

# 高斯噪声中的参数盲估计

王惠刚,李志舜

(西北工业大学航海工程学院,陕西西安 710072)

**摘 要:** 盲信号处理方法中常忽略噪声的影响,而实际问题中噪声的影响是存在的.本文主要讨论了在协方差矩阵未知的加性高斯噪声中混合系数的盲估计问题.本文以最大似然估计为基础,提出一种求解参数的最优化算法,给出了混合矩阵和协方差矩阵的计算式.采用高斯混合模型(GMM)来逼近源信号的概率密度函数,简化了算法中的积分,导出了一种基于EM算法的迭代式.仿真表明,算法不仅能稳定收敛,而且在低信噪比下也能获得良好性能.

**关键词:** 高斯噪声;盲信号处理;EM算法;高斯混合模型

**中图分类号:** TN911 **文献标识码:** A **文章编号:** 0372-2112(2003)07-0974-03

## Blind Estimation of Parameters in Gaussian Noise

WANG Hui-gang, LI Zhi-shun

(Marine Engineering College, Northwestern Polytechnic University, Xi'an, Shaanxi 710072, China)

**Abstract:** Generally, some methods of blind signal processing ignore noise, however, noise affects the performance of algorithms, especially seriously in some areas. This paper provides solutions to the problem that mixing matrix is estimated blindly in Gaussian noise with unknown covariance. Based on Maximum Likelihood estimation, the equations are given for solving the mixing matrix and covariance matrix. Gaussian Mixture Model (GMM) is used to approximate the pdf of sources and results in a practical EM algorithm. Computer simulation shows that this algorithm is convergent and has good performance in low SNR.

**Key words:** Gaussian noise; blind signal processing; EM algorithm; Gaussian mixture model

## 1 引言

盲信号分离(Blind Source Separation,简称BSS)是近几年来国内外研究的热点,它是从观测信号中恢复出统计独立的源信号,且源信号和通道参数的信息未知.这个问题在许多信号处理中会遇到,如鸡尾酒会问题、图像处理和被动声呐处理.假定 $m$ 个阵元的观测信号 $x(t)$ 来自于 $n(n \geq m)$ 个信号源 $s(t)$ 的线性混合,并包含了加性观测噪声 $n(t)$ ,忽略时间 $t$ ,其模型如下

$$x = As + n \quad (1)$$

其中 $x = (x_1, x_2, \dots, x_m)^T$ 是 $m$ 维的观测随机矢量, $s = (s_1, s_2, \dots, s_n)^T$ 是 $n$ 维的源信号矢量, $A$ 是 $m \times n$ 维的混合矩阵, $n$ 是具有高斯分布的加性白噪声.在一般的盲分离信号算法中,都假定 $A$ 是未知且为可逆的方阵(即 $m = n$ ),源信号是统计独立的而且是非高斯的,高斯噪声的影响可以忽略,通过一个分离矩阵 $W$ 来估计源信号 $\hat{s} = Wx$ .当真实分离矩阵 $W = A^{-1}$ 时,估计的源信号是真实的源信号.当矩阵 $A$ 和源信号未知时,根据不同的准则,有许多估计 $W$ 的算法,如Bell的Infomax算法<sup>[1]</sup>,Comon的独立成分分析(ICA)<sup>[2]</sup>,Amari的自然梯度算法<sup>[3]</sup>等.这些算法都是以最小互信息(Minimum Informa-

tion)为准则,以非线性函数或高阶统计量来逼近源信号的概率密度函数,在忽略噪声的前提下,这些算法的性能都是近似最优的或次最佳的.

然而实际中噪声的影响是存在的,特别是在水声信号处理中,噪声的影响更明显,因此研究噪声下盲处理方法具有广泛的应用前景.针对噪声的盲处理,A. Hyvarinen也给出了一些解决方法<sup>[4,5]</sup>,但应用受到限制.本文提出的方法是以最大似然概率为基础,提出了一种求解参数的EM(Expectation Maximum)算法,EM算法能保证迭代的稳定收敛.同时以高斯混合模型(GMM)来逼近信号的概率密度函数,当其模型参数已知时,可大大简化算法的计算复杂性.

## 2 参数的最大似然估计

假定我们获得了观测向量 $x(t)$ , $t = 1, 2, \dots, T$ 的 $T$ 个观测值,噪声 $n(t)$ 是均值为零的高斯噪声,满足正态分布 $N(0, \Sigma)$ , $\Sigma$ 是未知的噪声协方差矩阵.假定 $s(t)$ 的概率密度函数为 $f(s)$ ,则联合概率密度函数为

$$P(x, s | A, \Sigma) = P(x | s, A, \Sigma) \cdot P(s) = (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} (x - As)^T \Sigma^{-1} (x - As)\right) \cdot f(s) \quad (2)$$

其中  $| \cdot |$  定义为矩阵的行列式,  $(x - As)^{-1} = (x - As)^H$   
 $^{-1}(x - As)$ , 上标  $H$  定义为共轭转置. 观测数据的似然函数为

$$P(x|A, s) = \int P(x, s|A, s) ds = (2\pi)^{-m/2} | \Sigma |^{-1/2} f(s) \cdot e^{-\frac{1}{2}(x - As)^H \Sigma^{-1}(x - As)} ds \quad (3)$$

对式(3)进行最大化可求出参数  $A, s$  的最大似然解. 实际应用中, 更倾向于使用对数似然函数, 求解对数似然函数对参数的导数

$$\begin{aligned} \frac{\partial \ln P(x|A, s)}{\partial A} &= \frac{1}{P(x|A, s)} \frac{\partial}{\partial A} P(x, s|A, s) f(s) ds \\ &= \frac{\partial}{\partial A} \ln P(x, s|A, s) \cdot P(s|x, A, s) ds \\ &= E_{s|x} \left\{ \frac{\partial}{\partial A} \ln P(x, s|A, s) \right\} \end{aligned} \quad (4)$$

其中  $P(s|x, A, s)$  是后验概率密度函数,  $E_{s|x}$  是后验期望. 同样可推得, 对参数  $s$  也有类似的结构

$$\frac{\partial \ln P(x|A, s)}{\partial s} = E_{s|x} \left\{ \frac{\partial}{\partial s} \ln P(x, s|A, s) \right\} \quad (5)$$

从式(4), (5)可以看出, 参数  $A, s$  是与  $s(t)$  紧密相连的.

### 3 参数的期望最大化算法

期望最大化算法 (EM Algorithm) 是一种求解似然估计的迭代最优化算法, 它的每一次迭代都能保证似然函数是非递减的, 具体分析参见文献[6]. 其迭代可分为两步, 一是根据参数初始值或前一次迭代值来计算似然函数的期望 (E Step), 其次将似然函数最大化以获得新的参数值 (M Step).

(1) (E Step) 在给定的模型和观测数据下, 计算似然函数的期望值. 在上一节就已经描述了似然函数期望值的计算, 只要知道了前一次迭代的参数  $A_k$  和  $s_k$  就可根据式(3)计算似然函数.

(2) (M Step) 通过最大化似然估计来计算参数  $A$  和  $s$ :  
 $A_{k+1} = \arg \max_A P(x|A, s_k), \quad s_{k+1} = \arg \max_s P(x|A_k, s)$  (6)

由于参数  $A$  和  $s$  是确定性的未知参数, 分别求似然函数对参数  $A$  和  $s$  的导数公式(4)、(5), 并令其为零, 可求出参数的迭代式如下

$$A_{k+1} = R_{xs}^{(k)} \cdot R_{ss}^{(k)-1}; \quad s_{k+1} = R_{xx} - A_{k+1} \cdot R_{xs}^{(k)H} \quad (7)$$

其中

$$\begin{aligned} R_{xx} &= E_t \{ x(t) x(t)^H \} = 1/T \cdot \sum_{t=1}^T x(t) x(t)^H \\ R_{xs}^{(k)} &= E_t \{ x(t) \cdot E_{s|x} \{ s(t)^H | x(t), A_k, s_k \} \} \\ &= 1/T \cdot \sum_{t=1}^T x(t) E_{s|x} \{ s(t)^H | x(t), A_k, s_k \} \\ R_{ss}^{(k)} &= E_t \{ E_{s|x} \{ s(t) s(t)^H | x(t), A_k, s_k \} \} \\ &= 1/T \cdot \sum_{t=1}^T E_{s|x} \{ s(t) s(t)^H | x(t), A_k, s_k \} \end{aligned} \quad (8)$$

$E_t \{ \cdot \}$  示时间平均. 可看出, 参数  $A$  和  $s$  的最大似然估计与源信号分布密度函数紧密相关, 需要知道源信号的先验信息.

### 4 基于高斯混合模型的迭代算法

从前面给出的算法来看, 后验概率密度函数的计算必须

要求源信号的概率密度函数  $f(s)$  易于积分才行. 在盲处理中这些信息一般是未知的, 但有时可通过实验或对目标信号源的分析来获得这些信息. 常用的方法是在观测数据中估计  $f(s)$  的模型参数, 一种有效的方法是用高斯混合模型 (GMM) 来逼近信号的密度函数  $f(s)$ , 它不仅能拟合单峰的密度函数, 也能拟合多个峰值的密度函数[8].

假定用  $K$  个均值为  $\mu_i$ 、方差为  $\sigma_i^2 (i=1, \dots, K)$  的高斯函数来拟合概率密度函数, 其混合加权为  $\pi_i, i=1, \dots, K$ , 且

$\sum_{i=1}^K \pi_i = 1$ . 若设参数  $\theta_i = \{ \mu_i, \sigma_i^2 \}$ , 则信号的概率密度分布为

$$p(s|i) = \pi_i \cdot N_s(\mu_i, \sigma_i^2) \quad (9)$$

若假定每个信号源都由  $K$  个高斯函数混合而成, 则  $n$  个独立信号源的联合密度函数为

$$p(s|) = \prod_{j=1}^n \sum_{i=1}^K \pi_{ij} \cdot N_{s_j}(\mu_{ij}, \sigma_{ij}^2) \quad (10)$$

若令  $\tilde{q} = [q_1, \dots, q_n]^T, u_q = [\mu_{q_1}, \dots, \mu_{q_n}]^T, V_q = \text{diag}(\sigma_{q_1}^2, \dots, \sigma_{q_n}^2)$ , 其中  $q$  代表了所有  $q_1, \dots, q_n$  的组合, 且定义  $q = [q_1, \dots, q_n]$ . 则式(10)可写为:

$$p(s|) = \sum_q \tilde{q} \cdot N_s(u_q, V_q) \quad (11)$$

如何根据信号源的观测值来计算 GMM 参数, 在参考文献[8]中有详细的讨论, 限于篇幅, 在这里不作分析. 在后面的讨论中都假设这些参数已知. 把式(11)代入式(2)和(3), 可计算出联合似然函数

$$p(x, s|A, s) = \sum_q \tilde{q} N_s(u_q, V_q) N_x(As, s) \quad (12)$$

同样, 我们也很容易写出观测向量的似然函数

$$p(x|A, s) = \sum_q \tilde{q} N_x(Au_q, AV_q A^T + \Sigma) \quad (13)$$

由于似然函数与信号无关, 依据 Bayes 原理, 信号的后验分布  $p(s|x, A, s) = \sum_q \tilde{q} N_s((x - Au_q) + u_q, V_q - AV_q A^T)$  (14)

其中  $\Sigma = V_q A^H (\Sigma + AV_q A^H)^{-1} \Sigma$ ,  $\tilde{q} = \tilde{q} / p(x|A, s)$ . 根据式(14), 可以很容易地计算出其后验均值和二阶矩,

$$E\{s^H | x, A, s\} = \sum_q \tilde{q} [(x - Au_q)^H + u_q^H] \quad (15)$$

$$\begin{aligned} E\{ss^H | x, A, s\} &= \sum_q \tilde{q} [V_q - AV_q A^T + (x - Au_q + u_q) \\ &\quad \cdot (x - Au_q + u_q)^H] \end{aligned} \quad (16)$$

因此, 如果知道信号的参数  $\tilde{q}$ , 结合式(7)、(15)和(16), 就可通过观测值来计算混合矩阵  $A$  及噪声协方差矩阵  $\Sigma$  的迭代估计式.

在实际应用中, 并不关心协方差矩阵  $\Sigma$  的值, 更关心混合矩阵和重构的源信号. 当估计的混合矩阵  $A_{k+1}$  收敛后, 则源信号的最小均方估计为

$$s_{k+1}(t) = \text{pinv}(A_{k+1}) \cdot x(t) \quad (17)$$

为了评价算法的性能, 定义了参数估计误差  $e_A$ , 用来评价参数  $A$  的估计精度. 令  $Q(k+1) = \text{pinv}(A) \cdot A_{k+1}$ , 则误差

的定义为

$$e_A = \left| \frac{|Q_{ij}|}{\max_k |Q_{ik}|} - 1 \right|^2 + \left| \frac{|Q_{ij}|}{\max_k |Q_{kj}|} - 1 \right|^2 \quad (18)$$

当  $A_{k+1}$  是  $A$  的真实估计时,  $Q$  将逼近一个排列阵和对角阵的乘积, 估计误差  $e_A$  趋于零。

## 5 仿真实验

在计算机仿真中, 我们选择了三个信号源, 分别服从均匀分布、拉普拉斯分布和双模态 (bimodal) 分布。我们分别用 GMM 模型来拟合这三个信号源, 每个信号源由 5 个高斯成分, 即  $K=5$ ,  $K$  选取过大, 计算比较复杂, 选取过小, 逼近误差会增大, 导致算法的不稳定。分别抽取 800 个样本, 利用文 [8] 中的方法, 可以计算出对应的模型参数  $w_{ij}$ ,  $\mu_{ij}$ ,  $\sigma_{ij}^2$ 。对应的参数如表 1 所示, 由 GMM 模型产生的概率密度函数如图 1 所示。

表 1 三种分布函数对应的 GMM 参数 (5 个高斯成分)

参数	$w_{ij}$	$\mu_{ij}$	$\sigma_{ij}^2$
均匀分布	[0.220, 0.100, 0.220, 0.315, 0.145]	[-1.27, 0.094, 1.29, 0.091, -0.19]	[0.09, 0.42, 0.092, 0.422, 0.386]
双模态	[0.127, 0.166, 0.003, 0.217, 0.487]	[-0.76, -0.89, -1.14, -1.156, 1.02]	[0.73, 0.41, 0.169, 0.558, 0.12]
拉普拉斯	[0.005, 0.430, 0.365, 0.092, 0.108]	[-0.06, 0.002, 0.157, -0.05, -0.42]	[0.36, 1.91, 0.278, 0.03, 0.355]

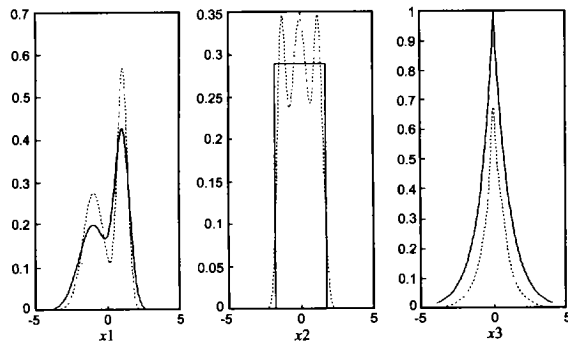


图 1 用 GMM 模型拟合的密度函数 (5 个高斯成分)。

实线为真实的分布曲线; 虚线为逼近的分布曲线

在仿真中, 假定信号和噪声都是零均值、单位方差的随机变量, 其中噪声服从高斯分布且通过系数  $\gamma$  迭加到观测信号上, 即  $x(t) = As(t) + n(t)$  系数  $\gamma$  用来调整信噪比, 信噪比的计算为

$$SNR = \frac{m}{n} \sum_{j=1}^n A_{ij}^2 / (m^2)$$

在仿真中, 选取一个  $3 \times 3$  非奇异的混合矩阵, 并假定各阵元的接收噪声是不相关的, 即协方差矩阵为对角阵

$$A = \begin{bmatrix} 0.95 & 0.486 & 0.4565 \\ 0.231 & 0.89 & 0.019 \\ 0.607 & 0.762 & 0.821 \end{bmatrix}, \quad \Sigma = \text{diag}(1, 1, 1)$$

为了使算法具有更快的收敛速度, 采用文 [2] 中给出的 ICA 算法计算的结果作为算法 (7) 的初始值。图 2 是在四个不

同信噪比下的误差收敛曲线。图 3 是在不同信噪比下的收敛稳定后的误差。计算机仿真表明, 算法在经过近 40 次的迭代后, 均能稳定收敛。

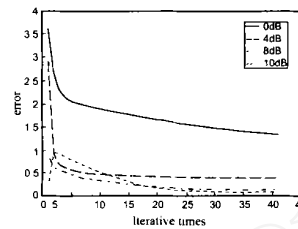


图 2 信噪比为 0、4、8、10dB 下的误差收敛曲线

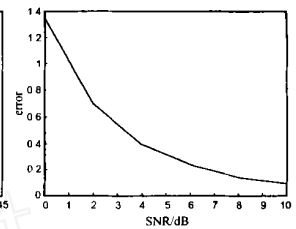


图 3 不同信噪比下的稳定收敛后的误差

## 6 结论

本文讨论了在高斯噪声下信号混合参数的盲估计问题, 以最大似然估计为基础, 给出了参数的后验估计公式。用高斯混合模型来逼近源信号的概率密度函数, 避免了积分, 简化了算法的复杂性。计算机仿真结果表明, 算法在较低信噪比下能较好地估计出混合矩阵。

## 参考文献:

- [1] A J Bell, T J Sejnowski. An information-maximization approach to blind separation and blind deconvolution [J]. Neural Computation, 1995, 7: 1129 - 1159.
- [2] P Comon. Independent component analysis: a new concept? [J]. Signal Processing, 1994, 36(3): 287 - 314.
- [3] S Amari, A Cichocki. Adaptive blind separation processing: neural network approaches [J]. Proc of IEEE, 1998, 86(10): 2026 - 2048.
- [4] A Hyvarinen. Independent component analysis in the presence of gaussian noise by maximizing joint likelihood [J]. Neurocomputing, 1998, 22: 49 - 67.
- [5] A Hyvarinen. Gaussian moments for noisy independent component analysis [J]. IEEE SP Lett, 1999, 6(6): 145 - 147.
- [6] S Amari. Information geometry of the EM and EM algorithms for neural networks [J]. Neural Networks, 1995, 8(9): 1379 - 1408.
- [7] Harry L Van Trees. Detection, Estimation, and Modulation Theory, Part [M]. New York: John Wiley & Sons, Inc, 1968.
- [8] L Xu, M Jordan. On convergence properties of the EM algorithm for gaussian mixtures [J]. Neural Computation, 1996, 8: 129 - 151.

## 作者简介:



王惠刚 男, 1974 年生于湖北孝感, 西北工业大学航海工程学院博士研究生, IEEE 学生会会员, 主要研究方向为统计信号处理、阵列信号处理及盲信号处理。