

文章编号: 1000-8152(2001)01-0105-04

一种奇异混合信号盲分离的神经网络模型*

李远清 张丽清 刘永清 胡刚

(华南理工大学自动控制工程系·广州, 510640)

摘要: 混合信号盲分离问题是一类很难而又具有很强应用背景的问题, 以往对这类问题的研究均在混合矩阵为非奇异的条件下进行. 本文给出一种神经网络模型及相应算法, 处理混合矩阵为奇异的混合信号盲分离问题.

关键词: 信号盲分离; 奇异性; 神经网络; 噪声

文献标识码: A

A Kind of Neural Network Model for Blind Separation of Singularly Mixed Sources

LI Yuanqing, ZHANG Liqing, LIU Yongqing and HU Gang

(Department of Automatic Control Engineering, South China University of Technology, Guangzhou, 510640, P. R. China)

Abstract: Blind separation of sources is a problem arising in many practical fields, which is difficult to research. In the existing references, the mixing matrix of sources is always supposed to be nonsingular. This paper presents an algorithm and a corresponding networks to deal with blind separation of sources in the case of singular mixing matrix.

Key words: blind separation; singularity; neural networks; noise

1 引言(Introduction)

在很多情形下, 我们通过感知器(天线, 声纳, 录音设备等)接收到的信号是一些源信号的混合, 有时可能会附加随机噪声, 其混合过程常可简单表示为^[1-3]

$$Y(t) = AX(t), \quad (1.1)$$

此处, $X(t) = [x_1(t), \dots, x_n(t)]^T$ 为 n 个源信号组成的向量, $Y(t) = [y_1(t), \dots, y_n(t)]^T$ 为接收到的 n 个混合信号组成的向量, A 为一 $n \times n$ 常数矩阵, 称为混合矩阵, 记为:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}. \quad (1.2)$$

实际情况是, 我们对源信号及其混合过程知之甚少, 即在式(1.1)中, $A, X(t)$ 是未知的, 一般只假定 x_1, \dots, x_n 互相独立, 均值为零, 要根据已知的 $Y(t)$, 将源信号 $X(t)$ 分离出来, 这就是信号盲分问题, 这是一个很难又具有重要意义的问题. 如在生物医学信号处理 (EEG, MEG, ECG, EMG), 无线通讯, 水下声学, 天文学, 天体物理等诸多领域都有重要应用.

信号盲处理是 80 年代中期出现的一种新技术,

其涵盖的内容十分广泛, 包括信号盲分离、信号盲解卷、信道盲辨识、图像盲恢复等^[4,5]. 信号盲分离是信号盲处理的一个重要方面, 目前还处于发展中. 从我们掌握的文献来看, 以往处理的问题一般假设(1.1)中 A 为方阵, 即接收到的信号个数等于源信号个数, 且假设 A 为非奇异的, 即 A 的逆矩阵存在; 以往处理盲分问题的神经网络模型一般是单层线性模型, 这也限制了网络处理能力; 另外, 由于传送的滞后, 我们接收到的信号往往是当前信号与过去信号的混合, 而发生滞后的形式又多种多样, 对这类信号盲分问题的研究, 又称为动态分量分析, 目前也有很多研究结果在不断涌现.

本文在假设 A 为奇异阵的情形下, 给出一种神经网络模型及相应算法, 进行混合信号盲分. 本文分五个部分, 第二部分给出理论解, 第三部分给出网络模型及算法, 第四部分给出仿真结果, 第五部分总结.

本文总假定 x_1, \dots, x_n 中的随机信号互相独立, 均值为零.

2 理论解(Theoretical solution)

关于 A , 首先我们给出一假设.

假设 2.1 A 中 n 个行向量记为 a_1, a_2, \dots, a_n .

* 基金项目: 国家自然科学基金重点基金(69574009), 国家自然科学基金(60004004), 广东省自然科学基金(990584)资助项目.

收稿日期: 1998-11-30; 收修改稿日期: 2000-06-26.

其中有一行向量, 设为 a_i , 满足

$$[a_{i1}, \dots, a_{i(j-1)}, a_{i(j+1)}, \dots, a_{in}] = \sum_{l=1, l \neq i}^n k_l [a_{l1}, \dots, a_{l(j-1)}, a_{l(j+1)}, \dots, a_{ln}], \quad (2.1)$$

但是, $a_i \neq \sum_{l=1, l \neq i}^n k_l a_l$.

本文中, 不妨设 a_1 满足假设 2.1, 即:

$$[a_{12}, \dots, a_{1n}] = k_2 [a_{22}, \dots, a_{2n}] + \dots + k_n [a_{n2}, \dots, a_{nn}], \quad (2.2)$$

但是, $a_1 \neq \sum_{l=2}^n k_l a_l$, 即 a_1 不是其它行向量的线性组合. 例如, 下面奇异矩阵便满足上述条件:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}. \quad (2.3)$$

上述混合矩阵的含义是, 第一个混合信号为三个源信号的简单叠加, 第二个混合信号与第三个混合信号一样为两个源信号的简单叠加. 此处实际上只收到两个混合信号. 如果对矩阵 A 中任意列向量 a_i , 假设 2.1 均不满足, 即 A 中任一行向量均为其它行向量的线性组合, 则本文算法失效.

在上述假设下, 有:

$$\begin{aligned} \begin{bmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_n(t) \end{bmatrix} &= \begin{bmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{21}}{b} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{a_{n1}}{b} & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 & -k_2 & \dots & -k_n \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} Y(t) = \\ &= \begin{bmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{21}}{b} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{a_{n1}}{b} & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 & -k_2 & \dots & -k_n \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \\ &= \begin{bmatrix} b & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} bx_1(t) \\ a_{22}x_2(t) + \dots + a_{2n}x_n(t) \\ \vdots \\ a_{n2}x_2(t) + \dots + a_{nn}x_n(t) \end{bmatrix}, \quad (2.6) \end{aligned}$$

即

$$\begin{cases} z_1(t) = y_1(t) - k_2 y_2(t) - \dots - k_n y_n(t) = bx_1(t), \\ z_2(t) = y_2(t) - \frac{a_{21}}{b} z_1(t), \\ \vdots \\ z_n(t) = y_n(t) - \frac{a_{n1}}{b} z_1(t). \end{cases} \quad (2.7)$$

$$\begin{bmatrix} 1 & -k_2 & \dots & -k_n \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} - \sum_{j=2}^n k_j a_{j1} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}. \quad (2.4)$$

记: $b = a_{11} - \sum_{j=2}^n k_j a_{j1}$, 显然 $b \neq 0$. 又

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{21}}{b} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{a_{n1}}{b} & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} b & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} b & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad (2.5)$$

故

故采用上述方法, 便可将源信号 x_1 分离出来, 即 $z_1 = bx_1$, 与 x_1 只相差一系数. 剩下的混合信号 z_2, \dots, z_n , 可继续利用上述方法进行信号抽取. 注意到剩下的混合信号 z_2, \dots, z_n 已与 z_1 独立, 这正是我们在本文第三部分进行网络权重训练的依椐.

3 模型与算法 (Model and algorithm)

我们采用如下网络模型:

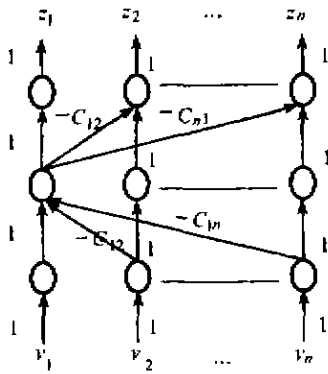


图1 神经网络模型
Fig. 1 Neural network model

其数学表达式为:

$$\begin{cases} z_1(t) = y_1(t) - c_{12}(t)y_2(t) - \dots - c_{1n}(t)y_n(t), \\ z_2(t) = y_2(t) - c_{21}(t)z_1(t), \\ \vdots \\ z_n(t) = y_n(t) - c_{n1}(t)z_1(t). \end{cases} \quad (3.1)$$

权重 $c_{12}, \dots, c_{1n}, c_{21}, \dots, c_{n1}$ 根据如下规则进行训练:

$$\begin{cases} c_{12}(k+1) = c_{12}(k) + \alpha_{12}z_1(k)z_2(k), \\ \vdots \\ c_{1n}(k+1) = c_{1n}(k) + \alpha_{1n}z_1(k)z_n(k), \\ c_{21}(k+1) = c_{21}(k) + \alpha_{21}z_1(k)z_2(k), \\ \vdots \\ c_{n1}(k+1) = c_{n1}(k) + \alpha_{n1}z_1(k)z_n(k). \end{cases} \quad (3.2)$$

其中, $\alpha_{12}, \dots, \alpha_{n1}$ 均为正常数. 由式(3.2)可见, 当 $E(z_1z_2) = 0, \dots, E(z_1z_n) = 0$ 时, 权 $c_{ij}(t)$ 稳定下来. 如果 x_1, \dots, x_n 全部为随机信号, 根据第一部分的假设, 它们相互独立, 且均值为零, 而 z_1, \dots, z_n 为 x_1, \dots, x_n 的线性组合, 故 $E(z_1) = 0, \dots, E(z_n) = 0, E(z_1z_2) = 0, \dots, E(z_1z_n) = 0$ 蕴含了 z_1 分别与 z_2, \dots, z_n 不相关.

4 仿真结果(Simulation results)

本节我们进行两个仿真实验, 源信号均为: $[x_1(t), x_2(t), x_3(t)]^T$, 其中 $x_1(t) = \sin(t)$ 或 $\sin(3t)$, 随机信号 $x_2(t)$ 服从 $[-0.5, +0.5]$ 上的均匀分布, $x_3(t)$ 服从 $N(0,1)$ 标准正态分布, 如图 2, 3 的第一组图.

仿真 4.1 奇异混合矩阵 A 取

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

如图 2, 混合信号 Y 如第二组图所示, 利用第四部分的算法进行信号分离, 其中, 权重 c_{ij} 的初值均为 0, $\alpha_{12} = \alpha_{13} = 0.001, \alpha_{21} = \alpha_{31} = 0.00001$, 计算结果 z_1, z_2, z_3 如第三组图所示, 而第四组图分别表示 $z_1 - x_1, z_1 - x_2, z_1 - x_3$, 可见本文的算法是有效的.

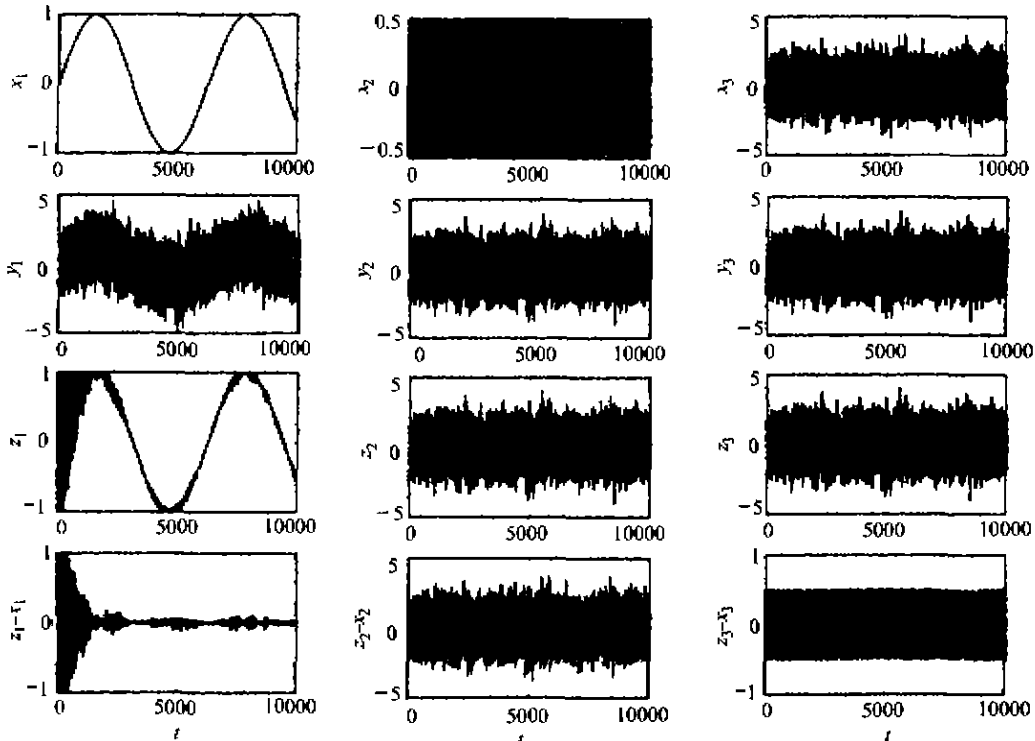


图2 仿真 4.1
Fig. 2 Simulation 4.1

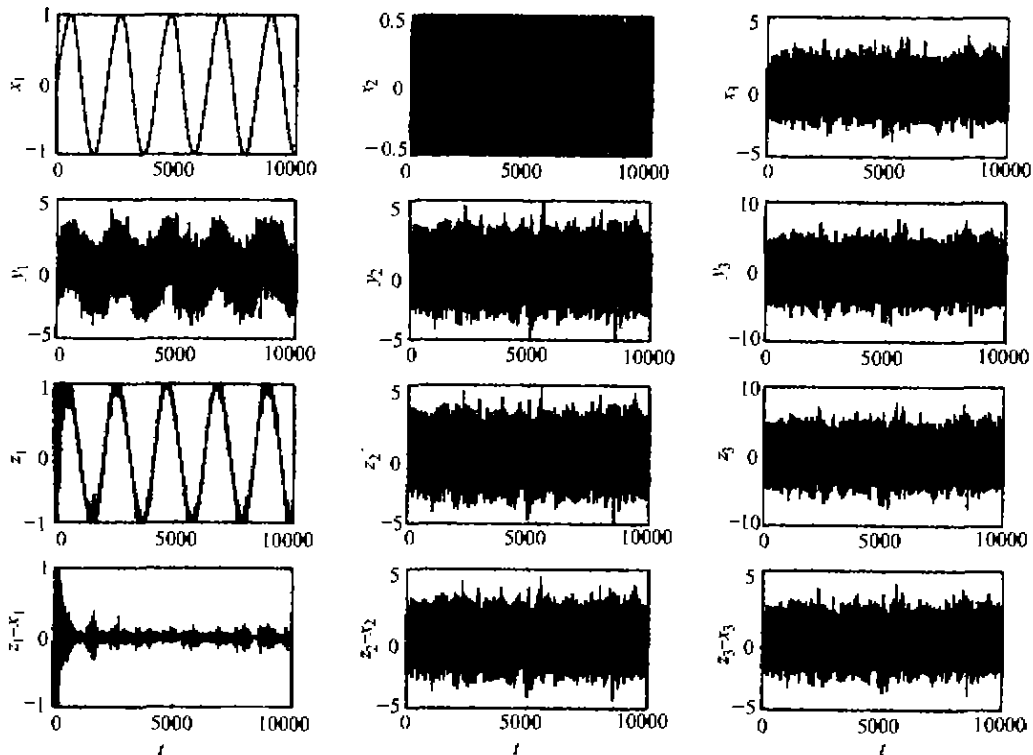


图3 仿真 4.2

Fig. 3 Simulation 4.2

仿真 4.2 奇异混合矩阵 A 为

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

此矩阵为非奇异的,源信号 $x_1(t) = \sin(3t)$, x_2, x_3 同上.如图 3,混合信号 Y 如第二组图所示,利用第四部分的算法进行信号分离,其中,权重 c_{ij} 的初值均为 0, $\alpha_{12} = \alpha_{13} = 0.001$, $\alpha_{21} = \alpha_{31} = 0.00001$, 计算结果 z_1, z_2, z_3 如第三组图所示,而第四组图分别表示为 $z_1 - x_1, z_1 - x_2, z_1 - x_3$, 可见,本文的算法对非奇异混合矩阵也是有效的.

5 结论(Conclusion)

1) 本文的主要目的是处理混合矩阵为奇异的情形,但对混合矩阵为非奇异的情形也是适用的.只是权重的初值及步长 a_{ij} 的选取很有讲究,有时需要变步长,这涉及到寻优算法的收敛性与稳定性.

2) 若假设 2.1 不满足,则本文算法失效,但非奇异矩阵总满足假设 2.1.

3) 本文的盲分是逐步进行的,一次计算只抽取一个源信号,剩下的混合信号可继续用本文算法进行处理,此时的混合矩阵已较原来的少了一阶.如果混合矩阵非奇异, n 次处理可将源信号完全分离,如果混合矩阵奇异,多次处理也只能分离出部分源信号.这样分步抽取源信号,即使不能将源信号完全分离,也是有意义的,因为实际情况往往是只要有用的源信号,去掉附加噪声,就像仿真实验 4.1 一样.

4) 至于源信号个数与接收到信号个数不相等

的情形,也可归结为混合矩阵为奇异的问题来处理.如仿真 4.1,第二个接收信号与第三个相等,实际上只有两个不同的接收信号.

参考文献(References)

- [1] Jutten C and Herault J. Blind separation of sources, Part I: an adaptive algorithm based on neuromimetic architecture [J]. Signal Processing, 1991, 24(1): 1-10
- [2] Jutten C and Herault J. Blind separation of sources, Part II: problems statement [J]. Signal Processing, 1991, 24(1): 11-20
- [3] Jutten C and Herault J. Blind separation of sources, Part III: Stability analysis [J]. Signal Processing, 1991, 24(1): 21-29
- [4] Bell A J and Sejnowski T J. An information-maximization approach to blind separation and blind deconvolution [J]. Neural Computation, 1995, 7(4): 1129-1159
- [5] Cichocki A, Shun-ichi Amari and Thawonmas R. Blind signal extraction using self-adaptive nonlinear Hebbian learning rule [A]. Proc. the 1996 International Symposium on Linear Theory and its Application [C], Kochi, Japan, 1996, 337-380

本文作者简介

李远清 1966 年生, 1994 年华南师范大学硕士毕业, 1997 年华南理工大学博士毕业. 现为华南理工大学电子与信息学院副教授. 目前研究方向: 滞后广义系统, 信号盲处理.

张丽清 华南理工大学电子与信息学院教授, 博士生导师. 现正在日本 RIKEN 研究中心从事信号盲处理方面研究工作.

刘永清 1930 年生, 华南理工大学电子与信息学院教授, 博士生导师. 主要研究方向为复杂系统及系统工程.

胡刚 广东工业大学研究生处副处长, 副教授. 现为华南理工大学自动控制系博士生. 感兴趣研究方向: 复杂系统的控制.