

含噪独立分量分析的期望最大化算法

张和发, 李立萍

(电子科技大学电子工程学院 成都 611731)

【摘要】把期望最大化(EM)算法应用到含噪ICA模型中,即假定源信号具有统计独立性,并将其放在贝叶斯估计框架中,提出一种解决含噪独立分量分析(ICA)的期望最大化(EM)算法。在含噪ICA模型中,假设源信号的均值和方差服从更为一般的均匀分布,提出的EM算法将混合矩阵和超参数交替进行处理,可以有效地估计混合矩阵和超参数在一定模型下的模型参数,从而能够估计出源信号。仿真结果说明,该方法能够很好地解决含有噪声ICA模型下的盲源分离问题。

关键词 贝叶斯方法; 盲源分离; 期望最大化算法; 独立分量分析

中图分类号 TN912

文献标识码 A

doi:10.3969/j.issn.1001-0548.2012.04.009

Expectation-Maximization Algorithm for Noisy Independent Component Analysis

ZHANG He-fa and LI Li-ping

(School of Electronic Engineering, University of Electronic Science and Technology of China Chengdu 611731)

Abstract Expectation-maximization (EM) algorithm is applied in the noisy independent component analysis (ICA) model, i.e., the source signals are assumed statistical independent and formulated in a Bayesian estimation framework. A Bayesian approach with EM algorithm for noisy ICA is proposed. In the noisy ICA model, supposing the means and variances of source signals are uniform, the proposed EM algorithm can efficiently estimate the model parameters of the mixing matrix and hyperparameters under a certain model, and then estimate the sources by processing the mixing matrix and hyperparameters alternatively. Simulation results show that the proposed method can perform blind source separation (BSS) with the noisy ICA model.

Key words Bayesian approach; blind source separation; expectation-maximum algorithm; independent component analysis

盲源分离(BSS)问题是一个非常病态的反问题,在没有先验信息的情况下是不可解的。目前,人们已经提出了很多解决盲源分离问题的方法。其中,独立分量分析(ICA)方法是基于源信号之间的相互独立性这一假设,大部分该类方法都是在没有噪声的情况下得到的,不同之处在于推断独立性的途径之间的差异性。

近年来,对ICA的研究日趋深入^[1-5]。但研究工作都是在没有噪声,并且只假设源信号之间相互独立性的情况下进行的,没有充分利用源信号的统计模型等先验信息。通过联合先验信息或者包括附加的模型参数,可以得到应用贝叶斯定理^[6-7]实现的贝叶斯方法。该方法把盲源分离问题看作是统计推断问题,并且联合了信号模型和模型参数概率的先验信息,通过计算基于先验信息和接收的混合信号的

模型参数的后验概率来得到源信号分离算法。它结合模型中包括的信息、模型参数的先验信息和所有观测得到的信息来表示关于未知量的信息状态的后验概率。在已知源信号及噪声的统计模型时,对于含噪的ICA模型,贝叶斯方法不仅能够利用源信号的统计独立性,而且还可以利用源信号及噪声的统计模型特性等先验信息,估计出模型参数从而准确地估计出源信号。贝叶斯方法中的期望最大化(EM)算法^[8]可以从非完整数据集中对参数进行极大似然估计(MLE),在解决包含非完整数据的统计估计和诸如混合估计^[9]等类似的问题中是一个有效的工具。文献[10-11]研究了高斯混合信号的梯度学习和惩罚EM源信号分离方法,并且具有一定的分离效果。文献[8]仅仅给出了一个整体上的框架,对不同问题的研究还需做具体的分析;文献[10]采用的是基

于梯度的学习算法,计算量比较大;文献[11]采用惩罚EM算法,其假设源信号的方差服从逆伽玛分布,这使得惩罚EM算法的性能对初始条件具有敏感性。

本文把期望最大化(EM)算法应用到含噪ICA模型中,即假定源信号具有统计独立性,并将其放在贝叶斯估计框架中,提出一种解决含噪独立分量分析(ICA)的期望最大化(EM)算法。在含噪ICA模型中,假设源信号的均值和方差服从更为一般的均匀分布,提出的EM算法将混合矩阵和超参数交替进行处理,这样可以估计混合矩阵和超参数在一定模型下的模型参数,从而能够估计出源信号,对初始条件稳健,其计算量比基于梯度的算法要小。仿真结果说明了所提出的方法能够很好地解决含有噪声ICA模型下的盲源分离问题。

1 盲源分离(BSS)模型和贝叶斯原理

考虑如下问题,即从含有噪声的线性混合信号:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

中找出混合的源信号 $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_M(t)]^T$ 。假设观测信号和源信号的个数都为 M , \mathbf{A} 是一个 $M \times M$ 维未知混合矩阵, $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T$ 是加性噪声。

当只考虑观测信号的有限样本时,可以把数据模型写成如下的矩阵形式:

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (2)$$

每个时间指标矩阵包含一个采样序列向量,即:

$$\mathbf{S} = [\mathbf{s}(1), \mathbf{s}(2), \dots, \mathbf{s}(T)] \quad (3)$$

由于混合矩阵和源信号之间一般不存在什么联系,因此可以假设混合矩阵 \mathbf{A} 和源信号 \mathbf{S} 之间是相互独立的。

在BSS的贝叶斯方法中,首先选择一个描述如何从发射的源信号 \mathbf{S} 产生观测的混合信号 \mathbf{X} 的信号模型;然后需要一个概率分布来量化假设的源信号对观测的混合信号影响的可能性。对于诸如式(2)的模型,可以指定 $p(\mathbf{X} | \mathbf{A}, \mathbf{S}) = p_n(\mathbf{X} - \mathbf{A}\mathbf{S})$, 其中 p_n 是加性噪声的概率。其次,依据已有的信息给模型参数指定先验概率,如指定源信号概率为 $p(\mathbf{S})$, 混合矩阵 \mathbf{A} 的元素的概率为 $p(\mathbf{A})$, 可以得到所有未知模型参数的后验概率:

$$p(\mathbf{A}, \mathbf{S} | \mathbf{X}) = \frac{p(\mathbf{X} | \mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S})}{p(\mathbf{X})} \propto \frac{p(\mathbf{X} | \mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S})}{p(\mathbf{X})} \quad (4)$$

其中, \propto 表示成比例,且:

$$p(\mathbf{X}) = \iint p(\mathbf{X} | \mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S})d\mathbf{A}d\mathbf{S} \quad (5)$$

式(5)表示了考虑观测信号后未知参数的信息状态。最后,推断出 \mathbf{A} 和 \mathbf{S} 的值或者所有的联合参数。

贝叶斯定理给出了所有未知参数的后验概率。运用一些诸如模、均值、中值的统计量,也可以通过定义一个相对高概率的区域^[12]来刻画后验概率。通常找出后验概率的模要容易一些。

2 含噪ICA的EM算法

在BSS问题中,至少有两个未知量 \mathbf{A} 和 \mathbf{S} , 它们构成了一个高维空间。可以用后验概率 $p(\mathbf{A}, \mathbf{S} | \mathbf{X})$ 估计模型参数 (\mathbf{A}, \mathbf{S}) 的值。 (\mathbf{A}, \mathbf{S}) 的估计值为:

$$(\hat{\mathbf{A}}, \hat{\mathbf{S}}) = \arg \max_{(\mathbf{A}, \mathbf{S})} \{J(\mathbf{A}, \mathbf{S}) = \ln p(\mathbf{A}, \mathbf{S} | \mathbf{X})\} \quad (6)$$

先验概率 $p(\mathbf{A})$ 是对于混合矩阵可能值知识的一种量化。假设源信号 $\mathbf{S} = [\mathbf{s}(1), \mathbf{s}(2), \dots, \mathbf{s}(T)]$ 相互独立且时不变,把幅度分布 $p(\mathbf{S})$ 的先验知识写成如下的形式:

$$p(\mathbf{S}) = p(s_{1,2,\dots,T}) = \prod_i p(s(t)) = \prod_i \prod_j p_j(s_j(t)) \quad (7)$$

式中, p_j 是第 j 个源信号 $s_j(t)$ 幅度的先验概率。

在选定源信号模型 $p(\mathbf{S})$ 后,联合分布函数就取决于参数集 θ_s 。因为 θ_s 对于描述模型参数是额外的参数,所以将其定义为超参数。这样联合源信号的先验可以表示为:

$$p(\mathbf{S}, \theta_s) = p(\mathbf{S} | \theta_s)p(\theta_s) \quad (8)$$

该情形与混合矩阵 $p(\mathbf{A} | \theta_A)$ 和噪声 $p(\mathbf{N} | \theta_n)$ 是一样的,从而影响似然函数 $p(\mathbf{X} | \mathbf{A}, \mathbf{S}, \theta_n)$ 。把这些超参数放到一起并用一个向量符号 $\theta = (\theta_n, \theta_s, \theta_A)$ 表示。在监督式学习中,可以假定这些超参数是已知的。但是,在通常情况下,超参数则被认为是未知而需要估计的。如上面的联合源信号的先验所指出的,贝叶斯方法需要为其指定一个先验概率 $p(\theta)$, 其可以分解为一连串独立的项为:

$$p(\mathbf{A}, \mathbf{S}, \theta) \propto p(\mathbf{X} | \mathbf{A}, \mathbf{S}, \theta_n)p(\mathbf{A} | \theta_A)p(\mathbf{S} | \theta_s)p(\theta) \quad (9)$$

贝叶斯方法的优点在于这些假设都具有显式,并且在不适当的情形下^[6]可以很容易地修正。

通常可以在更一般的情形下假定先验和似然性。假设观测到的误差是独立、中心化和高斯的,有如下关系式:

$$p(\mathbf{X} | \mathbf{A}, \mathbf{S}, \Sigma_n) = \prod_i \mathbf{N}(\mathbf{A}\mathbf{s}(t), \Sigma_n) \propto \exp\left[-\frac{1}{2} \sum_i [\mathbf{x}(t) - \mathbf{A}\mathbf{s}(t)]^T \Sigma_n^{-1} [\mathbf{x}(t) - \mathbf{A}\mathbf{s}(t)]\right] \quad (10)$$

矩阵 \mathbf{A} 的元素反映了源信号和探测器之间的耦合程度,典型的是高斯假设为:

$$p(\mathbf{A}_{ij}) = \mathbf{N}(M_{ij}, \sigma_{a,ij}^2) \rightarrow p(\mathbf{A} | \theta_{\mathbf{A}}) = \mathbf{N}(\mathbf{M}_{\mathbf{A}}, \Sigma_{\mathbf{A}}) \quad (11)$$

能够分离出源信号的关键是每个源信号在某些方面与其他源不同, 而这些不同能够产生分离算法的不相同。已经证明, 对于高斯分布的源信号, 如果它们在时域或谱域的性质是相同的, 那么是不可能分离出来的。对于非高斯信号也是一样, 如果它们的先验概率或者似然性相同时也是不可分的。许多非高斯模型已经得到广泛应用, 最常用的是高斯混合(MoG)^[13]模型。

高斯混合模型把源信号的幅度分布 $s_j(t)$ 描述为一些高斯信号的如下线性组合:

$$p(s_j | \alpha_{jk}, m_{jk}, \sigma_{jk}) = \sum_{k=1}^{K_j} \alpha_{jk} \mathbf{N}(m_{jk}, \sigma_{jk}^2) \quad (12)$$

其中, $\sum_{k=1}^{K_j} \alpha_{jk} = 1$, 由 $3 \times K_j$ 个参数所定义, 包括 $m_j = \{m_{jk}\}$ 和 $\sigma_j = \{\sigma_{jk}^2\}$, $k = 1, 2, \dots, K_j$ 。则源信号可以用参数 $\mathbf{K} = \{K_1, K_2, \dots, K_M\}$ 、 $\alpha = \alpha_{jk}$ 和 $\sigma = \{\sigma_j\}$ 表示, $j = 1, 2, \dots, M$ 。

高斯混合模型的一个重要性质是把 $\alpha_{jk} = P(z_j(t) = k)$ 解释为隐藏离散值变量 $z_j(t)$ 的概率, 使得:

$$\begin{cases} p(s_j(t) | z_j(t) = k, m_{jk}, \sigma_{jk}) = \mathbf{N}(m_{jk}, \sigma_{jk}^2) \\ P(z_j(t) = k) = \alpha_{jk} \end{cases} \quad (13)$$

这样, 问题简化为估计 $\mathbf{s}(t) = \{s_1(t), s_2(t), \dots, s_M(t)\}$ 、 $\mathbf{z}(t) = \{z_1(t), z_2(t), \dots, z_M(t)\}$ 以及参数 $(\mathbf{A}, \Sigma_{\mathbf{n}})$ 和 $\theta_s = (\mathbf{m}, \sigma)$ 。

当源信号是高斯混合模型时, 观测信号 $x_i(t)$ 也可以用一个包含 $K^i = \prod_j K_j$ 个分量的高斯混合模型描述:

$$p(x_i(t) | \mathbf{A}, \mathbf{z}(t), \mathbf{m}, \sigma) = \mathbf{N}\left(\sum_j A_{ij} m_{jz_j(t)}, \sum_j A_{ij} \sigma_{jz_j(t)}^2 A_{ji} + \sigma_{n_i}^2\right) \quad (14)$$

通常不能得到似然函数的解析表达式。有一类旨在得到最大似然(ML)解的方法, 期望最大化(EM)算法, 其主要思想是把源信号 \mathbf{S} 视为隐藏变量, 把观测信号 \mathbf{X} 视为不完全数据, 而把整体 (\mathbf{X}, \mathbf{S}) 视为完全数据。EM算法通过以下两步来实现: 1) E步骤(期望)包括求出观测数据 \mathbf{X} 和隐藏数据 \mathbf{S} 相对于参数 \mathbf{A} 和 θ 联合分布的似然函数, 求出其对于 \mathbf{X} 和在前面的迭代中估计的 $\hat{\mathbf{A}}$ 和 $\hat{\theta}$ 的条件期望, 2) M步骤(最大化)包含对得到的似然函数相对于参数 \mathbf{A} 和 θ 的最大化来得到其估计值 $\hat{\mathbf{A}}$ 和 $\hat{\theta}$:

$$\begin{cases} \text{E: } Q(\mathbf{A}, \theta, \hat{\mathbf{A}}, \hat{\theta}) = E_{\mathbf{x}, \mathbf{s}} \left[\ln p(\mathbf{X}, \mathbf{S}, \mathbf{A}, \theta) | \mathbf{X}, \hat{\mathbf{A}}, \hat{\theta} \right] \\ \text{M: } (\hat{\mathbf{A}}, \hat{\theta}) = \arg \max_{(\mathbf{A}, \theta)} \{Q(\mathbf{A}, \theta, \hat{\mathbf{A}}, \hat{\theta})\} \end{cases} \quad (15)$$

联合分布 $p(\mathbf{X}, \mathbf{S}, \mathbf{A}, \theta)$ 可以分解为 $p(\mathbf{X}, \mathbf{S}, \mathbf{A}, \theta) = p(\mathbf{X} | \mathbf{A}, \mathbf{S}) p(\mathbf{A}) p(\mathbf{S} | \theta) p(\theta)$ 。 $P(\mathbf{X}, \mathbf{S}, \mathbf{A}, \theta)$ 可以看作是关于 \mathbf{A} 和 θ 可分解的函数, 从而把函数 $Q(\mathbf{A}, \theta, \hat{\mathbf{A}}, \hat{\theta})$ 分解成两个分别关于 \mathbf{A} 和 θ 的函数的和:

$$Q(\mathbf{A}, \theta, \hat{\mathbf{A}}, \hat{\theta}) = Q_{\mathbf{A}}(\mathbf{A} | \hat{\mathbf{A}}, \hat{\theta}) + Q_{\theta}(\theta | \hat{\mathbf{A}}, \hat{\theta}) \quad (16)$$

其中,

$$\begin{cases} Q_{\mathbf{A}}(\mathbf{A} | \hat{\mathbf{A}}, \hat{\theta}) = E \left[\ln p(\mathbf{X} | \mathbf{A}, \mathbf{S}) + \ln p(\mathbf{A}) | \mathbf{X}, \hat{\mathbf{A}}, \hat{\theta} \right] \\ Q_{\theta}(\theta | \hat{\mathbf{A}}, \hat{\theta}) = E \left[\ln p(\mathbf{S} | \theta) + \ln p(\theta) | \mathbf{X}, \hat{\mathbf{A}}, \hat{\theta} \right] \end{cases} \quad (17)$$

当得到 \mathbf{S} 的估计时, 通过数学计算, 可以得到 \mathbf{A} 的估计值表达式为:

$$\hat{\mathbf{A}}^{(k+1)} = \text{Mat}([\mathbf{A}^* + T\hat{\mathbf{R}}_{\text{SS}}^* \Gamma^{-1} [\mathbf{A}^* \text{Vect}(\mathbf{M}) + T \text{Vect}(\hat{\mathbf{R}}_{\text{XS}})]) \quad (18)$$

式中, $\text{Vect}(\cdot)$ 是把矩阵转换为一个按行依次串联的向量; $\text{Mat}(\cdot)$ 是 $\text{Vect}(\cdot)$ 的逆变换; \mathbf{A}^* 是一个对角向量为 $\text{Vect}(\mathbf{A})$ 的 $M^2 \times M^2$ 对角矩阵; $\hat{\mathbf{R}}_{\text{SS}}^*$ 是块对角线为 $\hat{\mathbf{R}}_{\text{SS}}$ 的 $M^2 \times M^2$ 对角矩阵。

$$A_{ij} = \sigma_{\mathbf{n}}^2 / \sigma_{a,ij}^2 \quad (19)$$

$$\hat{\mathbf{R}}_{\text{XS}} = \frac{1}{T} \sum_{t=1}^T \mathbf{x}(t) \hat{\mathbf{s}}(t)^T \quad (20)$$

$$\hat{\mathbf{R}}_{\text{SS}} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{s}}(t) \hat{\mathbf{s}}(t)^T + \frac{1}{T} \sum_{t=1}^T (\mathbf{A}^T \Sigma_{\mathbf{n}}^{-1} \mathbf{A} + \Sigma_{\mathbf{z}}^{-1})^{-1} \quad (21)$$

$$\Sigma_{\mathbf{z}} = \text{diag}(\sigma_{1z_1}^2, \sigma_{2z_2}^2, \dots, \sigma_{Mz_M}^2) \quad (22)$$

首先, 通过下式估计 $\hat{z}_j(t)$ 得到数据分类更新:

$$\hat{z}_j(t) = \arg \min_{z_j(t)} \{-\ln p(z_j(t) | \mathbf{x}(t), \hat{\mathbf{A}}, \hat{\mathbf{s}}_{1 \neq j}(t), \hat{\theta})\} \quad (23)$$

当均值为均匀分布时, 求 Q_{θ} 相对于 α 和 m 的最大值, 可得:

$$\hat{\alpha}_{jz} = \frac{\sum_{t=1}^T p(z_j(t) | \mathbf{x}(t), \hat{\mathbf{s}}_{1 \neq j}(t))}{T} \quad (24)$$

$$\hat{m}_{jz} = \frac{\sum_{t=1}^T \mu_{jz}(t) p(z_j(t) | \mathbf{x}(t), \hat{\mathbf{s}}_{1 \neq j}(t))}{p(z_j(t) | \mathbf{x}(t), \hat{\mathbf{s}}_{1 \neq j}(t))} \quad (25)$$

其中,

$$\mu_{jz}(t) = E[s_j(t) | \mathbf{x}(t), z] \quad (26)$$

当方差为均匀分布时, 求 Q_{θ} 关于 σ 的最大值, 可得:

$$\hat{\sigma}_{jz} = \frac{\sum_{t=1}^T (V_{jz}(t) + \mu_{jz}^2(t) - 2\hat{m}_{jz}\mu_{jz}(t) + \hat{m}_{jz}^2)}{\sum_{t=1}^T p(z_j(t) | \mathbf{x}(t), \hat{\mathbf{s}}_{l \neq j}(t))} \quad (27)$$

其中,

$$V_{jz}(t) = E[s_j^2(t) | \mathbf{x}(t), z] \quad (28)$$

在下一节的仿真试验中,为了说明EM算法的有效性,用文献[14]所定义的归一化Amari指标 $I_A \in [0,1]$ 作为性能度量:

$$I_A = \frac{1}{2N(N-1)} \times \left[\sum_{i=1}^N \left(\sum_{j=1}^N \frac{|p_{ij}|}{\max_k |p_{ik}|} - 1 \right) + \sum_{j=1}^N \left(\sum_{i=1}^N \frac{|p_{ij}|}{\max_k |p_{kj}|} - 1 \right) \right] \quad (29)$$

式中, p_{ij} 表示矩阵 $\mathbf{P} = (\hat{\mathbf{A}})^{-1} \mathbf{A}$ 的第 (i,j) 个元素。 I_A 值越小,分离性能越好, $10\lg I_A > -10$ dB 表示分离不够充分。

3 仿真实验

为了说明本文所提算法对含噪ICA的分离性能,给出合成的高斯混合信号,每个分量含有两个均值为 ± 2 的高斯信号,方差 $\sigma^2 = 0.25$ 。这些源信号通过矩阵 $\mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ -0.7 & 1 \end{bmatrix}$ 相混合。附加方差为 $\sigma_n^2 = 0.12$ (SNR=15 dB) 的噪声,并取 $T=10\ 000$ 个观测值。

给出超参数如下: $\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,

$\Sigma_A = \begin{bmatrix} 0.01 & 100 \\ 100 & 0.01 \end{bmatrix}$, $\mathbf{H} = (\alpha_{jz}) = \begin{bmatrix} 0.3 & 0.7 \\ 0.6 & 0.4 \end{bmatrix}$ 。源信号的相空间分布如图1所示,混合信号的相空间分布如图2所示。

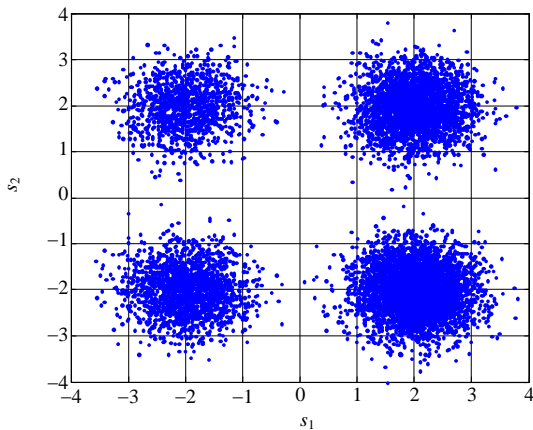


图1 源信号的相空间分布

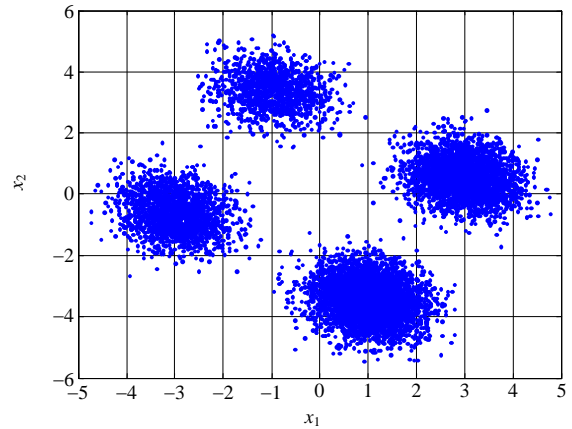


图2 混合信号的相空间分布

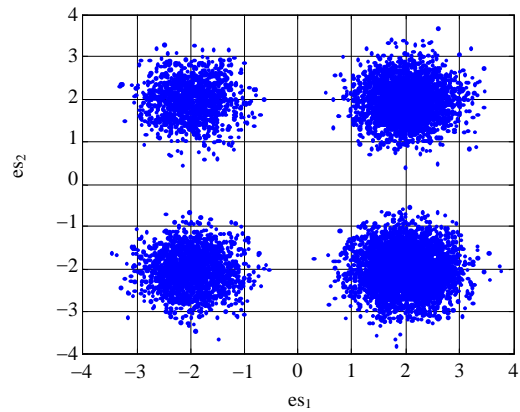


图3 估计信号的相空间分布

为了进行EM算法,首先选定初始条件: $\mathbf{A}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\Sigma_s^{(0)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $\mathbf{M}^{(0)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{S}^{(0)}$, 其中 $\mathbf{S}^{(0)}$ 的产生方式为: $\mathbf{s}_j^{(0)} \sim \sum_{z=1}^{K_j} \alpha_{jz} \mathcal{N}(\mathbf{M}_{jz}^{(0)}, \sigma_{jz}^2)^{(0)}$ 。这样,通过所提出的EM算法,可以得到分离信号。

分离后的结果如图3所示,将图3和图1及图2比较,可以看到所提出的EM算法能够收敛到正确解,从而能在含噪ICA模型中有效地分离高斯混合信号。

在上述仿真中,将初始条件选为:

$$\mathbf{A}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma_s^{(0)} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix},$$

$$\mathbf{M}^{(0)} = \begin{bmatrix} -0.5 & 0.5 \\ -0.5 & 0.5 \end{bmatrix},$$

其他条件不变,通过所提出的EM算法也能收敛到正确解。

为了定量地比较文献[11]的方法和本文所提方法分离性能差异,进行下面的实验。比较100次蒙特

卡洛实验后在前100次迭代中, 每次迭代后分离性能 $10\lg I_A$ 的平均值, 图4为两种方法在该分离性能下的比较图。从图4可以看到, 虽然文献[11]经过一定的迭代次数后能够分离出源信号, 但本文所提的方法比文献[11]中的方法具有更好的分离性能, 有近2 dB 的提高, 且比较稳定。

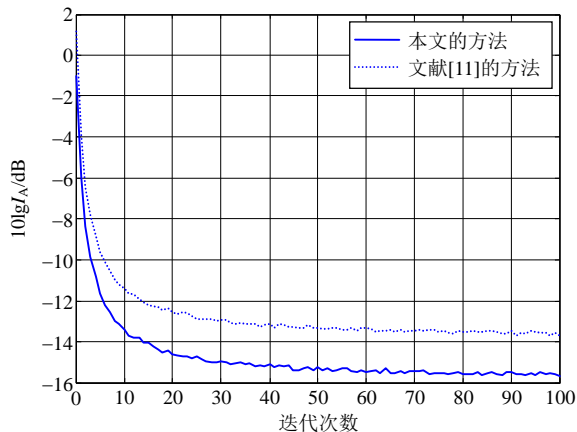


图4 分离性能比较

4 结 论

本文把期望最大化(EM)算法应用到含噪ICA模型中, 即假定源信号具有统计独立性, 并将其放在贝叶斯估计框架中, 提出一种解决含噪独立分量分析(ICA)的期望最大化(EM)算法。在含噪ICA模型中, 假设源信号的均值和方差服从更为一般的均匀分布, 提出的EM算法将混合矩阵和超参数交替进行处理, 可以有效地估计混合矩阵和超参数在一定模型下的模型参数, 从而能够估计出源信号, 对初始条件稳健, 其计算量比基于梯度的算法要小。仿真结果说明了所提出的方法能够很好地解决含有噪声ICA模型下的盲源分离问题。

参 考 文 献

[1] NOVEY M, ADALI T. On extending the complex FastICA algorithm to noncircular sources[J]. *IEEE Transactions on Signal Processing*, 2008, 56(5): 2148-2154.
 [2] NESTA F, SVAIZER P, OMOLOGO M. Convolutional BSS of short mixtures by ICA recursively regularized across frequencies[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2010, (99): 1-1.

[3] MIKHAEL W B, RANGANATHAN R, YANG T. Complex adaptive ICA employing the conjugate gradient technique for signal separation in Time-Varying flat fading channels [J]. *Circuits, Systems, and Signal Processing*, 2010, 29(3): 469-480.
 [4] TANG Y, LI J. Normalized natural gradient in independent component analysis[J]. *Signal Processing*, 2010, 90(9): 2773-2777.
 [5] 洪英, 韩周安. 偏压高斯语音信号有效分离ICA方法研究 [J]. *电子科技大学学报*, 2008, 37(5): 693-697.
 HONG Ying, HAN Zhou-an. Method research on effective separated ICA algorithm to sub-Gaussian distribution audio signal[J]. *Journal of University of Electronic Science and Technology of China*, 2008, 37(5): 693-697.
 [6] MOHAMMAD-DJAFARI A, KNUTH K. Bayesian approaches in handbook of blind Source separation, independent component analysis and Applications[M]. [S.l.]: Academic Press, 2009.
 [7] KNUTH K H, TSE M K, CHOINSKY J, et al. Bayesian source separation applied to identifying complex organic molecules in space[C]//*IEEE/SP 14th Workshop on Statistical Signal Processing*. [S.l.]: IEEE, 2007: 346-350
 [8] DEMPSTER A P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the em algorithm[J]. *Journal of the Royal Statistical Society: Series B*, 1977, 39(1): 1-38.
 [9] MCLACHLAN G, KRISHNAN T. The EM algorithm and extensions[M]. New York: John Wiley & Sons, 1996.
 [10] SNOUSSI H, MOHAMMAD-DJAFARI A. Bayesian source separation with mixture of Gaussians prior for sources and Gaussian prior for mixture coefficients[C]//*Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 20th International Workshop, AIP Conference Proceedings*. Gif-sur-Yvette, France: American Institute of Physics, 2001.
 [11] KUUSELA M, RAIKO T, HONKELA A, et al. A gradient-based algorithm competitive with variational Bayesian EM for mixture of Gaussians[C]//*Proceedings of International Joint Conference on Neural Networks*. Atlanta, Georgia, USA: [s.n.], 2009.
 [12] ROBERT C. The bayesian choice: a decision-theoretic motivation (springer texts in statistics)[M]. New York: Springer Verlag, 1996.
 [13] SNOUSSI H, MOHAMMAD-DJAFARI A. Fast joint separation and segmentation of mixed images[J]. *Journal of Electronic Imaging*, 2004, 13(2): 349-361.
 [14] AMARI S-I, CICHOCKI A, YANG H H. A new learning algorithm for blind signal separation. in *Advances in Neural Information Processing Systems*[M]. Cambridge MA: MIT Press, 1996.