

The Research of Multi-Source Data Separation Technology

Ke Wang

Graduate School of China University of Geosciences, Wuhan Hubei
Email: 1216684257@qq.com

Received: Apr. 5th, 2018; accepted: Apr. 18th, 2018; published: Apr. 26th, 2018

Abstract

This paper treats the separation of the double-source data as a blind signal separation problem, and introduces the basic theory of blind signal separation. In this paper, for the case where the number m of observed signals and the number n of sources are the same, a simple blind signal separation process is designed and the numerical simulation is carried out to prove the effectiveness of the proposed method.

Keywords

Multi-Source, Blind Signal Separation

多震源盲分离技术研究

王珂

中国地质大学研究生院, 湖北 武汉
Email: 1216684257@qq.com

收稿日期: 2018年4月5日; 录用日期: 2018年4月18日; 发布日期: 2018年4月26日

摘要

本文将双震源数据分离当成一个盲信号分离(Blind Signal Separation, BBS)的问题, 介绍了盲信号分离基本理论, 重点简述了线性瞬时模型的分离方法。对于混合数据个数 m 和单震源数 n 相同的情况, 本文设计了一种基于去相关的盲信号分离处理方法, 并对该分离方法做了数值模拟, 证明了方法的有效性。

关键词

多震源, 盲分离

Copyright © 2018 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

多震源地震信号分离方法主要是通过分离多个震源激发而产生的混叠地震记录, 从而得到混叠记录中的源信号的方法。由于多炮点激发相比单炮激发可以大幅度提高采集效率, 降低成本, 因此越来越受到重视[1]。而对于采集到的混叠信号, 将其分离再进一步分析的方式是一种重要的研究方法。

而其中盲源分离(Blind Source Separation 简称 BSS)技术, 是一种近几年迅速发展起来的多维信号处理方法。它以统计独立的基本原则建立目标函数, 通过优化算法将观测信号分解为若干独立成分, 来达到提取有用信息的目的, 在通信、医学、勘探等多个领域获得了广泛的应用。本文在介绍其基本原理和具体算法研究的基础上, 以该方法对双源地震混叠信号进行分离, 发现其可以有效地将独立的源信号进行分离, 取得了令人满意的结果。

2. 多震源混合信号概述

我们不妨将接收器接收到的单震源卷积混合数据看成一个独立的地震信号, 当多震源激发时, 不同震源的卷积混合数据显然是线性叠加的[2], 那么我们不妨将多震源地震数据表示为以下形式:

$$P_{bl} = \Gamma P \quad (1)$$

其中 P_{bl} 表示了多震源混合数据, P 为单个震源地震数据, Γ 为混合算子。用离散的形式表示有:

$$P_{bl}^j = \sum_{k=1}^N \Gamma_k^j P_k \quad (2)$$

其中, $j = 1, 2, \dots, M$, M 表示扫描次数, N 为道集数。

而通过多震源波场数据 P_{bl}^j , 通过计算得到单震源波场数据 P_k 的过程, 称作多震源地震波分离。

若 Γ_k^j 已知时, 实际上它是一个一个线性反演问题, 而当混合算子 Γ_k^j 未知, 则此波场分离是一个盲过程, 也就是我们上面提到的 BSS 技术。

3. BSS 简介

3.1. BSS 基本原理

我们来考虑这样一种情况, 现在有一组信号, 它们由几个不同的源发出, 并存在多个接收器, 分别接收了不同权重的信号的混合。

具体来说, 我们可以假设共有三个源信号及三个接收器, 观测信号我们不妨假设为 $x_1(t)$, $x_2(t)$, $x_3(t)$; 而未知的源信号不妨设为 $s_1(t)$, $s_2(t)$, $s_3(t)$; 这样, 观测信号为原始信号的加权和, 不妨设为:

$$\begin{aligned} x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\ x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\ x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t) \end{aligned} \quad (3)$$

由于 a_{ij} 和 $s_i(t)$ 都是未知的，因此我们不能对它们进行直接估计。

我们现在要做的，就是在对于原始信号及其传播的物理过程都知之甚少 的情况下，利用接收器的混合信号对原始信号进行估计。这就是盲源分离(BBS)问题。“盲”，意味着我们所知道的原始信号信息很少[3]。

而实际上，混叠信号一般可以有 两种混叠方式，一种是我们上面提到的线性混合，即：

$$x_i(t) = \sum_{j=1}^N a_{ij} s_j(t), \quad i = 1, 2, \dots, N \quad (4)$$

另外一种 是卷积混合，即混合信号为源信号的不同时延信号的混叠，即：

$$x_i(t) = \sum_{j=1}^N \sum_{k=0}^P a_{ij} s_j(t-k), \quad i = 1, 2, \dots, N \quad (5)$$

3.2. 信号预处理

在对我们由接收器得到的混合信号进行 ICA 前，我们首先要对信号进行预处理，预处理的过程主要由两部分组成：去均值和白化。

3.2.1. 去均值过程

在对我们由接收器得到的混合信号进行 ICA 前，我们首先要对信号进行预处理，预处理的过程主要由两部分组成：去均值和白化。

由于在大多数的 ICA 算法中，对原始信号都有信号均值为 0 的假设，因此我们对具体问题做 ICA 前，需要首先去均值，即中心化。

因此，我们首先需要设 $\bar{x}_i(t)$ ，使得：

$$\bar{x}_i(t) = x_i(t) - \frac{1}{N} \sum_{i=1}^N x_i(t), \quad i = 1, 2, \dots, N \quad (6)$$

3.2.2. 白化过程

白化总是可以实现的，一般可以利用协方差矩阵的特征值分解：

$$E\{\mathbf{xx}^T\} = \mathbf{E}\mathbf{D}\mathbf{E}^T \quad (7)$$

式中， \mathbf{E} 是 $E\{\mathbf{xx}^T\}$ 的特征向量的正交矩阵， \mathbf{D} 是相应的特征向量的对角矩阵。这样，白化过程可以利用以下的白化矩阵来实现：

$$\mathbf{V} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T \quad (8)$$

式中，矩阵 $\mathbf{D}^{-1/2} = \text{diag}\left(d_1^{-1/2}, \dots, d_n^{-1/2}\right)$ 。

由于白化过程中，新的混合矩阵 $\tilde{\mathbf{A}} = \mathbf{V}\mathbf{A}$ 是正交的：

$$E\{\mathbf{zz}^T\} = \tilde{\mathbf{A}}E\{\mathbf{ss}^T\}\tilde{\mathbf{A}}^T = \mathbf{I} \quad (9)$$

这就意味着我们可以把对混合矩阵的搜索范围限制到正交矩阵的空间中。我们可以无须估计原始矩阵 \mathbf{A} 的全部 n^2 个参数，只要估计一个正交混合矩阵即可。一个正交混合矩阵包含 $n(n-1)/2$ 个自由度。即是说，正交矩阵包含的可变参数个数基本上只有不到一半的任意矩阵参数个数。

4. 算法介绍及数据模拟

4.1. 线性混合盲分离算法

关于线性混合盲分离算法有很多，我们在这里提出一种基于二阶统计量的盲分离算法。该盲分离算

法也被称为去相关算法，并对源信号一般有一定的要求，即具有统计不相关性、非白性。就双震源混合数据的盲分离问题而言，其系统的输入输出关系可以表示为

$$\begin{cases} y_1(t) = x_1(t) - \omega_1 x_2(t) \\ y_2(t) = x_2(t) - \omega_2 x_1(t) \end{cases} \quad (10)$$

令 $r_{y_1 y_2}(t_1, t'_1) = E[y_1(t_1)y_2(t'_1)] = r_{x_1 y_2}(t_1, t'_1) - \omega_1 r_{x_2 y_2}(t_1, t'_1) = 0$ ，得到

$$\omega_1 = \frac{r_{x_1 y_2}(t_1, t'_1)}{r_{x_2 y_2}(t_1, t'_1)} = \frac{r_{x_1 y_2}(t_1, \tau_1)}{r_{x_2 y_2}(t_1, \tau_1)} \quad (11)$$

令 $r_{y_1 y_2}(t_2, t'_2) = E[y_1(t_2)y_2(t'_2)] = r_{y_1 x_2}(t_2, t'_2) - \omega_1 r_{y_1 x_1}(t_2, t'_2) = 0$ ，得到

$$\omega_2 = \frac{r_{y_1 x_2}(t_2, t'_2)}{r_{y_1 x_1}(t_2, t'_2)} = \frac{r_{y_1 x_2}(t_2, \tau_2)}{r_{y_1 x_1}(t_2, \tau_2)} \quad (12)$$

分离矩阵如下：

$$W = \begin{pmatrix} 1 & -\omega_1 \\ -\omega_2 & 1 \end{pmatrix} \quad (13)$$

对于 $\omega_1, \omega_2 \in [-1, 1]$ 范围内，我们可以任意赋予初值并迭代。构成了双震源混合数据的盲分离算法。该去相关算法计算简单，但对于初值选择有依赖性，收敛速度较慢。

本文依据白化处理和旋转处理算法，设计了一种新的基于去相关的盲分离方法，分离矩阵 W 的更新公式表示如下

$$W^{(i)}W \rightarrow W$$

$$W^{(i)} = \begin{bmatrix} \cos \theta_{ij} & \sin \theta_{ij} \\ -\sin \theta_{ij} & \cos \theta_{ij} \end{bmatrix} \quad (14)$$

其中， θ_{ij} 为分离信号 y_i, y_j 之间的角度。通过白化处理和旋转处理可以分别消除数据的二阶相关性以及高阶相关性。

4.2. 算法步骤

假设源信号的相关矩阵 $R_{xx}(0) = E[SS^T] = I$ ，即其具有零均值、单位方差、互不相关的性质。本文设计的盲信号分离处理方法，具体需要以下几个步骤来完成：

- 1) 将观测得到的混合数据用 X 表示；
- 2) 求出零时间延迟相关矩阵 $R_{xx}(0) = E[XX^T]$ ；
- 3) 对 $R_{xx}(0)$ 进行奇异值分解，得到 $R_{xx}(0) = UAU$ ， U 为正交矩阵， A 为对角矩阵；
- 4) 计算预白化矩阵 $B = A^{-1/2}U^T$ ，则得到预白化后和混合数据为 $Z = BX$ ；
- 5) 初始化盲分离信号 Y ，选取一对成分 y_i, y_j ，计算角度 θ_{ij} ；
- 6) 如果 $\theta_{ij} > \varepsilon$ ，其中 $\varepsilon \ll 1$ ，求出 $W^{(i)} = \begin{bmatrix} \cos \theta_{ij} & \sin \theta_{ij} \\ -\sin \theta_{ij} & \cos \theta_{ij} \end{bmatrix}$ ，更新分离矩阵，

即 $W^{(i)}W \rightarrow W$ ；

- 7) 旋转信号矩阵 $Z: Z \leftarrow W^{(i)}Z$ ，直到 $\theta_{ij} \leq \varepsilon$ ，否则跳回(5)。

上述步骤中的预白化处理的本质是使其相关矩阵 R_{xx} 变成单位矩阵，进而消除了混合信号的二阶相关性，最终达到分离混合数据的目的。

4.3. 数值模拟

从图 1 中可以看出,混合震源数据是来自 2 个震源的激发信号,无法按照常规的方法直接进行处理,故需将其按照本文所设计的盲分离方法进行分离处理。混合震源中包含的单震源个数为 2,则混合度为 2,混合数据预白化处理后,由于混合度较低,经过一次旋转处理后,基本完成了双源数据的分离任务。通过图 1 和图 2 的对比可以看出:将双震源混合数据(无加性噪音)按照本文设计的盲信号分离处理方法进行处理,其分离结果较好。因此,本文设计的盲信号分离处理方法能够有效地处理混合矩阵未知的双震源混合数据分离问题,成功完成了混合数据的分离任务。

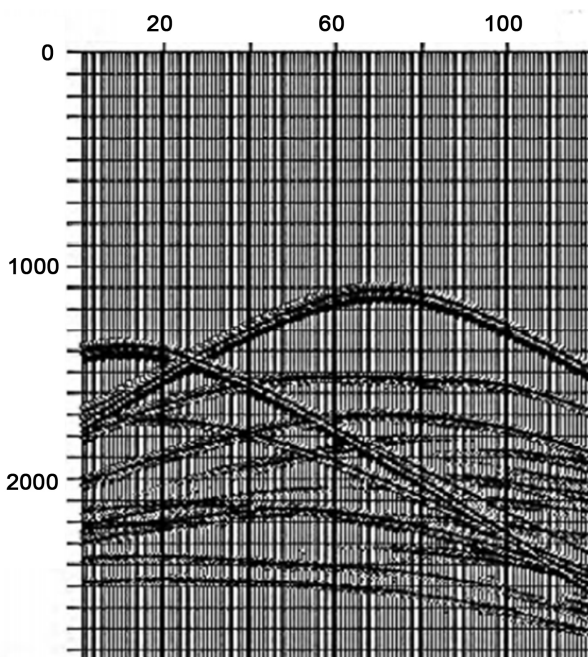


Figure 1. Multi-source data (dual source)
图 1. 混合震源数据(双震源)

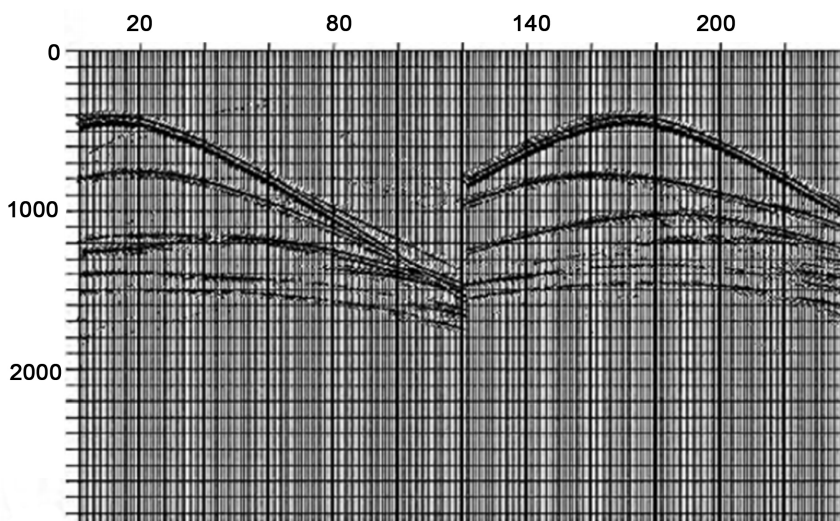


Figure 2. Separate data
图 2. 分离数据

5. 结语

本文将双震源混合数据的分离当成一个盲信号分离问题，在混合数据个数 m 和单震源数 n 相同的情况下，本文设计了一种基于去相关的盲源分离算法，成功实现了对混合数据的分离处理。事实上，混合数据通过预白化处理后，观测信号与源信号之间仅存在正交关系，故可以对处理后混合数据进行旋转处理，即可实现盲源分离。本文使用该去相关分离方法对混合数据分离做了数值模拟，证明了该方法的有效性，适用于混合矩阵 Γ 未知的多震源采集方式。本文提出的盲分离方法较传统的二阶盲分离法，简化了计算，并提高了算法的收敛速度。

参考文献

- [1] 倪宇东. 可控震源地震勘探采集技术[M]. 北京: 石油工业出版社, 2014.
- [2] 李剑锋. 基于盲源分离的地震信号处理方法研究及应用[D]: [硕士学位论文]. 青岛: 中国石油大学(华东), 2012.
- [3] Aapohyvarinen, Juhakarhunen, Erkk. 独立成分分析[M]. 北京: 电子工业出版社, 2014.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7991, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: aam@hanspub.org