Gao).

http://dx.doi.org/10.1016/j.jappgeo.2014.11.002
0926-9851/

2. The eigen-structure-based coherency algorithm

The algorithm was first proposed by Gersztenkorn and Marfurt (1999). Suppose that the sub-volume of seismic traces to be analyzed includes J traces (for example, 3 in-line traces by 3 cross-line traces for a total of 9 traces) and N samples per trace when coherency is estimated for a point in a 3D cube. Organizing the amplitude in this suite of traces by sample index n and trace index j results in the data matrix D:

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1J} \\ d_{21} & d_{22} & \dots & d_{2J} \\ \vdots & \vdots & \vdots & \vdots \\ d_{N1} & d_{N2} & \dots & d_{NJ} \end{bmatrix}. \quad (1)$$

The covariance matrix of D for a point can be calculated and formulated by

$$C = D^T D = \begin{bmatrix} \sum_{n=1}^{N} d_{n1}^2 & \sum_{n=1}^{N} d_{n1} d_{n2} & \dots & \sum_{n=1}^{N} d_{n1} d_{nJ} \\ \sum_{n=1}^{N} d_{n1} d_{n2} & \sum_{n=1}^{N} d_{n2}^2 & \dots & \sum_{n=1}^{N} d_{n2} d_{nJ} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{n=1}^{N} d_{n1} d_{nJ} & \sum_{n=1}^{N} d_{n2} d_{nJ} & \dots & \sum_{n=1}^{N} d_{nJ}^2 \end{bmatrix}. \quad (2)$$

The covariance matrix C in Eq. (2) is symmetric and non-negative, i.e., all its eigenvalues are greater than or equal to zero. The eigenstructure-based coherency makes use of the covariance matrix C, denoted by $Tr(C)$, which demonstrates the total energy for the traces enclosed by the analysis cube, and is equal to the sum of the eigenvalues (Chopra and Marfurt, 2007; Gersztenkorn and Marfurt, 1999).

$$C_3 = \frac{\lambda_1}{Tr(C)} = \frac{\lambda_1}{\sum_{j=1}^{J} c_{jj}} = \frac{\lambda_1}{\sum_{j=1}^{J} \lambda_j}. \quad (3)$$

Eq. (3) defines the eigen-structure coherency as a ratio of the dominant eigenvalue λ_1 to the total energy within the analysis cube. The expression of the last term in Eq. (3) can have a significant impact on the computational efficiency of the coherency calculations for 3-D seismic data. Summing the diagonal entries of the covariance matrix C is much more efficient than computing all the eigenvalues followed by their summation.

In fact, the distribution of eigenvalues describes the degree of coherence between the seismic traces within the analysis window. For instance, in two extreme cases, waveforms will be completely consistent when only the dominant eigenvalue is greater than zero and all others are equal to zero (C_3 will be 1.0), or completely inconsistent when all the eigenvalues are the same (C_3 will be $^1/_J$ and the range of coherence value from $^1/_J$ to 1.0 in Eq. (3)).

In addition, inspired by the information theory (Anastassiou, 2004; Qiao and Minematsu, 2008; Salicrú, 1994), there are different ways for discontinuity measurement, such as information measurement, which plays a key role in information divergence. In the following section, we will define a measure of coherency using the information divergence and will focus on discussing the information divergence coherency algorithm (IDCA).

3. The information divergence coherency algorithm

Based on the statistics and information theory, one of the important issues in the application of probability theory is to find an appropriate measure of difference (distance or discrimination) between two probability distributions. Information divergence for this purpose provides a measure of the difference or distance between two probability distributions. Let X and Y be two discrete random variables with distribution being $\mathbf{p}_1 = (p_{11}, p_{12}, \dots, p_{1J})$ and $\mathbf{p}_2 = (p_{21}, p_{22}, \dots, p_{2J})$, respectively. The information divergence of the probability distribution between $\mathbf{p_1}$ and $\mathbf{p_2}$ is given by (Zhang and Gao, 2011; Yang et al., 2013)

$$D_\phi(\mathbf{p}_1, \mathbf{p}_2) = \sum_{n=1}^{J} \mathbf{p}_{1n} \phi\left(\frac{\mathbf{p}_{2n}}{\mathbf{p}_{1n}}\right) \quad (4)$$

where $\phi : (0, \infty) \to R$ is a convex function and $\phi(1) = 0$.

For a convex function ϕ defined on R, applying classical the Jensen inequality for variable $X(n) = \mathbf{p}_{2n}/\mathbf{p}_{1n}$, then, $X(n)$ satisfies

$$\mathrm{E}[\phi(X)] \geq \phi(\mathrm{E}[X]) \quad (5)$$

here E [] denotes the mathematical expectation.

Therefore, using the Jensen's inequality for Eq. (4), we obtain

$$D_\phi(\mathbf{p}_1, \mathbf{p}_2) = \sum_{n=1}^{J} \mathbf{p}_{1n} \phi\left(\frac{\mathbf{p}_{2n}}{\mathbf{p}_{1n}}\right) \geq \phi\left(\sum_{n=1}^{J} \mathbf{p}_{1n} \frac{\mathbf{p}_{2n}}{\mathbf{p}_{1n}}\right) = \phi(1) = 0. \quad (6)$$

But the equality holds if and only if $\mathbf{p}_1 = \mathbf{p}_2$.

Considering the above eigenvalue-based algorithm, if we choose $\mathbf{p_1}$ and $\mathbf{p}_2$ as

$$\mathbf{p}_{1n} = \frac{\lambda_n}{Tr(A)} n = 1, 2, \dots, J \quad (7)$$

$$\mathbf{p}_{2n} = \frac{1}{J} n = 1, 2, \cdots, J \quad (8)$$

where A is a covariance matrix formed from a sub-volume of seismic traces within an analysis window, λ_n denotes the *nth* eigenvalue of A, $Tr(A)$ is the trace of matrix A, J is the dimension of A. $\mathbf{p}_2$ indicates the situation in which all the traces within an analysis window have different shapes. Thus, we can see that $D_\phi(\mathbf{p}_1, \mathbf{p}_2)$ is nonnegative, and will become larger when more traces within the analysis window have the same shapes.

If $\phi(x) = x \log x$, we have the Kullback–Liebler divergence, as given by

$$D^{KL}(\mathbf{p}_1, \mathbf{p}_2) = \sum_{n=1}^{J} p_{1n} \log_a \frac{\mathbf{p}_{1n}}{\mathbf{p}_{2n}} \quad (a > 1). \quad (9)$$

Substituting Eqs. (7) and (8) into Eq. (9), we obtain

$$\begin{aligned} D^{KL} &= \sum_{n=1}^{J} \mathbf{p}_{1n} \log_a \frac{\mathbf{p}_{1n}}{\mathbf{p}_{2n}} = \sum_{n=1}^{J} \frac{\lambda_n}{Tr(A)} \left(\log_a J + \log_a \frac{\lambda_n}{Tr(A)} \right) \\ &= \sum_{n=1}^{J} \frac{\lambda_n}{Tr(A)} \left[\log_a J + \frac{1}{\ln a} \ln(1 + \left(\frac{\lambda_n}{Tr(A)} - 1 \right) \right] \\ &= \sum_{n=1}^{J} \frac{\lambda_n}{Tr(A)} \left[\log_a J + \frac{1}{\ln a} \sum_{n=1}^{\infty} (-1)^{m-1} \frac{\left(\frac{\lambda_n}{Tr(A)} - 1 \right)^m}{m} \right] \\ &= \sum_{n=1}^{J} \frac{\lambda_n}{Tr(A)} \left[\log_a J + \frac{1}{\ln a} \sum_{n=1}^{\infty} (-1)^{m-1} \frac{\left(1 - \frac{\lambda_n}{Tr(A)} \right)^m}{m})(-1)^m \right] \\ &= \sum_{n=1}^{J} \frac{\lambda_n}{Tr(A)} \left[\log_a J - \frac{1}{\ln a} \sum_{n=1}^{\infty} \frac{\left(1 - \frac{\lambda_n}{Tr(A)} \right)^m}{m} \right] \\ &= \frac{1}{Tr(A)} \left[Tr(A) \log_a J - \frac{1}{\ln a} \sum_{n=1}^{J} \sum_{m=1}^{\infty} \frac{\left(1 - \frac{\lambda_n}{Tr(A)} \right)^m}{m} \lambda_n \right] \\ &\approx \frac{1}{Tr(A)} \left[Tr(A) \log_a J - \frac{1}{\ln a} \sum_{n=1}^{J} \sum_{m=1}^{M} \frac{\left(1 - \frac{\lambda_n}{Tr(A)} \right)^m}{m} \lambda_n \right]. \end{aligned} \quad (10)$$