

基于自适应 Metropolis 算法的 α 稳定分布参数估计

郝燕玲, 单志明, 沈 锋

(哈尔滨工程大学自动化学院, 黑龙江 哈尔滨 150001)

摘要: 基于马尔可夫链蒙特卡罗(Markov chain Monte Carlo, MCMC)方法的 α 稳定分布参数估计具有良好的性能, 但不合适的提议函数常导致算法不收敛或混合性能不好。针对提议函数难以选择的问题, 提出了一种基于自适应 Metropolis 算法的非对称 α 稳定分布参数估计新方法。该方法利用 Markov 链的历史信息自动调整提议函数的协方差矩阵, 使其不断地逼近目标分布, 从而获得更好的估计结果。理论分析和仿真结果表明, 此方法不仅能准确地估计出 α 稳定分布的 4 个参数, 而且具有良好的鲁棒性和灵活性。

关键词: 信号处理; α 稳定分布; 马尔可夫链蒙特卡罗; Metropolis-Hastings 算法; 参数估计

中图分类号: TN 911.7

文献标志码: A

DOI: 10.3969/j.issn.1001-506X.2012.02.04

Parameter estimation of α -stable distributions based on adaptive Metropolis algorithm

HAO Yan-ling, SHAN Zhi-ming, SHEN Feng

(College of Automation, Harbin Engineering University, Harbin 150001, China)

Abstract: Markov chain Monte Carlo (MCMC) methods for the parameter estimation of α -stable distributions have good performance, but an improper choice of proposal distributions can often lead to unexpected results. Aiming at the difficulties to choose an effective proposal distribution, a novel method based on adaptive Metropolis (AM) algorithm is proposed for non-symmetric α -stable distributions. The method uses the full history (cumulated so far) of the chain to tune the covariance of the proposal distribution suitably. This adaptation strategy can approach an approximation of the target distribution, which increases the efficiency of the simulation. Theoretic analysis and simulation results show that this method can not only estimate the four parameters of α -stable distributions, but also perform very accurately and robustly.

Keywords: signal processing; α -stable distribution; Markov chain Monte Carlo (MCMC); Metropolis-Hastings (M-H) algorithm; parameter estimation

0 引言

高斯模型在传统的信号处理中占据主导地位, 原因有二: 其一, 理论上中心极限定理的支持; 其二, 这种假设给算法的研究带来极大方便, 使得算法存在简单的表达式, 便于进行理论分析^[1]。然而, 在实际应用中所遇到的许多信号或噪声往往具有显著的尖峰脉冲特性, 例如水声信号、低频大气噪声、生物医学信号和金融数据等^[2-4]。这类信号的统计特性显著偏离高斯分布, 特别是其概率密度函数具有比高斯分布更厚的拖尾, 此时由于高斯模型不能与实际情况相匹配, 从而导致信号处理器性能显著退化。

在描述非高斯信号方面, α 稳定分布作为目前唯一的

一类满足广义中心极限定理的分布得到了广泛的研究与应用^[5-9]。 α 稳定分布又称为非高斯稳定分布、重尾分布, 它是一种更加广义化的高斯分布, 能够保持自然噪声过程的产生机制和传播条件, 非常好地与实际数据相吻合^[1]。因此, 通常用 α 稳定分布模型来描述这类具有显著尖峰脉冲波形和较厚概率密度函数拖尾的随机信号。由于 α 稳定分布的特性完全由其特征函数的 4 个参数来确定, 因此如何由稳定分布的一个样本实现来估计这 4 个参数成为实际应用中首先要解决的问题。

目前, 常用的 α 稳定分布的参数估计方法主要有最大似然法^[10-11]、样本分位数法^[12-13]、样本特征函数法^[14-15]、渐进极值法^[16]、负阶矩法和对数法^[17]以及基于贝叶斯的参数

收稿日期: 2011-03-01; 修回日期: 2011-07-29。

基金项目: 国家自然科学基金(61001154, 61102107); 中国博士后科学基金(20100480979)资助课题

作者简介: 郝燕玲(1944-), 女, 教授, 主要研究方向为惯性导航与组合导航系统、电子导航图系统和信息融合滤波技术、统计信号处理。

E-mail: haoyanling@hrbeu.edu.cn

估计方法^[2,18]等。最大似然方法属于复杂的非线性优化问题,并且没有初始值选择原则和收敛性分析可供利用;样本分位数法受到查找表的限制,仅适用于特征指数大于 0.6 的情况;渐进极值法不能正确地估计较大的 α 值;特征函数法、负阶矩法和对数法都有良好的性能,且计算量小容易实现,但仅适用于对称稳定分布(symmetric α -stable, $S_{\alpha}S$)的情况,因此无法估计对称参数 β 。然而在实际应用中,由于观测值往往反映出分布的某种偏斜特性($\beta \neq 0$),当 α 值较小时,样本均值与真实位置参数也相去甚远,这在一定程度上都影响了分布模型的准确性。而贝叶斯方法不仅实现了 α 稳定分布参数的同时估计,而且具有很好的性能,但是提议函数及其方差的选择会影响算法性能。为此,本文提出一种基于自适应 Metropolis(adaptive Metropolis, AM)算法的非对称 α 稳定分布参数估计方法,该方法利用马尔可夫链的历史信息自动优化提议函数的协方差矩阵,马尔可夫链蒙特卡罗(Markov chain Monte Carlo, MCMC)模拟方法不再依赖于参数的初始建议分布,使算法在保持良好遍历性的同时提高了抽样效率。文中首先介绍了 α 稳定分布的基本理论,然后建立稳定分布的贝叶斯推理模型,在分析标准 Metropolis-Hastings (M-H)算法的基础上,提出了提议函数的自适应调整策略,最后通过仿真验证了该算法的正确性和高效性。

1 α 稳定分布

α 稳定分布是广义的高斯分布,它比高斯分布具有更广泛的适用性,甚至可以描述许多不满足中心极限定理的数据。但与其他统计模型不同,除少数情况外, α 稳定分布没有统一封闭的概率密度函数,通常用其特征函数和 4 个参数 $\{\alpha, \beta, \gamma, \mu\}$ 描述。

$$\varphi(t) = \begin{cases} \exp \{j\mu t - |\gamma t| [1 - j\beta \text{sign}(t) \tan(\frac{\pi\alpha}{2})]\}, & \alpha \neq 1 \\ \exp \{j\mu t - |\gamma t| [1 + j\beta \text{sign}(t) \frac{2}{\pi} \ln |t|]\}, & \alpha = 1 \end{cases} \quad (1)$$

式中, $\text{sign}(\cdot)$ 是符号函数; $\alpha \in (0, 2]$ 为特征指数, α 值越小,分布的拖尾也就越厚,因此脉冲特性越显著,这个特点使得 α 稳定分布可以很好地描述某些脉冲性比较显著的信号,用 α 稳定分布作为建模工具是非常灵活的,其原因在于 $0 < \alpha \leq 2$ 可以用于控制分布拖尾的厚度; $\beta \in [-1, +1]$ 为偏斜参数,表示分布的对称特性,当 $\beta = 0$ 时,称为对称 α 稳定分布(记为 $S_{\alpha}S$); $\gamma > 0$ 为分散系数,表示 α 稳定分布的分散程度,类似于高斯分布的方差; μ 为位置参数,对于 $S_{\alpha}S, \mu \in (-\infty, +\infty)$ 表示分布的均值($1 < \alpha \leq 2$)或中值($0 < \alpha \leq 1$),当 $\gamma = 1, \mu = 0$ 时, α 稳定分布称为标准 α 稳定分布。 α 稳定分布存在 3 种特殊情况:高斯分布($\alpha = 2, \beta = 0$)、柯西分布($\alpha = 1, \beta = 0$)和 Levy 分布($\alpha = 1/2, \beta = \pm 1$)。此处定义 $0 < \alpha < 2$ 为分数低阶 α 稳定分布,以区别于 $\alpha = 2$ 的高斯分布,若随机变量 X 服从稳定分布,则记为: $\mathbf{X} \sim S_{\alpha}(\gamma, \beta, \mu)$ 。

α 稳定分布的概率密度函数是其特征函数的傅里叶逆变换,因此可以通过计算式(2)积分得到:

$$f_{\alpha, \beta}(x | \gamma, \mu) = F^{-1}(\varphi(t)) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-jtx) \varphi(t) dt \quad (2)$$

除上述 3 种特殊情况外,式(2)不存在最终的解析表达式,这严重阻碍了 α 稳定分布的研究与应用。为此, Nolan 提出一种计算稳定密度的数值方法,该方法基于一种特殊的积分表达式推导出稳定密度的积分公式,其中被积函数是连续、有界、非震荡函数,并且积分区间有界^[19]。鉴于此,本文将 Nolan 的积分公式与自适应积分方法相结合,可以方便地计算 α 稳定分布概率密度函数。

2 α 稳定分布的贝叶斯推理模型

贝叶斯推理本质上是通过观测数据把参数的先验概率分布转化为参数的后验概率分布。它的一般模式:先验信息 + 样本信息 \rightarrow 后验信息。利用贝叶斯条件概率计算后验概率分布的计算式如下:

$$p(\mathbf{A} | \mathbf{B}) = \frac{p(\mathbf{B} | \mathbf{A}) p(\mathbf{A})}{p(\mathbf{B})} \quad (3)$$

式中, $p(\mathbf{A})$ 是未知量的先验密度; $p(\mathbf{B} | \mathbf{A})$ 是参数 \mathbf{A} 的似然函数; $p(\mathbf{A} | \mathbf{B})$ 是已知 \mathbf{B} 时 \mathbf{A} 的后验密度; $p(\mathbf{B}) = \int p(\mathbf{B} | \mathbf{A}) \cdot p(\mathbf{A}) d\mathbf{A}$ 是归一化常数。因此,贝叶斯思想也可表示为

$$p(\mathbf{A} | \mathbf{B}) \propto p(\mathbf{B} | \mathbf{A}) p(\mathbf{A}) \quad (4)$$

由此可见,贝叶斯定理提供了一种形式优美的估计未知量的方法论。推理过程既利用了似然函数,又利用了参数的先验信息,因此贝叶斯估计量具有更小的方差,能得到更精确的预测结果。

利用贝叶斯推理进行 α 稳定分布参数估计,假定未知量为 $\mathbf{A} = \{\alpha, \beta, \gamma, \mu\}$, 已知观测矢量为 $\mathbf{B} = \mathbf{X}$, 将 \mathbf{A}, \mathbf{B} 代入方程(4),如果未知量 \mathbf{A} 中 4 个参数相互独立,可推得 α 稳定分布的贝叶斯层次模型:

$$p(\alpha, \beta, \gamma, \mu | \mathbf{X}) \propto p(\mathbf{X} | \alpha, \beta, \gamma, \mu) p(\alpha, \beta, \gamma, \mu) = p(\mathbf{X} | \alpha, \beta, \gamma, \mu) p(\alpha | a) p(\beta | b) p(\gamma | g, h) p(\mu | \xi, \kappa^{-1}) \quad (5)$$

式中, a, b, g, h, ξ, κ 是未知参数 \mathbf{A} 先验分布的超参数。这种先验分布是贝叶斯推理的基础和出发点,它反映了实验前对总体分布的认识,在获得了样本信息后,这种认识的改变就反映在后验分布中。在本文中,假设特征指数 α 和偏斜参数 β 的先验分别是各定义域内的均匀分布,分散系数 γ 的先验是以 (g, h) 为超参数的逆伽马分布,位置参数 μ 的先验是以 (ξ, κ^{-1}) 为超参的正态分布,即满足

$$\begin{cases} p(\alpha | a) = \frac{1}{a} = \frac{1}{2}, & 0 < \alpha \leq 2 \\ p(\beta | b) = \frac{1}{b} = \frac{1}{2}, & -1 \leq \beta \leq 1 \end{cases} \quad (6)$$

$$\begin{cases} p(\gamma) \sim IG^{-1}(\gamma | g, h) \\ p(\mu) \sim N(\mu | \xi, \kappa^{-1}) \end{cases} \quad (7)$$

3 AM 算法

3.1 Metropolis-Hastings (M-H) 算法

尽管贝叶斯推理模式简单,但是一般难以获得具体的后验分布密度函数。主要原因是贝叶斯推理存在多维积分问题,且通常没有解析解。如:未知量 \mathbf{A} 的函数 $f(\mathbf{A})$ 的后验期望为

$$E[f(\mathbf{A} | \mathbf{B})] = \frac{\int f(\mathbf{A}) p(\mathbf{B} | \mathbf{A}) p(\mathbf{A}) d\mathbf{A}}{\int p(\mathbf{B} | \mathbf{A}) p(\mathbf{A}) d\mathbf{A}} \quad (8)$$

MCMC 方法又被称为动态 Monte Carlo 方法^[20],以动态构造 Markov 链为基础,通过遍历性约束来实现模拟目标分布的一类随机模拟方法,成功解决了贝叶斯分析中异常复杂的数值计算问题。目前,在贝叶斯分析中应用最广泛的 MCMC 方法主要有两种: M-H 算法和 Gibbs 抽样方法,本文主要讨论比 Gibbs 更具一般意义的 M-H 算法。利用 M-H 算法进行 α 稳定分布参数估计的实现步骤如下:

步骤 1 对于特征指数 α ,假设 Markov 链的当前状态为 $(t, \alpha^{(t)})$,从提议函数 $q(\cdot)$ 抽取候选点 α^{new} ,即 $\alpha^{\text{new}} \sim q(\alpha^{\text{new}} | \alpha^{(t)})$;

步骤 2 根据 M-H 算法计算接受概率 A_α :

$$A_\alpha = \min \left\{ 1, \frac{\prod_i p(x_i | \alpha^{\text{new}}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)})}{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)})} \times \frac{p(\alpha^{\text{new}}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)}) q(\alpha^{(t)} | \alpha^{\text{new}})}{p(\alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)}) q(\alpha^{\text{new}} | \alpha^{(t)})} \right\} \quad (9)$$

假定未知参数 $\{\alpha, \beta, \gamma, \mu\}$ 的先验分布相互独立,提议函数 $q(\cdot)$ 为对称分布且 $p(\alpha)$ 为均匀先验,那么接受概率式(9)便可简化为

$$A_\alpha = \frac{\prod_i p(x_i | \alpha^{\text{new}}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)})}{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)})} \quad (10)$$

对偏斜参数 β 、分散系数 γ 和位置参数 μ 使用与特征指数 α 相同的策略,令 $\beta^{\text{new}} \sim q(\beta^{\text{new}} | \beta^{(t)})$, $\gamma^{\text{new}} \sim q(\gamma^{\text{new}} | \gamma^{(t)})$, $\mu^{\text{new}} \sim q(\mu^{\text{new}} | \mu^{(t)})$,计算可得接受概率 A_β, A_γ, A_μ 分别为

$$A_\beta = \frac{\prod_i p(x_i | \alpha^{(t)}, \beta^{\text{new}}, \gamma^{(t)}, \mu^{(t)})}{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)})} \quad (11)$$

$$A_\gamma = \frac{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{\text{new}}, \mu^{(t)}) IG(\gamma^{\text{new}} | g, h)}{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)}) IG(\gamma^{(t)} | g, h)} \quad (12)$$

$$A_\mu = \frac{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{\text{new}}) N(\mu^{\text{new}} | \xi, \kappa^{-1})}{\prod_i p(x_i | \alpha^{(t)}, \beta^{(t)}, \gamma^{(t)}, \mu^{(t)}) N(\mu^{(t)} | \xi, \kappa^{-1})} \quad (13)$$

步骤 3 生成随机数 $u \sim U(0, 1)$,如果 $u \leq A$ 则接受候选点,Markov 链的状态更新,否则拒绝候选点,Markov 链仍保持当前状态不变。

重复步骤 1~步骤 3 至产生需要的样本数量为止。

3.2 AM 算法

M-H 算法的不足在于:必须预先根据经验确定参数的抽样区间和提议函数,调试过程繁琐且难以保证参数的各态历经性和算法的快速收敛。为此,利用自适应策略获取提议函数。AM 算法是一种全局自适应调整策略,它基于经典随机游走 Metropolis 算法,其基本思想是首先构建高斯提议函数,然后利用 Markov 链样本校正高斯分布的协方差矩阵^[21]。为了保持良好的遍历性,AM 算法利用 Markov 链的全部历史信息连续地调整提议函数,这种调整会改变提议函数的延展度和空间方向,使其不断地逼近目标分布函数,从而有效地提高了抽样效率。

假定初始非自适应段的长度为 N_0 ,当抽样次数 $i \leq N_0$ 时,协方差取固定值 \mathbf{C}_0 ;之后,协方差 \mathbf{C}_i 根据式(14)进行自适应更新,即

$$\mathbf{C}_i = \begin{cases} \mathbf{C}_0, & i \leq N_0 \\ s_d \text{Cov}(x_0, \dots, x_{i-1}) + s_d \epsilon \mathbf{I}_d, & i > N_0 \end{cases} \quad (14)$$

式中, s_d 是一个仅依赖于状态空间维数 d 的比例因子,可以确保接受概率在合适的范围内; $\epsilon > 0$ 是小常数,以确保协方差阵的非奇异性; \mathbf{I}_d 是 d 维单位矩阵;经验协方差矩阵根据式(15)计算,即

$$\text{Cov}(x_0, \dots, x_k) = \frac{1}{k} \left(\sum_{i=0}^k \mathbf{x}_i \mathbf{x}_i^T - (k+1) \bar{\mathbf{x}}_k \bar{\mathbf{x}}_k^T \right) \quad (15)$$

式中, $\bar{\mathbf{x}}_k = \frac{1}{k+1} \sum_{i=0}^k \mathbf{x}_i$, $\mathbf{x}_i \in \mathbf{R}^d$ 。将式(15)带入式(14),得到 $i > N_0$ 时协方差 \mathbf{C}_i 的递归计算式为

$$\mathbf{C}_{i+1} = \frac{i-1}{i} \mathbf{C}_i + \frac{s_d}{i} (i \bar{\mathbf{x}}_{i-1} \bar{\mathbf{x}}_{i-1}^T - (i+1) \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^T + \mathbf{x}_i \mathbf{x}_i^T + \epsilon \mathbf{I}_d) \quad (16)$$

由此可知,AM 算法的采样机制与所有的历史样本信息有关,可以保证算法的收敛性和遍历性^[21]。令 $\boldsymbol{\theta} = \{\alpha, \beta, \gamma, \mu\}$,则 AM 算法的采样过程描述为:① 初始化;② 根据式(14)计算协方差 \mathbf{C}_i ;③ 产生候选点 $\boldsymbol{\theta}^{\text{new}} \sim N(\boldsymbol{\theta}, \mathbf{C}_i)$;④ 根据式(10)~式(13)计算接受概率 A ;⑤ 生成随机数 $u \sim U(0, 1)$,若 $u \leq A$ 则接受候选点,令 $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}^{\text{new}}$;否则拒绝,令 $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}$ 。重复过程②~⑤至产生需要的样本数量为止。

4 仿真结果及分析

仿真实验 1 设样本服从分布 $\mathbf{X} \sim S_{1.5}(0.3, 0.5, 3, 25)$,取以 Markov 链当前时刻状态为均值,标准差 $\boldsymbol{\sigma}_0 = \{\sigma_\alpha, \sigma_\beta, \sigma_\gamma, \sigma_\mu\}$ 的高斯型提议函数,样本个数 $N = 1\ 000$,迭代次数为 10 000 次, burn-in 时间 1 000,未知量先验分布的超参数为 $g = 1, h = 1, \kappa = 1, \xi = \text{mean}(y)$,优化的提议函数标准差 $\boldsymbol{\sigma}_0 = \{0.12, 0.12, 0.1, 0.12\}$,自适应调整间隔为 100,比例因子 $\sqrt{s_d} = 1.2$,分别用 M-H 算法和 AM 算法对 α 稳定分布参数进行估计。

图 1 给出了 α 稳定分布参数的估计结果。从图中可以看出,在给定的提议函数及其初始方差下,M-H 算法(图 1(a))和 AM 算法(图 1(b))都能同时准确地估计出稳定分布的 4 个参数 $\{\alpha, \beta, \gamma, \mu\}$,并且具有良好的混合性能,主要表现为链能够快速远离初值并在支撑域附近剧烈地摆动。图 2 为提议函数标准差的自适应变化情况。从图中可以看出,由于采用了自适应策略,AM 算法提议函数的标准差不再是固定值,而是不断变化的。在初始阶段较大的方

差有利于加快状态空间搜索,后期较小的方差又提高了参数估计精度。图 3 给出了估计结果的自相关系数。从图 3 可以看出,两种算法的自相关系数图均为很陡的下降曲线,说明样本之间具有良好的独立性,后面的样本能够较快地遗忘先前的样本。由此可知,尽管给定的初始方差是合理的,能够保证 M-H 算法的可靠收敛,但它并非是最优的,而 AM 算法可以通过自适应获得更好的提议函数方差,确保良好的混合和收敛效果,从而极大地提高了算法效率。

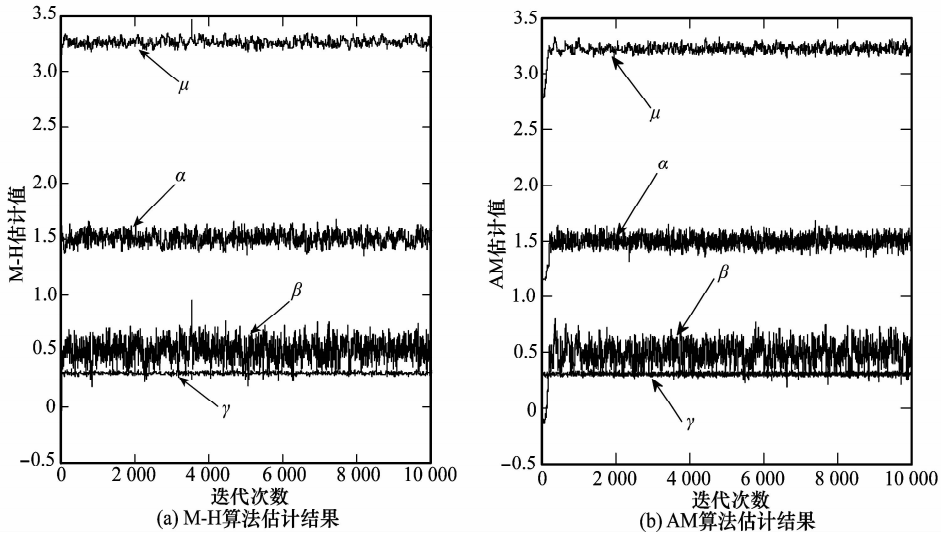


图 1 实验 1 参数估计结果

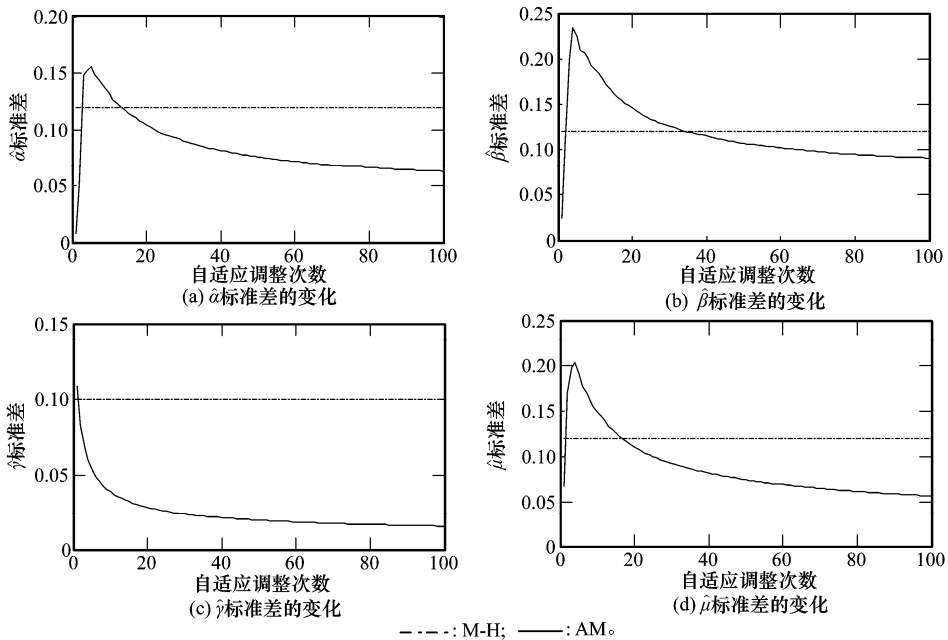


图 2 实验 1 提议函数标准差的变化

表 1 给出了仿真实验 1 的统计结果。统计结果亦表明在给定的提议函数及其方差下 M-H 算法和 AM 算法在估计精度方面性能相当,但是 AM 算法具有更高的接受概率和更小的自相关系数。由此可知,AM 算法通过自适应获

得的提议函数更加合理,高接受概率意味着达到相同的效果所需要的仿真时间较短,同时较小的自相关系数表明利用 AM 算法获得的 Markov 链包含更多的独立样本,算法效率也更高。

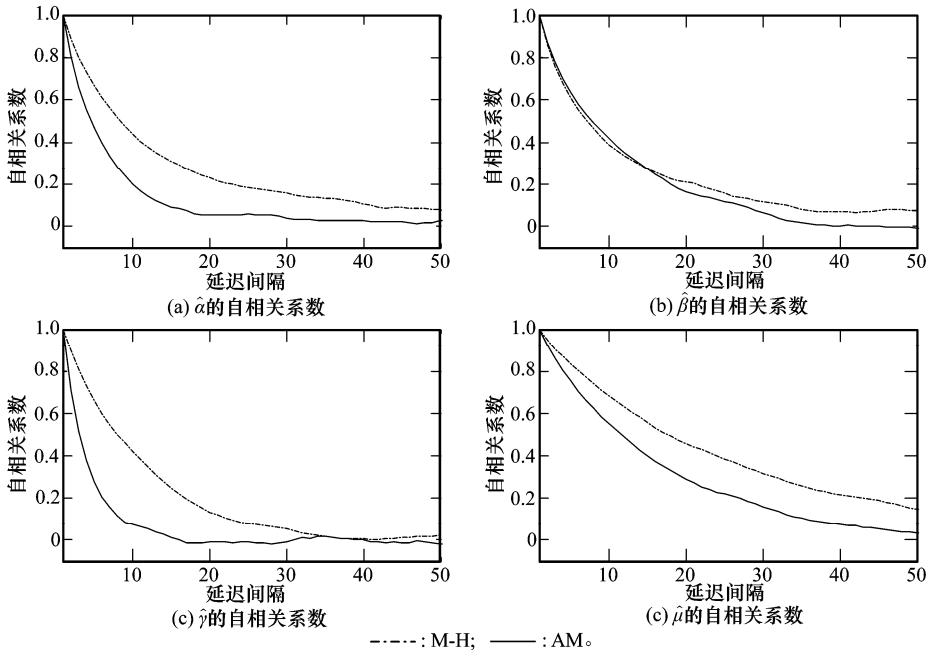


图 3 实验 1 Markov 链自相关系数

表 1 仿真 1 参数估计结果统计: $X \sim S_{1.5}(0.3, 0.5, 3, 25)$

仿真实验 1 参数	真实值 θ_T	初始值 θ_0	算法 M-H/AM	估计值 $\hat{\theta}$	估值标准差 σ_{θ}	接受概率 A	积分自相关时间
α	1.5	1.1	M-H	1.505 0	0.047 6	27.67	26.67
			AM	1.500 0	0.044 9	43.66	12.87
β	0.5	0.0	M-H	0.502 8	0.085 6	45.70	23.35
			AM	0.498 0	0.085 7	46.73	19.81
γ	0.3	1.0	M-H	0.303 2	0.009 7	10.87	19.47
			AM	0.308 7	0.010 1	43.18	6.07
μ	3.25	2.5	M-H	3.259 0	0.028 7	17.39	43.20
			AM	3.254 3	0.027 5	27.59	29.37

仿真实验 2 为避免实验的复杂性,这里将参数 α 和 γ 的标准差分别增大为原来的 5 倍,参数 β 和 μ 的标准差分别减小为原来的 1/20 倍,其他设置和参数均保持不变,分别应用 M-H 算法和 AM 算法对 α 稳定分布参数进行估计,以比较这两种算法的适应性。

图 4 给出了 α 稳定分布参数的估计结果。从图中可以看出,由于提议函数方差的改变导致 M-H 算法性能显著退化,参数 α, β 和 γ 的收敛时间明显增加,混合性能也变差, β 参数的估计值变得不稳定,参数 μ 的接受概率也显著降低。而 AM 算法的估计性能几乎不受影响,此时 Markov 链的收敛速度、混合性能和参数估计精度等方面都明显优于 M-H 算法。图 5 给出了估计结果的自相关系数。从图 5 可以看出,无论是增大方差还是减小方差, M-H 算法的自相关值都很大,说明有效样本数量明显减少,因此导致算法效率低下,而 AM 算法的自相关系数图仍是很陡的下降曲线,

从而保持了样本之间良好的独立性。图 6 为 AM 算法提议函数标准差的自适应变化情况。由于采用了自适应策略, AM 算法能够根据历史信息自动调整提议函数标准差使其更接近目标分布,因而 AM 算法的收敛速度和估计精度都要优于 M-H 算法。另外,这种自适应策略也使得初始方差的选取更加灵活,从而使繁琐复杂的调试和优化过程大为简化。

表 2 给出了仿真实验 2 的统计结果。结果表明 M-H 算法对提议函数方差的选取很敏感,若方差选择不当会导致算法性能显著退化,甚至得到错误的估计结果。而 AM 算法具有很强的韧性,在任意初始条件下都能够通过自适应调整获得合适的提议函数标准差,实现 α 稳定分布参数的准确估计,而且保持恰当的接受概率和良好的自相关特性。最后需要指出的是,该文方法适用于具有任意偏斜特性的 α 稳定分布参数估计,在对称条件下估计精度更高。

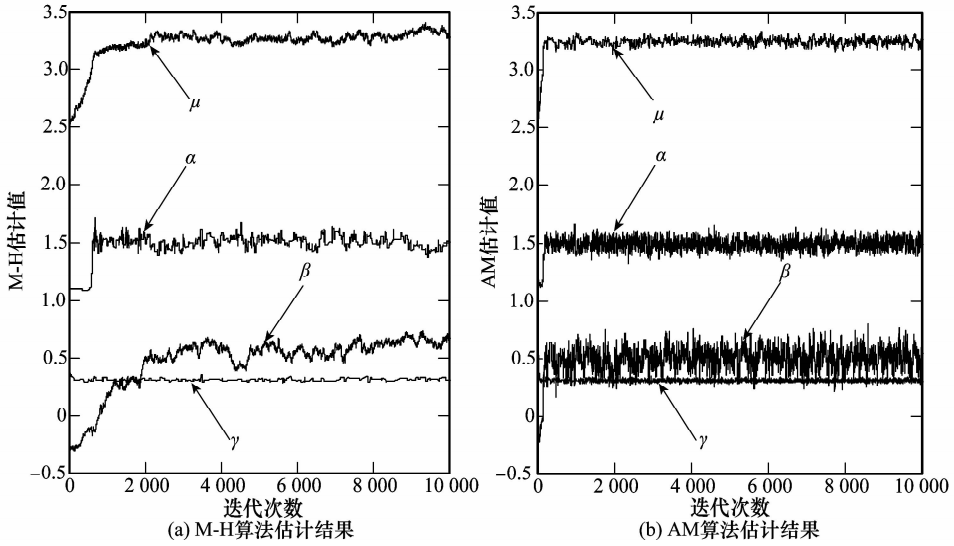


图 4 实验 2 参数估计结果

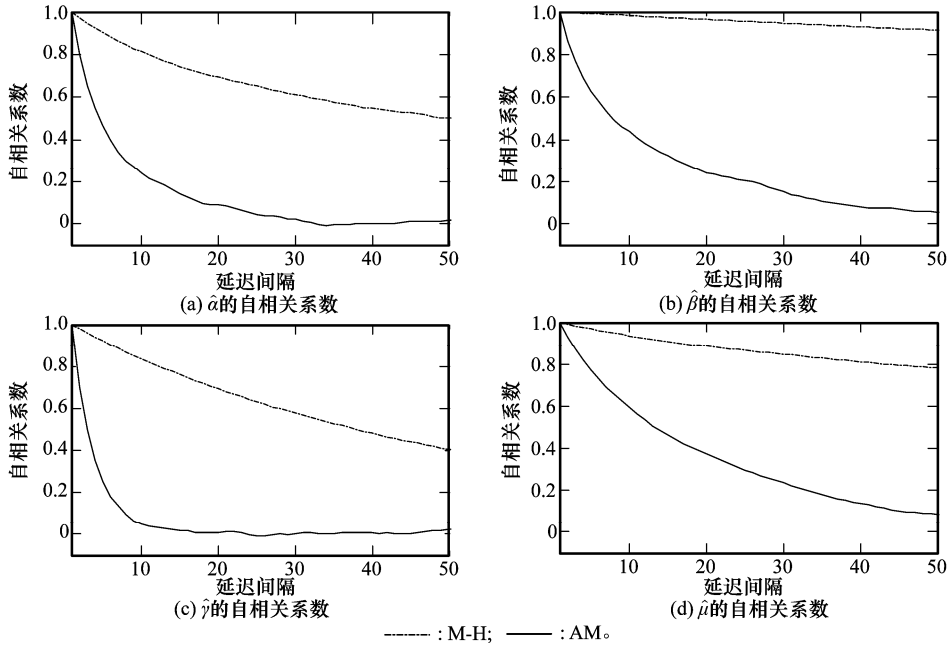


图 5 实验 2 Markov 链自相关系数

表 2 仿真实验 2 参数估计结果统计: $X \sim S_{1.5}(0.3, 0.5, 3, 25)$

仿真实验 2 参数	真实值 θ_T	初始值 θ_0	算法 M-H/AM	估计值 $\hat{\theta}$	估值标准差 $\sigma_{\hat{\theta}}$	接受概率 A	积分自相关时间
α	1.5	1.1	M-H	1.487 0	0.103 0	5.56	67.36
			AM	1.502 2	0.045 1	44.97	13.06
β	0.5	0.0	M-H	0.489 7	0.226 7	94.62	96.36
			AM	0.503 2	0.085 5	40.65	25.75
γ	0.3	1.0	M-H	0.309 9	0.014 9	2.30	64.96
			AM	0.308 1	0.009 9	39.40	7.43
μ	3.25	2.5	M-H	3.240 0	0.135 4	86.66	87.83
			AM	3.250 8	0.027 0	20.51	35.21

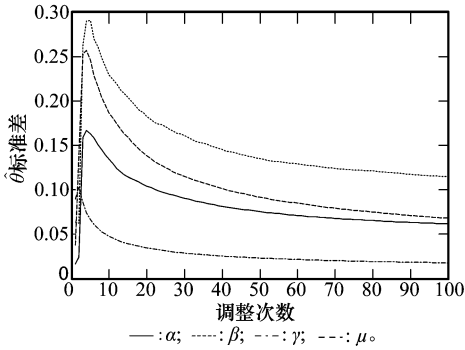


图 6 实验 2 AM 算法标准差的变化

5 结论

α 稳定分布是一种非常有效的统计信号处理工具,可以很好地描述具有显著尖峰脉冲波形和重尾的非高斯现象。要利用 α 稳定分布建模首先要解决参数估计问题,因此本文提出了基于 AM 算法的非对称 α 稳定分布参数估计新方法。仿真表明,AM 算法克服了 M-H 算法在实际应用中提议函数及其方差难以选择的困难,具有比 M-H 算法更好的韧性和灵活性,在任意初始条件下都能利用 Markov 链的历史信息。通过自适应调整获得合适的提议函数,实现所有稳定分布参数的准确估计,而且具有合适的接受概率和良好的自相关性,在效率、收敛速度和估计精度等方面都优于 M-H 算法。

参考文献:

[1] Shao M, Nikias C L. Signal processing with fractional lower order moments: stable processes and their applications[J]. *Proceedings of the IEEE*, 1993, 81(7): 986 - 1010.

[2] Casarin R. Bayesian inference for mixtures of stable distributions[D]. Pairs: University Pairs IX-Dauphine, 2004.

[3] Roenko A A, Lukin V V, Djurovic I. Two approaches to adaptation of sample myriad to characteristics of SaS distribution data[J]. *Signal Processing*, 2010, 90(7): 2113 - 2123.

[4] Salas-Gonzalez D, Kuruoglu E E, Ruiz D P. Modelling with mixture of symmetric distributions using Gibbs sampling[J]. *Signal Processing*, 2010, 90(3): 774 - 783.

[5] Arikan O, Belge M, Cetin A E, et al. Adaptive filtering approaches for non-Gaussian stable process[C]// *Proc. of the International Conference on Acoustics, Speech, and Signal Processing*, 1995: 1400 - 1403.

[6] Tsihrintzis G A, Nikias C L. Evaluation of fractional lower-order statistical-based detection algorithm on real sea-clutter data[J]. *IEE Proceedings on Radar, Sonar & Navigation*, 1997, 144(1): 29 - 38.

[7] Ma X, Nikias C L. Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics[J]. *IEEE Trans. on Signal Processing*, 1996, 44(11): 2669 - 2687.

[8] 朱晓波, 王首勇, 李旭涛, 等. 非高斯杂波中的 MIMO 雷达信号分离[J]. *系统工程与电子技术*, 2010, 32(6): 1210 - 1214. (Zhu X B, Wang S Y, Li X T, et al. MIMO radar signal separation in non-Gaussian clutter[J]. *Systems Engineering and Electronics*, 2010, 32(6): 1210 - 1214.)

[9] 唐洪, 邱天爽. Alpha 稳定分布噪声环境下广义恒模算法收敛性能的研究[J]. *电子学报*, 2009, 37(1): 118 - 121. (Tang H, Qiu T S. Convergence properties of the GCMA in alpha-stable noise environment[J]. *Acta Electronica Sinica*, 2009, 37(1): 118 - 121.)

[10] DuMouchel W H. Stable distribution in statistical inference[D]. New Haven: Yale University, 1971.

[11] Brorsen B W, Yang S R. Maximum likelihood estimates of symmetric stable distribution parameters[J]. *Communications in Statistics-Simulation and Computation*, 1990, 19(4): 1459 - 1464.

[12] Fama E F, Roll E R. Parameter estimates for symmetric stable distributions[J]. *Journal of the American Statistical Association*, 1971, 66(334): 331 - 338.

[13] McCulloch J H. Simple consistent estimators of stable distribution parameters[J]. *Communications in Statistics-Simulation and Computation*, 1986, 15(4): 1109 - 1136.

[14] Koutrouvelis I A. Regression-type estimation of the parameters of stable laws[J]. *Journal of the American Statistical Association*, 1980, 75(372): 918 - 928.

[15] Koutrouvelis I A. An iterative procedure for the estimation of the parameters of stable laws[J]. *Communications in Statistics-Simulation and Computation*, 1989, 10(1): 85 - 93.

[16] Tsihrintzis G A, Nikias C L. Fast estimation of the parameters of alpha-stable impulsive interference[J]. *IEEE Trans. on Signal Processing*, 1996, 44(6): 1492 - 1503.

[17] Ma X, Nikias C L. Parameter estimation and blind channel identification in impulsive signal environments[J]. *IEEE Trans. on Signal Processing*, 1995, 43(12): 2884 - 2897.

[18] Lombardi M J. Bayesian inference for α -stable distributions: a random walk MCMC approach[J]. *Computational Statistics & Data Analysis*, 2007, 51(5): 2688 - 2700.

[19] Salas-Gonzalez D, Kuruoglu E E, Ruiz D P. Finite mixture of α -stable distributions[J]. *Digital Signal Processing*, 2009, 19(2): 250 - 264.

[20] Laine M. Adaptive MCMC methods with applications in environmental and geophysical models[D]. Finland: Finnish Meteorological Institute, 2008.

[21] Haario H, Saksman E, Tamminen J. An adaptive Metropolis algorithm[J]. *Bernoulli*, 2001, 7(2): 223 - 242.