

# 一种基于内积运算的 ICA 新算法<sup>\*</sup>)

江宇闻 朱思铭

(中山大学数学系 广州 510275)

**摘要** 独立成分分析(ICA)是近几年发展起来用于解决盲源信号分离(Blind Source Separation)的一种基于信号高阶统计特性的分析方法。本文提出了一种基于向量间内积运算的解决独立成分分析问题的新算法,称之为 I-ICA。I-ICA 将混合数据在新的坐标轴上进行聚类,最后推断出原混合矩阵  $A'$ ,再通过  $A'^{-1}$  和已有的观测数据计算得到源信号。

**关键词** 独立成分分析,盲源信号分离,内积运算

## A New ICA Algorithm Based on Inner Product

JIANG Yu-Wen ZHU Si-Ming

(Department of Mathematics, Sun Yat-Sen University, Guangzhou 510275)

**Abstract** Independent Component Analysis (ICA) is developed in recent years, which is a solution to Blind Source Separation problem basing on the high-order statistics features of signals. This paper presents a new approach basing on inner product of vectors to ICA, named I-ICA. I-ICA uses mixed data to cluster in new coordinates, eventually infers the mixed matrix  $A'$ , and then computes the source signals through  $A'^{-1}$  and observed data.

**Keywords** ICA, Blind source separation, Inner product

### 1 简介

独立成分分析(Independent Component Analysis, ICA)是近几年发展起来的用于解决盲源信号分离(Blind Source Separation)的一种基于信号高阶统计特性的分析方法,可以应用于许多领域,如信号处理、生物医学、分子学、电子通信、声音处理、航天物理、纹理结构等,并且在生物医学方面取得了很大的应用成功。在 EGG 信号、fMRI 信号的分析处理方面,应用 ICA 方法进行分析,取得了令人满意的结果。

Jutten 和 Herault<sup>[1]</sup>首先在其文章中提出了 Independent Components Analysis(INCA)这个概念,其最初的目的是为了解决盲源信号分离的问题。INCA 后来被统一地称为 Independent Component Analysis(ICA),并且成为 BSS 问题的代名词。ICA 的经典解法及概念由 Sejnowski<sup>[2,3]</sup>提出,作者采用了最大化信息熵(Informax)的方法来解决 ICA 问题。随后 Amari<sup>[4]</sup>和 Cardoso<sup>[5]</sup>各自得出了具有相同更新方程的解法。于是,他们把各自的解法命名为“自然梯度法”和“相对梯度法”。此类基于梯度下降的学习方法被认为是 ICA 的一个里程碑。Puntonet 和 Prieto<sup>[7]</sup>提出了解决 ICA 问题的几何算法(Geo-ICA),该算法在计算机上非常容易实现,已经被用于生物弱电信号处理上。Hyvarinen<sup>[6]</sup>提出了基于不动点原理的 ICA 算法,由于 FastICA 的学习收敛速度非常之快,因此它是目前在实际运用中被应用得最为广泛的 ICA 算法。

目前,解决 ICA 问题用到的方法包括:极大似然法(Maximum-Likelihood)、基于峰度(Kurtosis)的学习算法、自然梯度法(Nature Gradient)、极大熵法(Entropy Maximization)、负熵法(Negentropy)、不动点快速解法(Fast Fixed-point)等。

本文提出了一种基于向量间的内积的 ICA 算法,称为 I-ICA。第 2 节简单介绍了 ICA 模型和 I-ICA 算法的思想;第 3 节详细给出了 I-ICA 算法;第 4 节采用人工产生的随机数据

进行了混合-分离实验;最后就 I-ICA 进行了总结。

### 2 模型及 I-ICA 算法

#### 2.1 ICA 模型

在一般的盲源信号分离(BSS)中,一个由  $n$  个相互独立的信号源产生的随机向量  $S: \Omega \rightarrow R^n$ ,通过一个混合矩阵  $A = (a_1, \dots, a_n)^T \in Mat(m \times n)$  生成了混合向量  $X: \Omega \rightarrow R^n$ :

$$X = AS$$

其中  $\Omega$  表示一个特定的概率空间。所谓的盲源信号分离是指:仅仅知道混合向量  $X$ ,要去通过  $X$  来得到源信号向量  $S$  和混合矩阵  $A$ 。我们令  $a_i = Ae_i$  表示矩阵  $A$  的列向量,其中  $i$  表示单位向量。我们设矩阵  $A$  是满秩的,任取两个列向量  $i, j$  都是线性独立的,  $i \neq j$ 。

#### 2.2 I-ICA 算法

I-ICA 与 Geo-ICA<sup>[7,8]</sup> 都是采用聚类方法进行学习的。I-ICA 的基本思想是在源信号空间  $\{s_1, \dots, s_n\} \in R^n$  中,混合矩阵  $A$  可以看作一个坐标变换矩阵。将源信号影射成在新的坐标系上的坐标,即把混合矩阵  $A$  看作是对源数据进行了一次坐标变换。因此,找到了  $n$  个变换后的坐标轴,就可以推断出混合矩阵  $A'$ 。基于这种思想, I-ICA 采用了类似 Kohonen 自组织图这样的非监督竞争学习方法进行数据点的聚类计算。然而,步长变化是依赖于向量变化的方向而不是依赖于距离,而且所有的更新都发生在单位球面上。

首先在单位球面上随机选取  $2n$  个单位向量  $w_1, \dots, w_n, w'_1, \dots, w'_n$ , 其中  $w_i = -w'_i$ 。然后选取一个观测样本  $x(t) \in R^n$ , 将此样本投影到单位球面上,然后比较  $w_i, w'_i (i=1, \dots, n)$  与  $x(t)$  在球面上投影之间的欧氏距离,找出距离最近的  $w_i$  或  $w'_i$ , 然后按照更新规律进行变化。

I-ICA 将  $n$  维空间中向量间的内积作为比较向量间的距离。在空间上的两个向量  $\alpha, \beta$  间的内积可以定义为:

<sup>\*</sup>) 本文得到国家自然科学基金资助(10371135)。江宇闻 博士研究生,主要研究方向:计算机网络、人工智能、动力系统;朱思铭 博士生导师,主要研究方向:计算机网络与人工智能、动力系统、混沌理论。

$$\alpha \cdot \beta = |\alpha| |\beta| \cos \theta_{\alpha\beta}$$

其中  $\theta_{\alpha\beta}$  表示向量  $\alpha$  和  $\beta$  间的方向角。

当把  $\alpha$  和  $\beta$  投影到单位球面上时, 则  $|\alpha| = |\beta|$ , 即它们之间的距离可以完全通过内积来决定。如果这两个向量重合, 则它们的内积最大; 如果它们正交, 则它们的内积为零。因此, 要求空间中与目标向量间距离最短的向量就等价于把这些向量投影到单位球面上后求出与目标向量间内积最大的那个。同时, 内积运算还满足交换律, 即  $\alpha \cdot \beta = \beta \cdot \alpha$ 。

Geo-ICA<sup>[7]</sup> 在比较  $w_i, w'_i (i=1, \dots, n)$  与  $x(t)$  在球面上的投影时需要计算向量间的欧氏距离, 在处理高维数据时付出计算量和存储空间的代价是很大的。另外, 当维数增加的时候, 需要准确地恢复混合矩阵所需要的样本数目将会呈指数增长。

由于上述的特性, 使得 I-ICA 在计算其向量的距离的时候不需要做 Geo-ICA 的符号变换和置换操作, 因此 I-ICA 只需要做  $n$  次内积运算而不是 Geo-ICA 做的  $2^n - 1$  次欧氏距离计算, 有效提高了运算的速度。

### 3 I-ICA 算法描述

I-ICA 算法可以归纳为下面的步骤:

Step1: 在  $n$  维单位球面上选择  $n$  个向量, 组成一个权重矩阵  $W_0 = [w_1, w_2, \dots, w_n]$ 。在该矩阵初始化的时候, 初始的点可以选择球面上的单位向量。

Step2: 在观测向量组中选择一个  $n$  维的样本  $x_i$ , 使得  $x_i \neq 0, i=1, 2, \dots, N$ 。将选出的样本  $x_i$  投影到单位球上, 即

$$y_i = \frac{x_i}{\|x_i\|}, i=1, 2, \dots, N$$

Step3: 分别计算  $y_i$  与  $W_0$  的列向量间的内积。通过下面的公式选出与  $y_i$  距离最近的列向量:

$$w_j^* = \arg \max_{w_j} |y_i^T \cdot w_j|$$

$$j=1, 2, \dots, n; i=1, 2, \dots, N$$

Step4: 根据距离最近的原则进行迭代更新

$$w_j(k+1)^* = \begin{cases} \Psi(w_j(k)^*) + \eta(t) \frac{\text{sgn}(y_i - w_j(k)^*)}{|y_i - w_j(k)^*|}, & y_i^T \cdot w_j^* > 0; \\ \Psi(w_j(k)^*) + \eta(t) \frac{\text{sgn}(-y_i - w_j(k)^*)}{|y_i - w_j(k)^*|}, & y_i^T \cdot w_j^* < 0; \end{cases}$$

其中  $\eta(t)$  是一个依赖于迭代次数  $t$  的学习率。  $\Psi(\cdot)$  是将权重矩阵中的列向量投影到单位圆上的非线性方程。

Step5: 更新权重矩阵:

$$\Delta W_i = W_i - W_{i-1}; i=1, 2, \dots, N$$

Step6: 若  $|\Delta W_i| < \epsilon$ , 其中  $\epsilon > 0$ , 则该算法收敛到它的结束状态, 最后得到的  $W_i$  即为所求的基向量矩阵。

Step7: 若算法未达到收敛条件, 则重复进行 2~6 步运算, 直到达到算法收敛。

最后得到权重矩阵  $W$ , 则:  $s' = W^{-1} \cdot x$ 。  $s'$  为所要恢复的源向量。

### 4 实验结果

为了研究 I-ICA 的表现, 我们随机产生两组 10000 个样本点的信号, 其中一个信号服从 Gaussian 分布, 另一个信号服从 Uniform 分布, 采用矩阵  $A$  进行混合:

$$A = \begin{pmatrix} 0.8 & 0.3 \\ 1.6 & -0.2 \end{pmatrix}$$

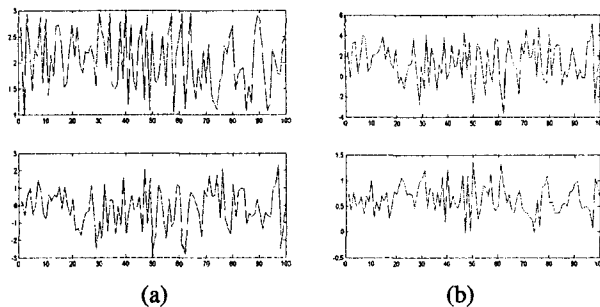


图 1 (a)一组服从 Gaussian 分布和一组服从 Uniform 分布的测试数据; (b)两组数据经过矩阵  $A$  进行混合的结果。

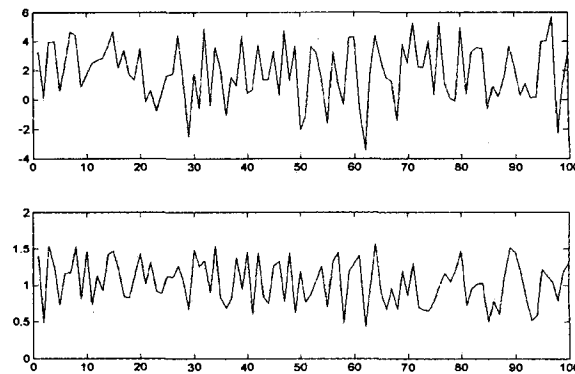


图 2 使用 I-ICA 进行分离的结果

经过大约 30 次的学习后, I-ICA 可以得到分离的结果。图 1(a)和(b)分别表示源数据和混合后的数据, 图 2 表示使用 I-ICA 进行分离后的结果。比较图 1 和图 2, 可以看出 I-ICA 可以很好地将混合的数据分离。

小结 本文提出了一种基于向量间内积运算的 ICA 算法。 I-ICA 算法的思路源自 Geo-ICA 算法, 但是在比较向量间的距离时, 采用向量间的内积来代替原来的计算欧氏距离。因此, I-ICA 在计算基向量的距离的时候不需要做 Geo-ICA 的符号变换和置换操作, 所以 I-ICA 只需要做  $n$  次内积运算, 而不是 Geo-ICA 做的  $2^n - 1$  次欧氏距离计算, 有效提高了运算的速度。从实验的结果可以看出, I-ICA 可以很好地对混合的数据进行分离。

### 参考文献

- 1 Jutten C, Herault J. Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. *Singal Processing*, 1991, 24(1): 1~10
- 2 Bell A J, Sejnowski T J. Blind separation and blind deconvolution: An information-theoretic approach. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1995. 3415~3418
- 3 Bell A J, Sejnowski T J. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 1995, 7, 1129~1159
- 4 Amari S I. Natural gradient works efficiently in learning. *Neural Computation*, 1998, 10, 251~276
- 5 Cardoso J F. The invariant approach to source separation. In: *International Symposium on Nonlinear Theory and ites Application (NoLTA)*, 1995. 55~60
- 6 Bingham E, Hyvarinen A. A fast fixed-point algorithm for independent component analysis of complex valued signals. 2000
- 7 Puntonet C, Prieto A. An adaptive geometrical procedure for blind separation of sources. *Neural Processing Letters*, 1995, 2
- 8 Theis F J, Jung A, Puntonet G. Linnear Geometric ICA: Fundamentals and Algorithms. *Neural Computation*, 2003, 15, 419~439