

混合高斯参数估计的动态簇算法

王平波, 蔡志明, 刘旺锁

(海军工程大学电子工程学院, 武汉 430033)

摘要: 混合高斯概率密度模型可以很好地拟合非高斯样本的概率密度。在各高斯分量概率密度互不重叠的条件下, 使用动态簇算法可以快速而精确地估计出混合高斯概率密度模型参数。这是一种基于最小均方差原则的递推算法, 在正向推导出各种可能的簇边界后, 再根据确定的最末边界值逆向推定各前导簇边界, 从而得到混合高斯概率密度模型参数估计值。描述模型及参数估计问题之后, 动态簇算法被推导出来。然后深入探讨了该算法的实质及适用条件。最后结合数值仿真实例, 分析了动态簇算法的估计性能。

关键词: 混合高斯; 累积方差; 动态簇算法

中图分类号: TJ630

文献标识码: A

文章编号: 1000-3630(2007)-04-0741-06

Dynamic cluster algorithm for Gaussian mixture parameter estimation

WANG Ping-bo, CAI Zhi-ming, LIU Wang-suo

(Electronic Engineering College, Navy Engineering University, Wuhan 430033, China)

Abstract: Probability density of non-Gaussian processes can be well fit by Gaussian mixture model whose parameters can be estimated through the dynamic clutter algorithm that is a recursion on the principle of minimum mean-square deviation. All possible boundaries of each clutter are forward-derived. Since the right boundary of final clutter is determinate, boundaries of previous clutters can be recurred backwards one by one. Thus the Gaussian mixture parameters are estimated. After descriptions of the model and the estimation problem, dynamic clutter algorithm for Gaussian mixture parameters is obtained. Its essential ideas and applicability are discussed in detail. A numerical example is presented to study the performance of estimation.

Key words: Gaussian mixture; cumulative variance; dynamic cluster algorithm

1 引 言

混合高斯模型(简记为 GM)因具有结构简明、参数精练、高效稳健等突出优点,被广泛于图像、医疗、语音、通信、雷达、声纳等信号处理领域对非高斯概率密度(PDF)进行拟合。这种模型的参数估计方法有很多种,比如最大似然估计(MLE)及其变种 EM 迭代算法。MLE 作为有效估计,能够取得很高

的估计性能,但一则其结构复杂难以实现,二则可能会因无法界定方差而得到无意义的奇异解^[1]。而 EM 算法则以求解完全数据的似然函数之期望的极值代替直接搜寻常规似然函数的极值,也可以取得较高的估计性能,但估计结果依赖于迭代初值的设置,一旦初值设定不合理,迭代就会收敛于局部的极值点,甚至不收敛,致使估计性能严重下降^[2]。文献[3]中提出了一种与上述各方法思想完全不同的 GM PDF 参数(简记为 GM 参数)估计全新方案—动态簇算法(简记为 DC)。这是一种非迭代、非极值化的另类求解方法。它基于最小累积方差准则把样本空间划分为几个互不重叠的高斯簇,然后基于此得到 GM 参数估计。但该文只是简要说明了 DC 的思想及实现

收稿日期: 2006-03-18; 修回日期: 2006-07-13

基金项目: 国家 973 基金项目(5132102ZZT32)

作者简介: 王平波(1976-),男,山东人,博士研究生,研究方向为水声信号处理。

通讯作者: 王平波, E-mail:blackberet@163.com

DC的几个重要公式,并未给出详细的DC算法流程,也未就此作专门的实验验证。

在计算机上实现了DC算法,并对其估计性能作了大量的仿真实验研究发现,DC并不适用于所有的GM情形,但只要PDF满足一定条件,它就能以更快的运算速度取得接近正确收敛的EM估计精度。综合考虑,它仍不啻于为一种优秀的GM参数估计算法。本文首先对GM模型及其参数估计问题进行了有针对性的描述,然后阐述了DC算法的思想和具体实现流程,明确了其适用前提,并分析了其局限性之根本所在。然后结合数值仿真实例,详细探讨了其估计性能。

2 GM模型及估计问题描述

一般地,GM模型的概率密度表述如式1所示,

$$f(u_n) = \sum_{m=1}^M \varepsilon_m \cdot f_m(u_n) \quad (1)$$

这里, u_n 为非高斯序列 $u=[u_1, u_2, \dots, u_N]^T$ 的第 n 个样本; M 为模型阶数; ε_m 为混合参数,满足式(2)所示的关系; f_m 为高斯分布 $N(u_m, \sigma_m^2)$ 的PDF,如式(3)所示。

$$\sum_{m=1}^M \varepsilon_m = 1 \quad (2)$$

$$f_m(u_n) = \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left[-\frac{(u_n - u_m)^2}{2\sigma_m^2}\right] \quad (3)$$

可以看到,拥有级数形式PDF的GM是一种纯数学模型,它的导出不基于特定的物理模型,它的使用不受具体物理意义的限制,这与仅适用于窄带PDF建模的Class A模型有着明显的区别。当然,它仍有自己明确的物理解释。事实上,式(1)所示的GM,恰能解释非高斯信号形成的物理机理,即非高斯信号是由多个高斯源信号依据一定的概率分布组成的。具体而言, M 代表了组成非高斯信号的高斯源(亦称高斯分量)数量; f_m 就是第 m 个高斯源的PDF; ε_m 为当前样本 u_n 来自第 m 个高斯源的概率。显然,通过简单地调整GM参数 $g=[\varepsilon_1, \dots, \varepsilon_M, \mu_1, \dots, \mu_M, \sigma_1^2, \dots, \sigma_M^2]^T$ 就可以拟合几乎任意的单钟型或多钟型PDF。

经过简单的数学推导,我们可以得到混合高斯序列 u 的一阶矩 μ (亦称为均值)和二阶中心矩 σ^2 (亦称为方差)分别如式(4)、(5)所示。

$$\mu = \sum_{m=1}^M \varepsilon_m \mu_m = 1 \quad (4)$$

$$\sigma^2 = \sum_{m=1}^M \varepsilon_m D_m \quad (5)$$

这里, D_m 称为第 m 个高斯源的类二阶中心矩(或类方差),如式(6)所示。显然,若均值 $\mu=0$,则 D_m 就变成了相应高斯源的真实方差,即 $D_m=\sigma_m^2$ 。

$$D_m = \int_{-\infty}^{\infty} (u_n - \mu)^2 f_m(u_n) du_n \quad (6)$$

图1是一段非高斯序列的PDF,粗实线为使用二阶混合高斯参数 $g=[\varepsilon_1, \varepsilon_2, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2]^T$ 描述的PDF曲线,而粗虚线则为使用高斯参数 $g=[\mu, \sigma^2]^T$ 描述的PDF曲线。显然,二者已经非常不一致了,这时如果无视序列的非高斯性,仍使用高斯PDF描述之,则定会造成极大误差。

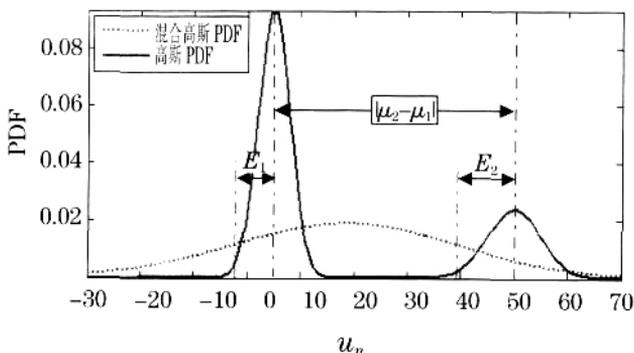


图1 混合高斯PDF示例

Fig.1 An example of GM PDF

图1中还标示出了三个量: $|\mu_2 - \mu_1|$ 、 E_1 、 E_2 。 $|\mu_2 - \mu_1|$ 就是两个高斯分量均值之距离;而 E_1 、 E_2 则称为各自的 α 样本量半宽,即 α 样本量宽(以均值为中心包含了 $\alpha \cdot 100\%$ 样本的区间宽度)的一半,这里 α 为取值在0~1之间的指定概率。一般地,高斯分布 $N(u_m, \sigma_m^2)$ 的 α 样本量半宽可按式(7)计算。

$$E_m = \mu_m - \Phi_m^{-1}[(1-\alpha)/2] \quad (7)$$

其中, $\Phi_m^{-1}(p)$ 为如式(8)所示的累积分布逆函数。

$$\Phi_m^{-1}(p) = \left\{ u_n \left| \frac{1}{\sqrt{2\pi\sigma_m^2}} \int_{-\infty}^{u_n} \exp\left[-\frac{(t-\mu_m)^2}{2\sigma_m^2}\right] dt = p \right. \right\} \quad (8)$$

可如式9所示使用这三个量来判断混合高斯PDF中两个相邻的高斯分量能否被以 $\alpha \cdot 100\%$ 的成功率清楚区分:

$$\begin{cases} \text{若 } |\mu_j - \mu_i| > (E_i + E_j), \text{ 则能够清楚区分} \\ \text{若 } |\mu_j - \mu_i| > (E_i + E_j), \text{ 则不能清楚区分} \\ \text{若 } |\mu_j - \mu_i| \approx (E_i + E_j), \text{ 则处于临界状态} \end{cases} \quad (9)$$

这里, i, j 是1~ M 间的数, $|i - j| = 1$ 。

PDF中的各高斯分量能否被清楚区分,这关系到DC算法能否取得好的估计性能,原因将在下文阐述,这里只给出能否清楚区分的判断方法。

这样,式(1)所示的GM模型参数估计问题可

描述为: 如何基于观测样本序列 $u=[u_1, u_2, \dots, u_N]^T$ 确定其混合高斯 PDF 参数 $g=[\varepsilon_1, \dots, \varepsilon_M, \mu_1, \dots, \mu_M, \sigma_1^2, \dots, \sigma_M^2]^T$ 。这里必须指出, GM 模型地定阶问题, 即如何确定 GM 模型阶数 M , 也是数据 GM 建模应当要解决的关键技术之一。限于篇幅, 本文对此不做研究, 只是简单地假定 M 为先验已知。

3 参数估计的 DC 算法

既然非高斯信号是由多个高斯源信号按一定概率分布组合而成的, 那么, 可以直观地想象, 若根据一定规则, 将非高斯序列 u 的样本空间划分为 M 个子集(簇), 每个子集都属于一个高斯源, 那么 GM 参数估计问题就会迎刃而解。关键是如何划分, 亦即如何确定划分的原则。一般而言, 划分出来的这 M 个高斯簇相互间会有重叠, 致使样本重复利用而估计不准确, 为使其尽量不相重叠, 要令如式 (10) 所示的这 M 个簇的累积方差达到最小。

$$J = \sum_{m=1}^M \varepsilon_m \sigma_m^2 \quad (10)$$

这就是所谓的最小均方差准则, 它是大多数应用的首选准则。下面我们就在这一准则下进行动态簇算法的推导。

将信号 u 的样本 u_1, u_2, \dots, u_N (简记为 u_i^N) 按照由小到大的顺序重新排列为 v_1, v_2, \dots, v_N (序列 v 样本则简记为 v_i^N)。而要将 u_i^N 划分为 M 个不相重叠的簇, 也就是在实整数轴上将 v_i^N 的下标 $1, 2, \dots, N$ 分割为 M 段, 分界点 n_1, n_2, \dots, n_{M-1} 。定义第 1 个簇的左边界为 $n_0=0$ 、第 M 簇的右边界为 $n_M=N$ 。同样简记边界下标序列为 n_m^M 。现在要做的就是, 确定(亦即估计)使得均方误差 J 为最小的边界下标的具体值 $\hat{n}_1, \hat{n}_2, \dots, \hat{n}_{M-1}$ 。

用区间 $(n_{m-1}, n_m]$ (不包括左边界点, 但包括右边界点) 表示第 m 簇, 簇内样本均值为 μ_m , 方差为 σ_m^2 , 混合参数为 ε_m , 定义如式 (11) 所示。

$$\begin{aligned} \mu_m|_{n_m+1}^{n_m} &= \frac{\sum_{n=n_{m-1}+1}^{n_m} v_n}{n_m - n_{m-1}} \\ \sigma_m^2|_{n_m+1}^{n_m} &= \frac{\sum_{n=n_{m-1}+1}^{n_m} (v_n - \mu_m|_{n_m+1}^{n_m})^2}{n_m - n_{m-1}} \\ \varepsilon_m|_{n_m+1}^{n_m} &= \frac{n_m - n_{m-1}}{N} \end{aligned} \quad (11)$$

对于给定最右边界 n_m 的前 m 簇样本, 其最小累积方差 $\omega_m(n_m)$ 可以通过一组边界点 n_i^M , 由第 1 簇到第 m 簇累加求得:

$$\omega_m(n_m) = \min_{n_1 < n_2 < \dots < n_{m-1}} \sum_{j=1}^m \varepsilon_j |_{n_j+1}^{n_j} \sigma_j^2 |_{n_j+1}^{n_j} \quad (12)$$

当 $m=M$ 时, 最小累积方差也就是最小均方差, 即 $\min J = \omega_M(n_M)$ 。

最小均方差可通过动态编程用下述递归法求出。设第 m 簇的最佳右边界点为 n_m , 则第 $m-1$ 簇的最佳右边界点为 \tilde{n}_{m-1} , 依此类推, 则动态编程算法如下:

第一步: 初始化 ($m=1$)

$$1 \quad n_1 \quad N - M + 1 \quad (13)$$

$$\tilde{n}_0 = n_0 = 0 \quad (14)$$

$$\omega_1(n_1) = \varepsilon_1 |_{n_1+1}^{n_1} \sigma_1^2 |_{n_1+1}^{n_1} \quad (15)$$

第二步: 递推迭代 ($m=m+1, 1 < m < M$)

$$m \quad n_m \quad N - M + m \quad (16)$$

$$\tilde{n}_{m-1} = \operatorname{argmin}_{n_{m-1}} \{ \omega_{m-1}(n_{m-1}) + \varepsilon_m |_{n_m+1}^{n_m} \sigma_m^2 |_{n_m+1}^{n_m} \} \quad (17)$$

$$\omega_m(n_m) = \omega_{m-1}(\tilde{n}_{m-1}) + \varepsilon_m |_{n_m+1}^{n_m} \sigma_m^2 |_{n_m+1}^{n_m} \quad (18)$$

第三步: 终止迭代 ($m=M$)

$$n_M = N \quad (19)$$

$$\tilde{n}_{M-1} = \operatorname{argmin}_{n_{M-1}} \{ \omega_{M-1}(n_{M-1}) + \varepsilon_M |_{n_M+1}^{n_M} \sigma_M^2 |_{n_M+1}^{n_M} \} \quad (20)$$

$$\omega_M(n_M) = \omega_{M-1}(\tilde{n}_{M-1}) + \varepsilon_M |_{n_M+1}^{n_M} \sigma_M^2 |_{n_M+1}^{n_M} \quad (21)$$

考察上述步骤, 当 $m=1$ 时, 式 (13) 表明 n_1 有 $N - M + 1$ 种可能值, 于是 $\omega_1(n_1)$ 也随之有 $N - M + 1$ 种可能值; 当 $m=2$ 时, 式 (16) 表明 n_2 也有 $N - M + 1$ 种可能值, 对应每一个 n_2 可能值, 根据式 (17) 所示的 argmin 运算可以求出 $N - M + 1$ 个 n_1 值中令花括弧内二阶累积均方差取得最小的 n_1 值 \tilde{n}_1 , 再根据式 (18) 求出对应的最小二阶累积均方差 $\omega_2(n_2)$, 而 n_2 具有 $N - M + 1$ 种可能值, 所以 $\tilde{n}_2, \omega_2(n_2)$ 也都相应具有 $N - M + 1$ 种可能。 $m=3, 4, \dots, M-1$ 时各变量值, 皆可以此类推获得。直到最后一步, 当 $m=M$ 时, 根据定义, n_M 只有一种可能值, 即 $\tilde{n}_M = n_M = N$, 于是根据式 (20) 即可确定出第 $M-1$ 个分界点 n_{M-1} 的值 $\tilde{n}_{M-1} = \tilde{n}_{M-1}$, 由此倒推回去, 可以逐次得到划分的最佳边界点估计为:

$$\begin{aligned} \hat{n}_{M-1} &= \tilde{n}_{M-1} \\ &\vdots \\ \hat{n}_{m-1} &= \tilde{n}_{m-1} \\ &\vdots \\ \hat{n}_1 &= \tilde{n}_1 \end{aligned} \quad (22)$$

根据上述迭代算法, 在最小均方差准则下得到序列 u 样本的最佳划分 $\{n_0, \hat{n}_1, \hat{n}_2, \dots, \hat{n}_{M-1}, n_M\}$ 后, 将之代入式 (11) 即可得到各高斯分量参数的估计值。

这就是序列 GM 参数估计的动态簇(DC)算法。

4 DC 算法的适用条件

DC 算法的巧妙之处在于它把排序后的样本划分成了几个互不重叠的子集,但这也正是它的缺陷之所在,决定了它只能适用于一定条件的 PDF 参数估计。比如说,两个均值很接近的高斯源信号 x 和 y , 按某一概率分布抽取样本,组合称为二阶 GM 序列 u , 则 u 的某个样本空间中势必既有来自于 x 的样本,又有来自于 y 的样本。如是,就无法把 u 的样本划分为两个互不重叠的区间,DC 算法得到的估计将很不准确。

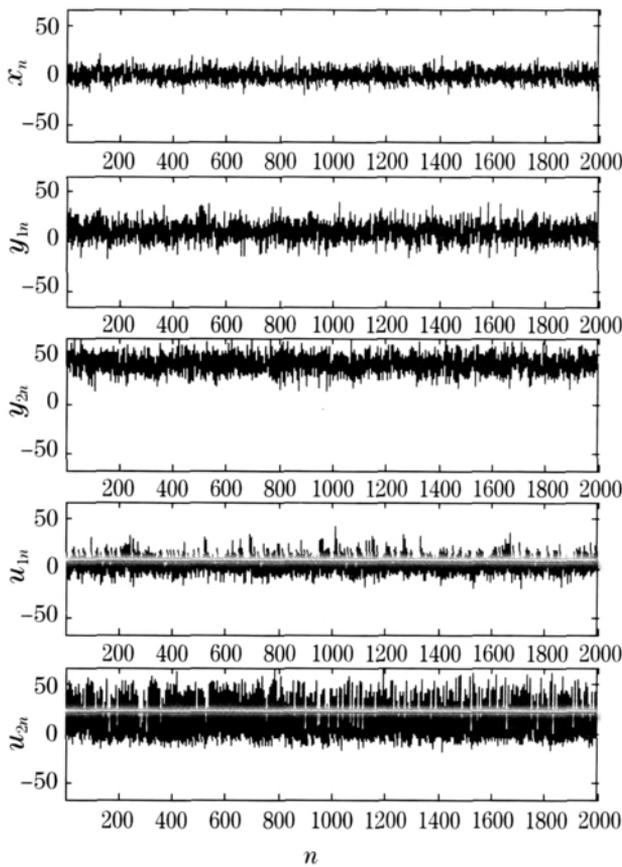


图 2 不同高斯源组成的 GM 序列波形

Fig.2 GM processes with different Gaussian sources

(u_{1n} comes from x and y_{1n} while u_{2n} comes from x and y_{2n})

以二阶 GM 模型为例,来说明 DC 算法的实质和应用限制。设 x, y_1, y_2 分别为服从分布 $N(0, 36), N(10, 81), N(40, 81)$ 的高斯序列,分别以 $[0.7, 0.3]$ 的概率从 x 和 y_1 中随机抽取样本组成了 GM 序列 u_1 ,以同样的概率分别从 x 和 y_2 中随机抽取样本组成了 GM 序列 u_2 。它们的波形及相应的 PDF 曲线分别如图 2 和 3 所示。为便于对比,波形和 PDF 都被分别归一化到同一坐标系中。

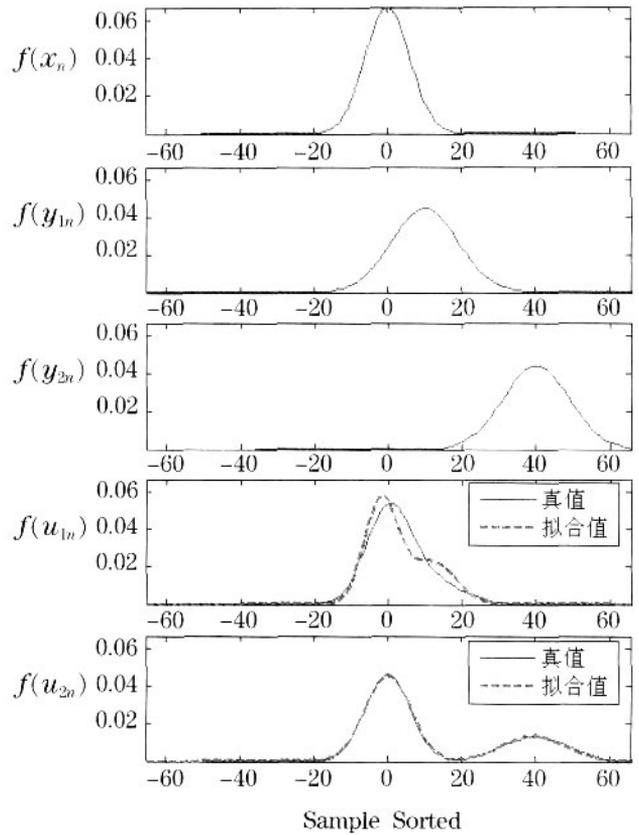


图 3 不同高斯源组成的 GM 序列 PDF

Fig.3 GM PDFs with different Gaussian sources

使用 DC 算法得到 u_1, u_2 的 GM 参数估计,并将据此绘出的 PDF 拟合曲线也相应绘制在图 3 中(真值用实线表示,拟合值用虚线表示)。

若设置两高斯源 PDF 清楚区分概率为 $\alpha=98\%$, 则根据式(7)计算可得 u_1 的两个高斯源的 α 样本量半宽分别为 $E_x=13.9581, E_{y_1}=20.9371$, 而其均值距离为 $|u_{y_1} - u_x|=10$, 显然属于式(9)中所示的不能清楚区分情形。更加直观地反映在图上,二阶 GM 序列 PDF 参数估计的 DC 算法,就相当于在其波形图上恰当位置划一条横线(如图 2 中 u_{1n} 波形上的立体线所示), 希望其值超过线的样本属于一个高斯源 y_1 , 而未超过线的属于另一个高斯源 x 。但遗憾的是,由于两个高斯源 x 和 y_1 的样本有大量重叠,这个“恰当位置”就难以确定。所以,可以想见,DC 算法得到的 u_1 的 GM 参数估计将不准确,图 3 中的 $f(u_{1n})$ 图印证了这一点: 图上使用参数估计值拟合的 PDF 曲线(虚线)与用真值描述的 PDF 曲线(实线)不一致,这表明参数估计不准确。高斯源 y_2 的 α 样本量半宽 E_{y_2} 与高斯源 y_1 的 α 样本量半宽 E_{y_1} 相等, 但 $|u_{y_2} - u_x|=40$, 已满足式(9)中所示的清楚区分条件。所以,应用 DC 算法将会得到 u_2 的高精度 GM 估计。反映在图 2 所示的波形上, y_2 和 x 的样本值本来就属于两个几乎毫不干涉的区间中, 于是, u_2 波形图上可以

找到一个恰当位置放置分隔横线(如图 2 中 u_{2n} 波形上的立体线所示)将样本值划分为两个不相重叠的簇, DC 算法将奏效。图 3 中的 $f(u_{1n})$ 图也印证了这一点: 图上使用参数估计值拟合的 PDF 曲线(虚线)与用真值描述的 PDF 曲线(实线)几乎完全重叠, 这表明参数估计很精确。

根据使用经验, 混合高斯 PDF 中任意相邻两高斯分量若能够以 $\alpha=98\%$ 的概率清楚区分, 则使用 DC 算法就能取得很高 GM 参数估计性能, 同时还能取得高于 EM 迭代三倍以上的运算速度。

那么, 在工程实用中应当如何判断一段数据能否使用 DC 算法进行 GM 建模呢? 也就是说, 对于一段数据, 应当怎样判断它的 PDF 是否由能够清楚分离的高斯分量所构成的呢? 要从理论上严格地推导出一个界定规则, 将是一件非常麻烦的事情。事实上, 对于某一领域要处理的专业数据, 我们完全可以根据此类数据样本分布的柱状统计图来先验地判定其是否可以使用 DC 算法进行 GM 建模。下节将结合实例来说明这种经验判定法。

5 实例

为便于考察估计效果, 使用参数可控的仿真数据。首先舍取抽样法仿真产生参数为 $g=[0.7, 0.3, 5, 50, 9, 25]^T$ 的 2 阶混合高斯白序列 u , 然后对之使用本文所述 DC 算法进行 GMAR 模型参数估计。每段观测样本长度为 2000 点, 共进行 500 次蒙特-卡罗试验。对于这样的参数设置, 由式(4)、(5)计算可得其真均值为 18.5, 真方差为 439.05。设置清楚区分概率 $\alpha=98\%$, 则由式(7)计算可得两个高斯源的 α 样本量半宽分别为 $E_1=6.9790$ 、 $E_2=11.6317$, 而 $|\mu_2 - \mu_1|=45$, 所以满足式(9)中所示的可清楚区分条件, 属于 DC 算法的适用范围。注意, 因为这是参数可控的仿真数据, 我们才可以如此根据参数判断出它是可以使用 DC 算法进行 GM 建模的。现在, 我们将说明如何使用柱状图统计法来判断 DC 算法是否可行。取其中某一段长度为 2000 点的样本序列, 其波形如图 4 所示。绘制出其样本分布的柱状统计图, 如图 5 所示。

这种柱状图的绘制方法比较简明: 首先把样本取值区间等分为 Q 个互不重叠的子集(图 5 中 $Q=40$), 然后统计每个子集中的样本个数作为该子集的响应。这实质上是一种对样本数据 PDF 的简单的非参数拟合。得到这种柱状图后, 究竟能否使用 DC

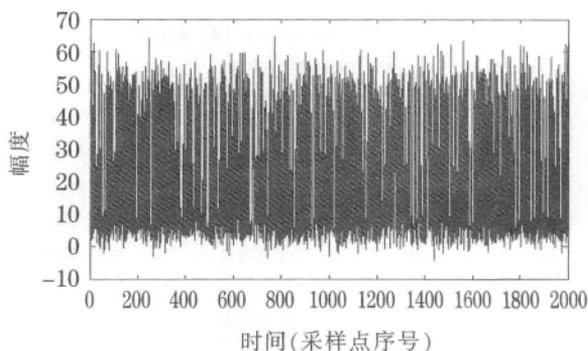


图 4 一段仿真数据波形

Fig.4 A segment of numerical data

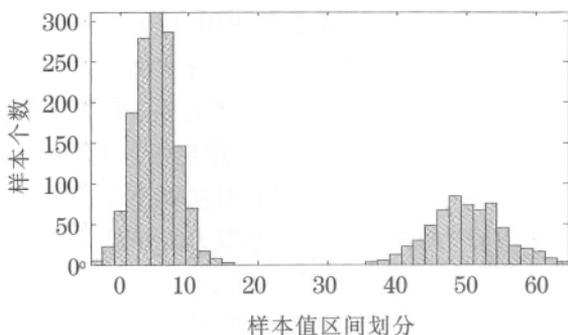


图 5 样本分布的柱状图

Fig.5 Histogram of samples' distribution

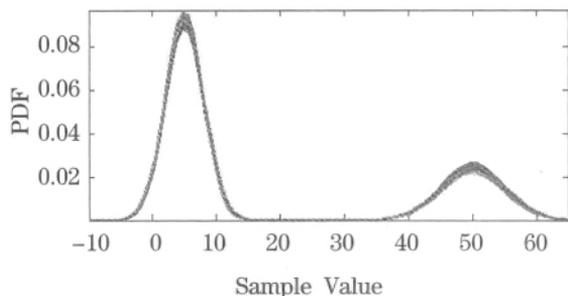


图 6 DC 算法得到的 PDF 估计(前 100 次试验叠合)

Fig.6 Fitting of PDF with DC algorithm
(100 estimations overlaid)

算法对数据 PDF 进行 GM 建模, 就一目了然了。显然, 对于具有图 5 所示柱状图的数据, 应该可以使用 DC 算法进行 GM 建模。

根据真值绘出的 PDF 如图 1 所示, 而根据前 100 次试验估计值绘出的 100 条 PDF 曲线叠合如图 6 所示。可以看到, 图 6 中的各条曲线高度重叠汇聚称为一条与图 1 真值极其一致的粗线, 这表明了参数估计的无偏性和有效性都很好。而表 1 为基于 500 次试验 GM 参数估计结果得到的统计数字。可以看到, 各 GM 参数估计均值非常接近其真值, 而估计方差也很低, 这再次印证了估计的无偏性和有效性之高。

表 1 DC 算法 GM 参数估计性能(500 次试验之统计)
Table 1 Performance of DC algorithm
(500 experiments)

项目	真值	均值	方差
ε_1	0.7	0.699 3	1.0308e- 4
ε_2	0.3	0.300 7	1.0308e- 4
μ_1	5	5.002 7	7.1519e- 3
μ_1	50	50.004 3	4.1046e- 2
σ_1^2	9	8.975 5	1.1575e- 1
σ_2^2	25	24.949 8	2.0452e+0

6 结束语

GM是对非高斯数据进行 PDF 建模的最有效模型之一,而当 GM PDF 中各高斯分量能够以 98%左右的概率清楚区分时,本文所述的 DC 算法将能取得很好的 GM 参数估计性能和理想的运算速度,这是依赖于初值设置的 EM 迭代算法所不能比拟的。

使用 DC 算法估计 GM 参数,在非高斯信号处理领域有着重要的实践意义。比如,在雷达、声纳等信号检测领域,可以使用估计得到的 GM 参数构建

高斯化滤波器以高斯化检测背景,从而提高后续匹配滤波(或相关检测)检测性能。再比如,非高斯有色序列预白滤波器的构建需要功率谱(PSD)参数估计,而它的精确获得必须与 PDF 参数估计相耦合,DC 算法就可以被结合进来。

使用柱状图法来判断可否使用 DC 算法对数据进行 GM 建模,还存在一定的局限性,比如:若柱状图中的某一独立“钟形包”本身即非高斯的,那么这种判断显然就是不准确的了。因此,这个问题还有待于以后进一步研究。

参 考 文 献

- [1] Redner R A, Walker H F. Mixture densities, maximum likelihood, and the EM algorithm[J]. SIAM Review, 1984, 26: 195-209.
- [2] Aaron A D Souza. Using EM to estimate a probability density with a mixture of Gaussians[DB/OL]. <http://cite-seer.ist.psu.edu>, 2000-1.
- [3] ZHAO Yunxin, ZHUANG Xinhua, TING Shenjin. Gaussian mixture density modeling of non-Gaussian source for autoregressive process[J]. IEEE Transactions on Signal Processing, 1995, 43(4): 894-903.