

一种基于 ICA 的盲信号分离快速算法

游荣义^{1,2}, 陈忠²,

(1. 集美大学计算科学与应用物理系, 福建厦门 361021; 2. 厦门大学物理系, 福建厦门 361005)

摘要: 基于 ICA(独立成分分析: Independent Component Analysis) 原则, 给出一种盲信号分离的快速学习算法. 通过寻求观测变量线性组合的四阶累积量(即 kurtosis 系数)局部极值, 得出该算法的模型和步骤. 将该算法用于盲信号分离实验, 实验结果表明, 该算法在盲信号分离和信号特征提取方面具有收敛速度快、无需动态参数等优点. 该算法能有效地分离出任意分布的非高斯盲源信号的各个独立成分, 是信号处理的一种新的、高效可靠的方法.

关键词: 盲信号分离; 独立成分分析; kurtosis; 神经算法

中图分类号: TN911, 957 **文献标识码:** A **文章编号:** 0372-2112 (2004) 04-0669-04

A Fast Algorithm of Blind Signal Separation Based on ICA

YOU Rong-yi^{1,2}, CHEN Zhong²

(1. Department of Computational Science & Applied Physics, Jimei University, Xiamen, Fujian 361021, China;

2. Department of Physics, Xiamen University, Xiamen, Fujian 361005, China)

Abstract: Based on ICA (Independent Component Analysis) principle, a fast learning algorithm for blind signal separation is presented. By seeking the local extrema of the fourth-order cumulants (i. e. kurtosis coefficients) of a linear combination of the observed variables, the model and the process of this algorithm are obtained and then used for the experiment of blind signal separation. The results of the experiment show that this algorithm has a great deal advantages in blind signal separation and features extraction such as fast convergence and needless in any dynamic parameter and the like. The algorithm can separates all independent components from blind source signals, which is non-Gaussian distribution. The algorithm is a new, highly efficient and reliable method in signal processing.

Key words: blind signal separation; independent component analysis; kurtosis; neural algorithm

1 引言

在神经网络应用、统计分析、信号处理等领域, 一个普遍关心和感兴趣的问题是如何借助某种适当的变换, 找到源数据的一个恰当表示. 如果源数据的表示有利于后续的分析工作, 这将是至关重要的. 从计算上的简单和概念上的清晰来考虑, 线性变换自然成为人们首先追寻的一种表示. 例如众所周知的主成分分析(PCA)、因素分析(FA)和投影追踪(PP)等线性变换方法已得到相当广泛的应用. 但以往的盲源信号分离技术大多采用基于二阶统计的方法(例如, 协方差、交叉相关以及主成分分析等). 近年, 一种在欧美和日本等国家十分流行的线性变换方法, 即独立成分分析技术引起广泛的兴趣和应用. ICA 技术的实质就是试图将一组随机变量表示成统计上独立的变量的线性组合. 基于 ICA 方法, 使得被分析信号各成分之间的统计依赖性得到最小化, 突出了源信号的本质结构. 目前, ICA 最富有前景的两项应用是盲源分离和特征提取. 本文介绍 ICA 技术在盲信号分离方面的应用, 利用信号的

高阶统计信息, 给出算法模型及其实验结果.

2 ICA 数据模型

线性 ICA 大致有三种基本模型(一般型、噪声型和无噪声型)^[1,2].

设有 M 个零均值源信号 $s_k(1), \dots, s_k(M), k = 1, 2, \dots$, 这些源信号(独立成分)对每一个离散时间 k 来说是统计上相互独立的. 许多信号(如混合语音信号)基本上符合或近似符合这种情况. 假设这些源信号都是无法直接被观测到的, 而我们所能观测到的是含有噪声的 L 种不同线性混迭信号 $x_k(1), \dots, x_k(L)$. 我们用矢量 $x_k = [x_k(1), \dots, x_k(L)]^T$ 表示 L 维第 k 个观测数据矢量, 它是源信号和噪声在离散时间点 k 时的线性混迭. 这样, ICA 数据模型可以表示为^[3]:

$$X_k = AS_k + n_k = \sum_{i=1}^M s_k(i) a(i) + n_k \quad (1)$$

其中 $S_k = [s_k(1), \dots, s_k(M)]^T$ 表示时间指数为 k 时, 由 M 个源信号组成的源矢量. $A = [a(1), \dots, a(M)]$ 为常量 $L \times M$ 混

收稿日期: 2002-04-16; 修回日期: 2004-01-08

基金项目: 国家自然科学基金重点项目(No. 10234070), 福建省自然科学基金计划资助项目(No. C0310028).

迭矩阵,其每个元素均为未知系数, A 的所有列 $a(i)$ 称为 ICA 的基矢.在实际使用中,噪声附加项 n_k 通常予以忽略,而采用无噪声 ICA 模型,因为要将噪声从源信号中完全分离几乎是不可能的.即便如此,采用无噪声模型也已证明具有足够的难度.所以,大多数研究者都采用无噪声 ICA 模型作为实际噪声模型的近似估计^[2,4]:

$$x_k = AS_k = \sum_{i=1}^M s_k(i) a(i) \quad (2)$$

对模型(2)来说,在没有关于矩阵 A 和源信号 S_k 的任何先验知识的情况下,要从观测信号 x_k 中分离出源信号几乎是不可能的,所以,对模型(2),一般附带如下几个假设^[5,6]: (1) L 应不小于源信号数 M ,通常,出于计算方便考虑,可取 $L = M$. (2) S_k 是零均值平稳随机过程,且统计上相互独立,或尽可能相互独立;最多只允许其中一个成分 $s_k(i)$ 具有高斯分布.这是因为无法将几个高斯源彼此分离. (3) A 为满秩矩阵,即有 $\det A \neq 0$.

ICA 的基本任务是从观测的混迭信号 x_k 中估计(分离)出源信号 S_k ,这等价于估计混迭矩阵 A .由于对矩阵 A 事先没有任何已知信息,所以,在确定展开式(2)时,独立性的假设是十分必要的.而式(2)也仅仅定义了 ICA 基矢 $a(i)$ ($i=1, 2, \dots, M$) 的方向,因为它们的大小和源信号 $s_k(i)$ 的幅度在式(2)中是可以互相交换的,即任何一个常数乘以 $s_k(i)$,都可以被矩阵 A 中对应的列除以相同的常数所抵消.此外,为了计算上的方便和使展开式(2)具有唯一性,我们可以要求每一个源信号 $s_k(i)$ 具有单位方差,或将基矢标准化为单位长度.

3 数据预处理

通常,ICA 算法需要对观测数据进行适当的预处理,如中心化和预白化处理.即便有些算法不需要进行这些预处理,但采用经预处理过的数据,计算过程的收敛性能常常会更好.本文假设式(2)中的观测数据 x_k 已经是中心化过的,即 x_k 是零均值的.而预白化处理通过将式(2)中的观测值矢量 x_k 进行如下线性变换来实现:

$$v_k = Qx_k \quad (3)$$

其中, v_k 表示第 k 个白化矢量, Q 为 $M \times L$ 白化矩阵,它的选取是使白化矢量 v_k 的各分量彼此之间是互不相关的,并都具有单位方差.这样, v_k 的相关矩阵(协方差矩阵)成为单位矩阵,即 $E\{v_k v_k^T\} = I$.这种变换一般总是可能的并且可以通过主成分分析(PCA)方法来实现^[5,6].如果 $L > M$,矩阵 Q 将同时把数据矢量 x_k 的维数从 L 降到 M ,此 M 即为独立成分数.当然,噪声的影响也相应被减弱.变换之后,由式(2)和(3)并取 $L = M$ 有:

$$v_k = Qx_k = QAS_k = BS_k \quad (4)$$

其中,矩阵 $B = QA$ 称为分离矩阵,是 $M \times M$ 正交矩阵,因为基于 S_k 的统计独立假设和对 v_k 的要求,有:

$$E\{v_k v_k^T\} = BE\{S_k S_k^T\} B^T = BB^T = I \quad (5)$$

这样一来,我们把寻求一个任意的满秩矩阵 A 的问题简化为寻求一个正交矩阵 B .同时,比较式(2)和式(4),注意到观测

数据矢量现在已经表示为 v_k .由式(4)我们有:

$$S_k = B^T v_k \quad (6)$$

一旦 B 找到,我们就可以从观测的 v_k 解决独立成分分离问题.如果 B 的第 i 列元素表示为 $b(i)$,那么,第 i 个独立成分便可以由 $s_k(i) = b(i)^T v_k$ 中计算出.而 B 的所有列元素 $b(i)$ 构成 ICA 的基矢.

4 算法描述

目前,流行的 ICA 算法大致可以分为两大类:一类是基于批运算某种相关的判据函数^[2],这些算法一般需要进行复杂的矩阵或张量运算.另一类是基于随机梯度方法的自适应算法^[7],这些算法可以用神经网络实现,主要问题是运算收敛速度慢,且其学习速率参数的正确选取对收敛起决定性的作用.鉴于上述诸多算法的研究分析,并根据盲源分离的基本任务,本文给出的 ICA 算法,是通过寻求观测变量的线性组合的四阶累积量(即 kurtosis 系数)局部极值的快速神经算法,称为定点迭代方法^[8].为此,先介绍有关随机变量的 kurtosis 系数的性质.

4.1 Kurtosis 系数

kurtosis 系数是信号的高阶统计量.利用 kurtosis 系数来解决 ICA 问题已经是一个十分流行的方案.对第 i 个零均值随机信号 $s(i)$, kurtosis 系数定义为^[6]:

$$\text{kurt}[s(i)] = E\{s(i)^4\} - 3[E\{s(i)^2\}]^2 \quad (7)$$

如果 $s(i)$ 是高斯随机信号,则 kurtosis 系数为零.随机信号超高斯分布时,其 kurtosis 为正数;次高斯分布时, kurtosis 为负数.对两个独立随机信号 $s(i)$ 和 $s(j)$,有 $\text{kurt}[s(i) + s(j)] = \text{kurt}[s(i)] + \text{kurt}[s(j)]$,对一个标量常数 β ,有 $\text{kurt}[\beta s(i)] = \beta^4 \text{kurt}[s(i)]$.

设存在观测信号 v_k 的一种线性组合 $w^T v_k$,使得它具有最大或最小的 kurtosis,且 w 的范数有界并约定为 $\|w\| = 1$.定义 $z = B^T w$,因 B 是正交矩阵,则有 $\|z\| = 1$.利用式(4)和 kurtosis 性质有:

$$\begin{aligned} \text{kurt}(w^T v_k) &= \text{kurt}(w^T B S_k) = \text{kurt}(z^T S_k) \\ &= \sum_{i=1}^M z_i^4 \text{kurt}[S_k(i)] \end{aligned} \quad (8)$$

在 $\|w\| = \|z\| = 1$ 条件下,式(8)存在若干局部极大和极小值.为简单起见,不妨假设式(8)右边关于矩量求和项中,至少包含一个独立成分,其 kurtosis 系数为正,并且至少包含一个独立成分,其 kurtosis 系数为负.则式(8)中的极值点将是规范基矢 $z = \pm e_j$,即,除了一个矢量元素为 ± 1 ,其余均为零^[9].相应地有, $w = Bz = Be_j = b(j)$ (可能还带一个负号),即, w 由正交矩阵 B 的一列元素构成.可见,在给定的约束条件下,通过极大化或极小化式(8)中的 kurtosis 系数,矩阵 B 的所有列元素便可以通过求 w 来获得.而线性组合本身便是独立成分之一:

$$w^T v_k = b(i)^T v_k = S_k(i) \quad (9)$$

这样,式(8)其实成为本问题的一个目标函数(objective function).同时显示,高斯成分不能用这种方法来分离,因为它们 $\text{kurt}[s(i)] = 0$.

4.2 算法模型

为了最大化或最小化 $kurt(w^T v_k)$, 采用基于梯度下降或上升的神经算法, 这时, w 可以理解为一个具有输入矢量为 v_k 的神经元的权矢量. 由于输入 v_k 已经预白化, 所以目标函数可以在式(8)的基础上简化为:

$$kurt(w^T v_k) = E\{(w^T v_k)^4\} - 3[E\{(w^T v_k)^2\}]^2 = E\{(w^T v_k)^4\} - 3\|w\|^4 \quad (10)$$

$E\{ \cdot \}$ 表示数学期望. 但考虑到约束条件 $\|w\| = 1$. 为此增加一个惩罚项 F , 则目标函数用 $G(w)$ 表示为:

$$G(w) = E\{(w^T v_k)^4\} - 3\|w\|^4 + F(\|w\|^2) \quad (11)$$

设 v_k 为观测时间序列, $\mu(k)$ 为学习速率因子序列, 并令 f 为 $F/2$ 对 w 的导数, 则基于梯度下降(或上升)的在线学习算法模型为:

$$w(k+1) = w(k) \pm \mu(k) [E\{v_k(w(k)^T v_k)^3\} - 3\|w(k)\|^2 w(k) + f(\|w(k)\|^2) w(k)] \quad (12)$$

其中, 括号[.]内各项由 $G(w)$ 对 w 的梯度求得. 括号[.]前的正号“+”意味着找到局部极大, 负号“-”则意味着找到局部极小. 式(12)模型的缺点在于学习收敛速度慢, 且依赖学习因子 $\mu(k)$ 的选择. 因此需要寻求更快速且可靠的学习方法. 定点迭代算法便是其中之一, 其基本思想是: 使学习规则(12)中 w 的变化为零(定点), 权的调整由[.]内各项来确定, 即

$$E\{v_k(w(k)^T v_k)^3\} - 3\|w(k)\|^2 w(k) + f(\|w(k)\|^2) w(k) = 0 \quad (13)$$

注意到, 只要 $F(\|w\|^2)$ 仅仅是 $\|w\|^2$ 的函数, 则其对 w 的梯度应为一个标量常数 c 乘以 w , 即有 $f(\|w\|^2) = c$. 故式(13)可以写成:

$$w(k+1) = c \times (E\{v_k(w(k)^T v_k)^3\} - 3\|w(k)\|^2 w(k)) \quad (14)$$

实际上我们知道, 对权矢量 w 来说, 其范数的大小是无关紧要的, 而其方向才是重要的, 所以, 式(14)右侧的常数 c 可以令其为 $c=1$. 又考虑约束条件 $\|w\| = 1$, 则定点迭代算法模型的最最终形式为:

$$w(k+1) = E\{v_k(w(k)^T v_k)^3\} - 3w(k) \quad (15)$$

4.3 算法实现

设 v_k 表示待进行盲源分离的随机观测矢量, 它是符合式

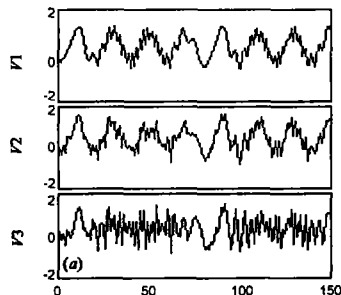
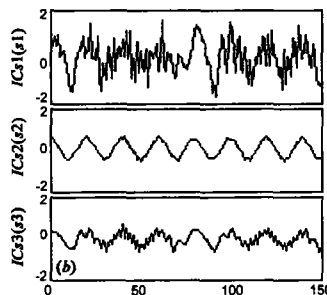


图 1 (a)三个观测信号



(b)混迭在一起的独立源信号

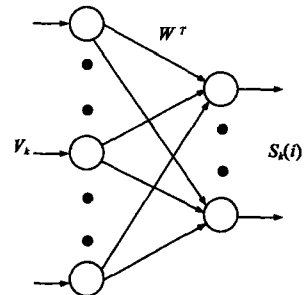


图 2 二层神经网络模型

(4)的若干未知独立成分的线性混迭信号集. 利用前面讨论的结果, 我们给出快速定点迭代算法步骤如下

(1)初始化: 置循环变量 $k = 1$, 随机地初始化权矢量 $w(k) = rand()$.

(2) $w(k)$ 标准化: $w = w / \|w\|$, $w_0 = 0$.

(3)迭代运算: 收敛检测(w 的两个方向), while $\|w(k) - w_0\| > \epsilon \wedge \|w(k) + w_0\| > \epsilon$, 其中, ϵ 是小于 1 的常数. $w_0 = w(k)$.

$$w(k+1) = E\{v_k(w(k)^T v_k)^3\} - 3w(k).$$

(4)重新标准化 $w, w = w / \|w\|$.

(5)循环直到结束.

注意到步骤(3)中收敛检测采取双向检测, 原因是式(14)中标量常数 c 既可以是正数也可能是负数. 该算法最终得到的权矢量 $w(k)$ 便是正交矩阵 B 的某一行元素 $b(i)$, 即为 ICA 基矢(记为 ICA_b). 在盲源分离情况下, 这意味着 $w(k)$ 分离了非高斯独立源信号之一, 即式(9).

为了估计(分离) M 个独立成分, 需要循环运行上述算法 M 遍. 为了保证每次分离的是不同的独立成分, 需要在上述循环迭代的基础上再增加一个简单的正交投影操作. 由于观测数据 v_k 已经预白化处理, 所以分离矩阵 B 的每一列元素都是正交的, 因此, 我们可以通过 Gram-Schmidt 正交化方法, 将当前的 w 投影到“早先”确定的 w 所张开的子空间之上. 即“早先”确定的, $b(i)$ 张成的子空间之上. 表示为: $w = w - BB^T w$. 并插入到上述步骤(2)和步骤(4)的开始处参加运算, 即可分离出 M 个独立成分.

5 实验结果

图 1(a)是三个源信号 s_1, s_2, s_3 混迭后得到的观测信号 v_1, v_2, v_3 ; 其中, $s_1(k) = \sin(20k + 0.04k^2)$, $s_2(k) = 2.0 \sin(\pi k/10)$, $s_3(k)$ 是 $[-1, 1]$ 区间均匀白噪声. 观测信号事先都经过预白化处理, 图 1(b)为分离出的源信号 ICA_1, ICA_2 和 ICA_3 , 即 s_1, s_2, s_3 . 可以看出, 分离出的信号次序发生了变化, 这是该算法的特点, 但其波形却能较真实地反映源信号的波形, 说明该算法的有效性之高. 相应于式(9)的神经网络模型是一个二层网络, 如图 2 所示.

6 结论

本文介绍的 ICA 算法是一种快速算法. 与传统的主成分

分析(PCA)方法和早期 Delfosse^[9]等人提出的基于随机梯度 ICA 算法相比最大的优势就是计算速度快. 传统的 PCA 方法基于最小平方技术, 但没有丝毫的鲁棒性. 而近年出现的鲁棒

PCA 则涉及数据样本训练预处理环节,限制了处理速度.而基于随机梯度的神经网络方法则涉及网络学习速率因子选择难和网络收敛速度慢的问题.因此,本文介绍的方法和其它盲分离算法相比具有:(1)算法的收敛速度极快,一般迭代次数仅 5~10.比 PCA 和随机梯度 ICA 至少快 10 倍以上,因此可以实现实时处理.(2)无需学习速率因子或其它动态调整参数,且可靠性高.(3)能同时分离多个盲源信号(包括超高斯信号和亚高斯信号),只需事先设定分离矩阵 B 的阶数.研究表明,ICA 算法在盲信号分离和信号特征提取方面已显示出它的巨大优势,在无导师神经网络学习和统计信号处理等方面具有广阔的应用前景,值得我们进一步探讨.

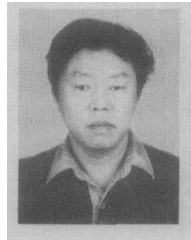
参考文献:

- [1] C Jutten, J Herault. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture [J]. Signal Processing, 1991, 24(1): 1 - 10.
- [2] P Comon. Independent component analysis-a new concept? [J]. Signal processing, 1994, 36(3): 287 - 314.
- [3] J Karhunen, E Oja, L Wang, R Vigario, J Joutsensalo. A class of neural networks for independent component analysis [J]. IEEE Trans. on Neural Networks, 1997, 8(3): 486 - 504.
- [4] A Hyvarinen, E Oja. Simple neuron models for independent component analysis [J]. Int. Journal of Neural Systems, 1996, 7(6): 671 - 687.
- [5] J Karkunen, L Wang, R Vigario. Nonlinear PCA type approaches for source separation and independent component analysis [C]. Proceedings of the '95 IEEE International Conference on Neural Networks, IC-

NN '95, Perth, Australia, 1995. 995 - 1000.

- [6] J Karhunen. Neural approaches to independent component analysis and source separation [C]. In ESANN'96, European Symposium on Artificial Neural Networks, Bruges, Belgium, 1996. 249 - 266.
- [7] A J Bell, T J Sejnowski. An information-maximization approach to blind separation and blind deconvolution [J]. Neural Computation, 1995, 7 (6): 1129 - 1159.
- [8] A Hyvarinen. A family of fixed-point algorithms for independent component analysis [A]. In Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97) [C], Munich, Germany, IEEE, 1997. 3917 - 3920.
- [9] N Delfosse, P Loubaton. Adaptive blind separation of independent sources: a deflation approach [J]. Signal Processing, 1995, 45(1): 59 - 83.

作者简介:



游菜义 男, 1957 年 7 月出生于福建建阳市, 1990 年在福建师范大学获得硕士学位, 现为集美大学副教授、厦门大学物理系博士研究生, 现主要从事智能信号处理、遥感图像处理、神经网络应用等研究工作。

陈忠 男, 1965 年 1 月出生于福建省福清市, 厦门大学物理系教授, 博士生导师, 主要从事信号处理, 神经网络应用等研究工作。