Dipartimento di Statistica
"Giuseppe Parenti"

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - www.ds.unifi.it

WORKING PAPER 2004/11

Bayesian inference for alpha-stable distributions: A random walk MCMC approach

Marco J. Lombardi

Università degli Studi di Firenze

Statistics

Bayesian inference for α-stable distributions: A random walk MCMC approach

Marco J. Lombardi*

Abstract

The α-stable family of distributions constitutes a generalization of the Gaussian distribution, allowing for asymmetry and thicker tails. Its practical usefulness is coupled with a marked theoretical appeal, given that it stems from a generalized version of the central limit theorem in which the assumption of the finiteness of the variance is replaced by a less restrictive assumption concerning a somehow regular behavior of the tails. The absence of the density function in a closed form and the associated estimation difficulties have however hindered its diffusion among practitioners.

In this paper I introduce a novel approach for Bayesian inference in the setting of α-stable distributions that resorts to a FFT of the characteristic function in order to approximate the likelihood function; the posterior distributions of the parameters are then produced via a random walk MCMC method. Contrary to the other MCMC schemes proposed in the literature, the proposed approach does not require auxiliary variables, and so it is less computationally expensive, especially when large sample sizes are involved. A simulation exercise highlights the empirical properties of the sampler; an application on audio noise data demonstrates how this estimation scheme performs in practical applications.

Keywords: α-Stable distributions, Infinite variance, MCMC

1 Introduction

1.1 α-Stable Distributions

The α-stable family of distributions stems from a more general version of the central limit theorem which replaces the assumption of the finiteness of the variance with a much less restrictive one concerning the regular behavior of the tails (Gnedenko & Kolmogorov 1954); the Gaussian distribution is thus a particular case of α-stable distribution. This family of distributions has a very interesting pattern of shapes, allowing for asymmetry and thick tails, that makes them suitable for the modelling of several phenomena, ranging from the

*I would like to thank Steve Brooks, Fabio Corradi and Federico M. Stefanini for their useful comments and especially my co-supervisor Fabrizia Mealli for her insightful suggestions and discussions. A preliminary version of this paper was presented at the SIS 2004 scientific meeting in Bari.

engineering (noise of degraded audio sources) to the financial (asset returns); moreover, it is closed under linear combinations.

The family is identified by means of the characteristic function

$$\phi_1(t) = \begin{cases} \exp\left\{\mathrm{i}\delta_1 t - \gamma^\alpha |t|^\alpha \left[1 - \mathrm{i}\beta \mathrm{sgn}(t) \tan\frac{\pi\alpha}{2}\right]\right\} & \text{if } \alpha \neq 1 \\ \exp\left\{\mathrm{i}\delta_1 t - \gamma |t| \left[1 + \mathrm{i}\beta \frac{2}{\pi} \mathrm{sgn}(t) \ln|t|\right]\right\} & \text{if } \alpha = 1 \end{cases} \tag{1}$$

which depends on four parameters: $\alpha \in (0, 2]$, measuring the tail thickness (thicker tails for smaller values of the parameter), $\beta \in [-1, 1]$ determining the degree and sign of asymmetry, $\gamma > 0$ (scale) and $\delta_1 \in \mathbb{R}$ (location).

While the characteristic function (1) has a quite manageable expression and can straightforwardly produce several interesting analytic results (Zolotarev 1986), it unfortunately has a major drawback for what concerns estimation and inferential purposes: it is not continuous with respect to the parameters, having a pole at $\alpha = 1$. The shorthand notation we will employ for the distribution is $\mathcal{S}_1(\alpha, \beta, \gamma, \delta_1)$.

An alternative way to write the characteristic function that overcomes this problem, due to (Zolotarev 1986), is the following:

$$\phi_0(t) = \begin{cases} \exp\left\{\mathrm{i}\delta_0 t - \gamma^\alpha |t|^\alpha \left[1 + \mathrm{i}\beta \tan\frac{\pi\alpha}{2} \mathrm{sgn}(t) \left(|\gamma t|^{1-\alpha} - 1\right)\right]\right\} & \text{if } \alpha \neq 1 \\ \exp\left\{\mathrm{i}\delta_0 t - \gamma |t| \left[1 + \mathrm{i}\beta \frac{2}{\pi} \mathrm{sgn}(t) \ln(\gamma |t|)\right]\right\} & \text{if } \alpha = 1 \end{cases} \tag{2}$$

In this case, the distribution will be denoted as $\mathcal{S}_0(\alpha, \beta, \gamma, \delta_0)$. The formulation of the characteristic function is, in this case, quite more cumbersome, and the analytic properties have a less intuitive meaning. Anyway, this formulation is much more useful for what concerns statistical purposes. The only parameter that needs to be "translated" according to the following relationship is δ:

$$\delta_0 = \begin{cases} \delta_1 + \beta\gamma \tan\frac{\pi\alpha}{2} & \text{if } \alpha \neq 1 \\ \delta_1 + \beta \frac{2}{\pi} \gamma \ln\gamma & \text{if } \alpha = 1 \end{cases} \tag{3}$$

On the basis of the above equations, a $\mathcal{S}_1(\alpha, \beta, 1, 0)$ distribution corresponds to a $\mathcal{S}_0(\alpha, \beta, 1, -\beta\gamma \tan\frac{\pi\alpha}{2})$, provided that $\alpha \neq 1$.

Another parameterization which is sometimes used is the following (Zolotarev 1986):

$$\phi_2(t) = \begin{cases} \exp\left\{\mathrm{i}\delta_1 t - \gamma_2^\alpha |t|^\alpha \exp\left[-\mathrm{i}\frac{\pi\beta_2}{2} \mathrm{sgn}(t) \min(\alpha, 2-\alpha)\right]\right\} & \text{if } \alpha \neq 1 \\ \exp\left\{\mathrm{i}\delta_1 t - \gamma_2 |t| \left[1 + \mathrm{i}\beta_2 \frac{2}{\pi} \mathrm{sgn}(t) \ln(\gamma_2 |t|)\right]\right\} & \text{if } \alpha = 1 \end{cases} \tag{4}$$

Also in this case, however, the density is not continuous with respect to α and presents a pole at $\alpha = 1$. Another unpleasant feature of this way of writing the characteristic function is that the meaning of the asymmetry parameter β changes according to the value of α: when $\alpha \in (0, 1)$ a negative β indicates negative skewness, whereas for $\alpha \in (1, 2)$ it produces positive skewness. For what concerns the "translation" of this parameterization into the others, we have, for $\alpha \neq 1$:

$$\begin{aligned} \beta &= \cot\frac{\pi\alpha}{2} \tan\left(\frac{\pi\beta_2}{2} \min(\alpha, 2-\alpha)\right) \qquad (5) \\ \gamma &= \gamma_2 \left[\cos\left(\frac{\pi\beta_2}{2} \min(\alpha, 2-\alpha)\right)\right]^{1/\alpha}, \end{aligned}$$

while δ and α remain unchanged.

1.2 Estimation issues

Unfortunately, (1) cannot be inverted to yield a closed-form density function except for a very few cases: $\alpha = 2$, corresponding to the normal distribution, $\alpha = 1$ and $\beta = 0$, yielding the Cauchy distribution, and $\alpha = \frac{1}{2}, \beta = \pm 1$ for the Lévy distribution.

This difficulty, coupled with the fact that moments of order greater than α do not exist whenever $\alpha \neq 2$, has made impossible the use of standard estimation methods such as maximum likelihood and the method of moments. Researchers have thus proposed alternative estimation schemes, mainly based on quantiles (McCulloch 1986), the performance of which is judged unsatisfactory in a number of respects. With the availability of powerful computing machines, it has become possible to exploit computationally-intensive estimation methods for the estimation of α-stable distributions parameters, such as maximum likelihood based on the FFT of the characteristic function (Mittnik, Rachev, Doganoglu & Chenyao 1999). Those methods, however, present some inconvenience: their accuracy is quite poor for small values of α because of the spikedness of the density function; furthermore, when the parameters are near their boundary, the distributions of the estimators become degenerate making frequentist inferential procedures unreliable.

The Bayesian approach has suffered from the same difficulties as the frequentist, as the absence of a closed-form density prevented from evaluating the likelihood function and thus constructing posterior inferential schemes. Also in this case, however, the availability of fast computing machines has made possible the use of MCMC methods. In particular, Buckle (1995) has shown that, conditionally on an auxiliary variable, it is possible to express the density function in closed form. With this result, he proposes a Gibbs sampling scheme for the α-stable distribution parameters.

This approach, however, presents some difficulties. It is unfortunately not straightforward to produce random numbers from this auxiliary variable and one must resort to rejection sampling. Since we need a random sample from the auxiliary variable of the same size as the observation vector for each iteration of the chain, it follows that this approach can be particularly slow, especially when large sample sizes are involved. Furthermore, since the parameterization involved is (4) and has poles at $\alpha = 1$ and $\beta = 0$, one needs to constrain a priori α and β on a portion of their support.

2 A random walk Metropolis sampler

I shall here introduce a novel approach for the construction of the posterior density of the (possibly) asymmetric α-stable law parameters that avoids the use of the auxiliary vector.

To put it on more formal grounds, we are looking for a computable expression for the likelihood function in order to be able to produce samples from the posterior distribution of the parameters according to Bayes' theorem, namely

$$p(\alpha, \beta, \gamma, \delta | \mathbf{x}) \propto p(\mathbf{x} | \alpha, \beta, \gamma, \delta) p(\alpha, \beta, \gamma, \delta).$$

As we have previously claimed, an approximate version of the likelihood function was used by Mittnik et al. (1999) to perform maximum likelihood optimization.

This approximation employs a inverse-FFT of the characteristic function that yields, for a given lattice of points, the exact values of the density at each abscissa. The densities of the observations in-between each abscissa are then obtained by linear interpolation. The fact that the points at which the inverse-FFT is computed need to be equally spaced can be a major shortcoming, since in order to cover observations in the extreme tails we might be forced to "waste" lots of computational resources. This drawback can be overcome by restricting the inverse FFT onto an interval which is likely to cover most of the observations and resorting to the series expansion of Bergstrøm (1952) for the observations outside that interval:

$$\begin{aligned} f_2(x;\alpha,\beta) &= \frac{1}{\pi}\sum_{k=1}^{+\infty}\frac{\Gamma(k\alpha+1)}{\Gamma(k+1)}\sin\left[\frac{k\pi\alpha}{2}K(\alpha,\beta)\right](-1)^{k-1}x^{-k\alpha-1} \\ f_2(x;\alpha,\beta) &= \frac{1}{\pi}\sum_{k=1}^{+\infty}\frac{\Gamma(k/\alpha+1)}{\Gamma(k+1)}\sin\left[\frac{k\pi\alpha}{2\alpha}K(\alpha,\beta)\right](-x)^{k-1}, \end{aligned} \quad (6)$$

with

$$K(\alpha,\beta) = \alpha + \beta\min(\alpha, 2-\alpha).$$

This piecewise approximation of the likelihood function can be combined with the prior in order to obtain the posterior distribution of the parameters. At this point, a simple random walk MCMC scheme can be employed to produce simulated samples from the posterior density.

The main advantage of this approach is its computational quickness, given that the FFT needs to be performed only once per iteration of the chain. Furthermore, although it is an approximation, the precision can be arbitrarily increased by simply reducing the spacing between each abscissa of the FFT. A similar scheme was proposed by Tsionas (1999) for a Metropolis subchain for the parameter α in the setting of a Gibbs sampler for symmetric α-stable regression models.

One of the fundamental issues in the implementation of a successful random-walk Metropolis approach is the appropriate choice of the variance-covariance matrix of the proposal distribution. Since in modelling heavy-tailed phenomena one usually deals with large sample sizes, it is known (DuMouchel 1973) that, in absence of an extremely strong prior and for values of α and β situated away from their bounds, the posterior distribution of the parameters should approximately behave as a multivariate Gaussian with variance-covariance matrix equal to the inverse of the information matrix. I have investigated the use, for the update step of the random-walk, of various Gaussian distributions with no correlation among the components and various choices of the individual component variances, but this eventually led to poor mixing of the chain. My proposal is thus to to first run a coarse maximum likelihood estimation, in order to get insights on both the correlation structure and the starting values of the chain, and use the inverse of the information matrix evaluated at the maximum likelihood estimate as covariance matrix of the Metropolis chain.

2.1 Simulation experiments

I will here report some results on simulated datasets about the empirical features of the proposed MCMC scheme. All results are based on C source code compiled

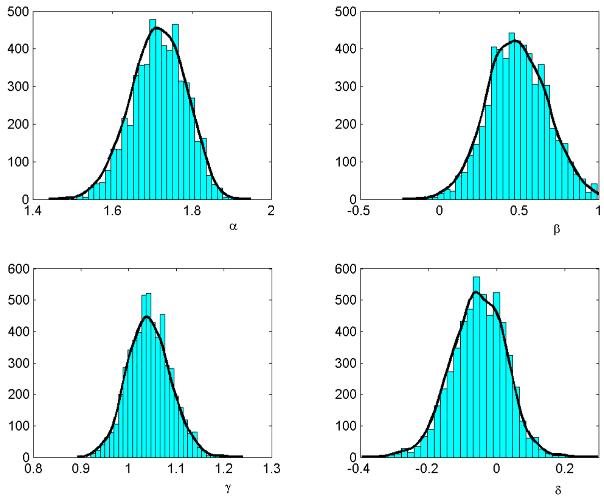

Figure 1: Histograms and kernel smoothed densities of the marginal posterior distributions, data set 1.

with Microsoft Visual Studio and run on a Intel® Pentium IV® processor at 2.66GHz with 512Mb RAM.

The first experiment aims at assessing the general properties of the approach. I have generated three synthetical random samples of size $N = 500$ from an α-stable distribution with parameters $\alpha = 1.7$, $\beta = 0.6$, $\gamma = 1$ and $\delta = 0$. A coarse (tolerance level 0.0001) maximum likelihood estimation with starting values close to the actual ones was run and the chain was then started from the estimated values of the parameters and run for 5500 iterations, with a burn-in of 500. The prior chosen were mild: uniform on the whole support for both α and β, inverse gamma with parameters $a = 2$ and $b = 3$ for γ and a $\mathcal{N}(0, 5)$ for δ. The resulting histograms and kernel smoothed posterior densities are displayed in figure 1; in figure 2, I report the behavior of the ergodic means for three different simulated data sets generated with the same parameter choice.

A visual examination of the ergodic mean behavior seems to indicate that the chain actually converged. To test this conjecture, I sampled 100 values from, respectively, the first and the second half of the realization of each chain (after discarding the burn-in period) and performed a Kolmogorov-Smirnov test for equal distribution of the two samples. Results, reported in table 1, clearly indicate that the chain has apparently converged: most of the p-values refer to values of the test statistic well outside any reasonable rejection region and only one, referring to δ, seems pathologic.

By plotting the autocorrelations of the chains (cf. figure 3), one actually notes that they all converge quickly to zero besides the one for which the Kolmogorov-Smirnov test rejected the null hypothesis. So one can conclude that the proposed algorithm possesses, in general, good mixing properties; pathologic behavior of the chain for certain datasets is however possible.

One of the most serious problems with MCMC algorithms is the "you've only seen where you've been" paradigm, that is the fact that the chain seems to have

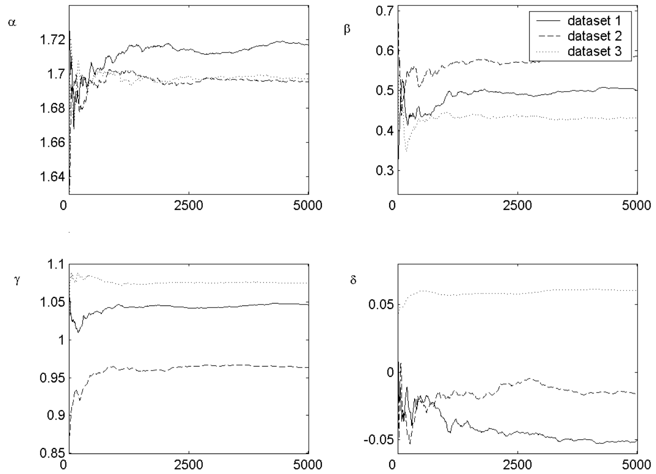

Figure 2: Evolution of the ergodic mean for three different simulated data sets.

Table 1: P-values of the Kolmogorov-Smirnov test for equal distribution of two samples of size 100 for each of the three data sets.

	α	β	γ	δ
Data set 1	0.2106	0.9062	0.6994	0.5806
Data set 2	0.2810	0.3667	0.8127	0.0783
Data set 3	0.3667	0.0541	0.1111	0.0000

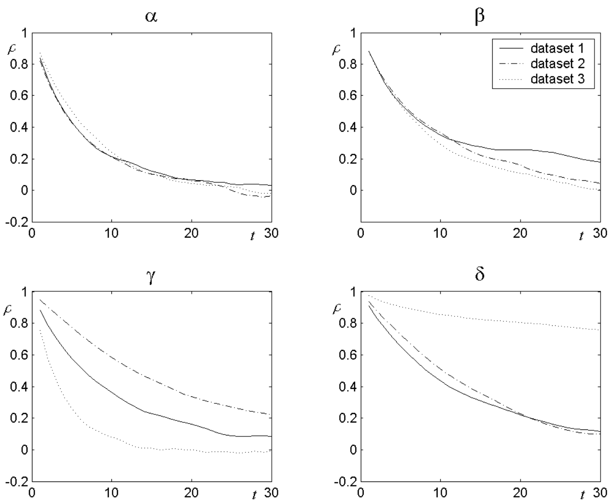

Figure 3: Autocorrelation structure of the chains for three different simulated data sets.

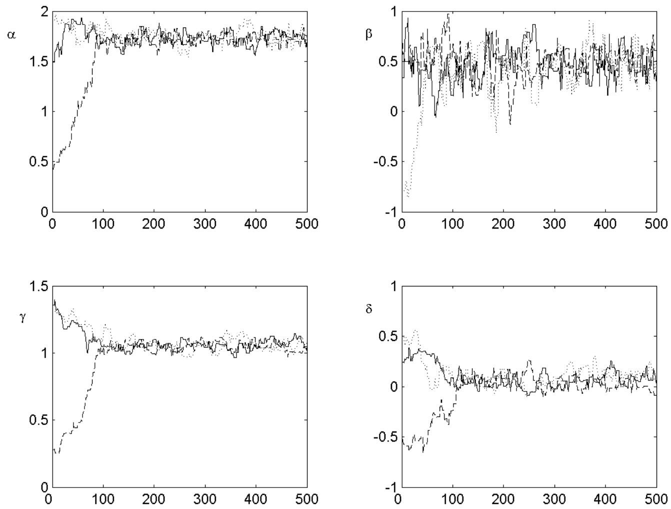

Figure 4: Behavior of three different chains for the same data set with different starting values.

converged but has failed to explore the whole sample space. Instead of a single long chain, several parallel chains running from widely dispersed starting points may overcome this problem. This approach is illustrated in figure 4, where I plot the first 500 runs of three different chains, for the same data set, starting from different and well spread points. One can observe that the chains converge quickly in the region where the posterior distribution has highest probability density.

I have claimed that the main advantage of the proposed approach is its computational quickness with respect to that of Buckle (1995). In the following example, I will compare the two methods. Since the random walk Metropolis and the Gibbs sampler are designed for two different parameterizations, respectively 0 and 2, the comparison was undertaken by converting the values produced by the Gibbs sampler into parameterization 0. A sample of size 500 was generated from a $\mathcal{S}_0(1.4, 0.3, 1, 0)$ distribution and two chains, based, respectively, on the Gibbs sampler of Buckle (1995) and on the proposed Metropolis random walk, were run for 1100 iterations, with burn-in of 100. The behavior of the ergodic means (figure 5) indicates that the chains behave in a similar way, although we have to note that the Gibbs-based chains seem to take much more time to convergence. From the point of view of the speed, the Gibbs takes as twice as the time as the proposed random walk: running the Gibbs sampler for 1100 iterations required 78.53 seconds against 34.03 seconds required by the proposed random walk Metropolis approach.

At this point, however, one might wonder why use a simulation-based approach instead of a less computationally intensive maximum-likelihood: after all, we know that the target posterior distribution is approximately multivariate normal, provided that we use a sufficiently large sample size and mild priors. The advantage of using a MCMC approach becomes however clear as we move

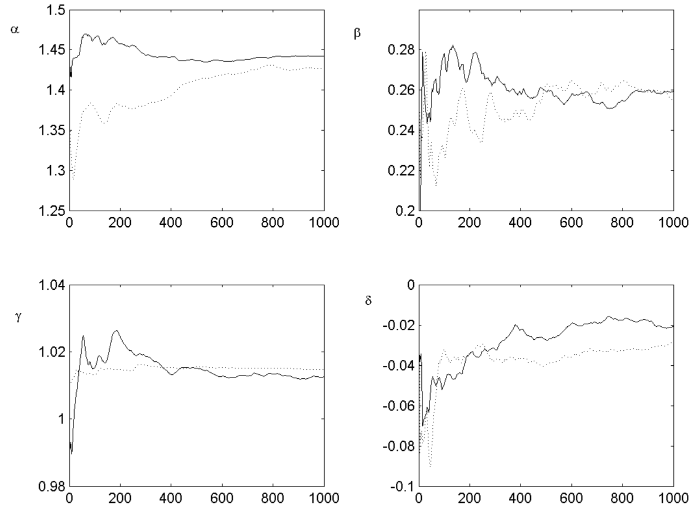

Figure 5: Behavior of the ergodic means for the random walk Metropolis (solid line) and the Gibbs sampler (dashed line)

to consider values of α and β close to their bounds. This is of great interest for practical applications, since in most cases the empirical distributions are heavy tailed but not too far from the Gaussian. In this case, as first pointed out by DuMouchel (1973), the distribution of the parameter estimators becomes degenerate and Gaussian-based confidence intervals are not reliable. The advantage of having a sample representative of the complete posterior distribution becomes then apparent. The issue is illustrated in the following example, in which two samples from, respectively, a $\mathcal{S}_0(1.95, 0.2, 1, 0)$ and a $\mathcal{S}_0(1.2, 0.95, 1, 0)$ were estimated with the above methodology. Priors for α and β were set to be uniforms on their support. The Gaussian posterior distribution and the empirically generated one are displayed in figure 6, from which we can observe that in both cases the normal approximation substantially underestimates the probability density on the left tail.

I have up to now employed mild priors. An unpleasant effect of the use of strongly informative priors is that, as the prior gets stronger, the posteriors moves away from normality: this causes the approximate covariance matrix for the Metropolis proposal, which is obviously based on a non-informative prior, to get more and more inaccurate, eventually leading to an inefficient behavior of the chain, i.e. a very high (or very low) rejection ratio. The situation can be however improved upon by employing the inverse Hessian matrix of the posterior density evaluated at its maximum instead of the information matrix of the maximum likelihood. Even if this approach has no sound theoretical justification – the posterior distribution is not Gaussian and thus it is not identified by its covariance matrix – it seems to improve the performance of the algorithm, as the following example shows. A random sample of 50 units was generated from a $\mathcal{S}_0(1.6, 0.2, 1, 0)$. The priors for α and β were set to be extremely strong,

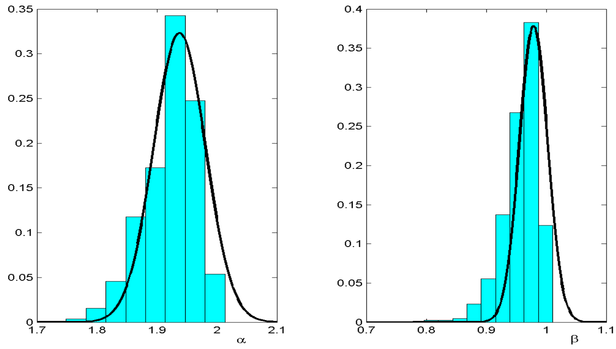

Figure 6: Gaussian and empirical posterior distributions for parameters $\alpha = 1.95$ and $\beta = 0.95$.

Table 2: Initial values of the chain (point estimates of the parameters), posterior means and rejection ratio for two chains with ML- and MAP-based covariance matrices for the random walk proposal.

	Starting values			
	α	β	γ	δ
ML	1.2811	-0.6547	0.8644	0.4280
MAP	1.7252	-0.5483	0.9727	0.1902
	Posterior means			
	α	β	γ	δ
ML	1.7187	-0.5372	1.0459	0.1402
MAP	1.7145	-0.5350	1.0153	0.1174
	Rejection ratio			
ML	0.935			
MAP	0.588			

namely $\mathcal{B}\rceil(200, 30)$ and $\mathcal{B}\rceil(30, 100)$; on the other hand, the priors for the scale and location parameters were chosen to be, respectively, $\Gamma(2, 3)$ and $\mathcal{N}(0, 5)$. The chain was run for 1100 iterations, with burn-in of 100, with both the ML and the MAP covariance matrices for the random walk proposal. Results are reported in table 2, and the corresponding ergodic means for parameters α and β are displayed in figure 7.

2.2 Convergence issues

It is known from the general MCMC theory that random walk Metropolis sampler cannot attain uniform convergence. To demonstrate that it attains geometric convergence, one should be able to prove that the target distribution has exponentially decaying tails. In the case of interest here, this is not trivial, since the posterior distribution cannot be expressed in closed form; however, one can exploit the fact that the distribution of the maximum likelihood estimator is asymptotically normal: provided one does not use a heavy-tailed prior, the

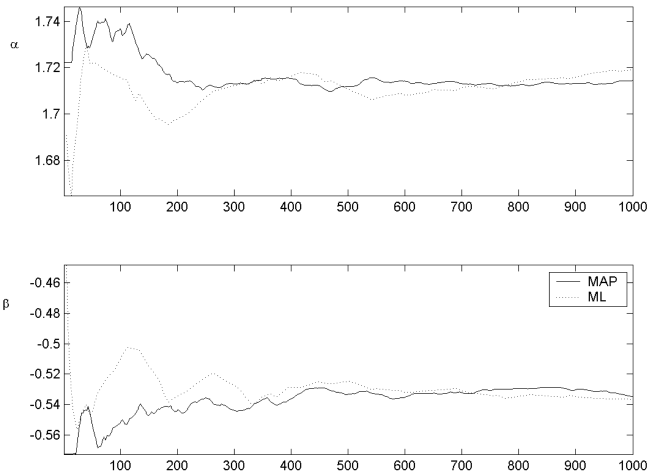

Figure 7: Ergodic means for α and β for two chains with MAP- and ML-based random walk proposal

posterior distribution should be asymptotically normal too. This means that, in the case of a finite number of observations, the posterior distribution should lie in the domain of attraction of a Gaussian law and thus possess finite second moment. In order to fulfill this requirement, its tails should be, at worst, proportional to $|x|^{-(3+\eta)}$, with $\eta > 0$ and arbitrarily small, and thus bounded by a multiple of $|x|^{-3}$. This means that the chain is polynomially ergodic with rate 1.

3 A practical example

As an illustration on how the proposed method works on real-world data, we have carried out the estimation of a sample of audio noise drawn from a set of recordings of songs taken by Robert Lachmann in Palestine in early twentieth century by means of a mobile recording studio (Lachmann 1929); as one might guess, the audio medium is very degraded and the noise is extremely heavy tailed. The audio sample, consisting of 44487 observations, is the same used in Lombardi & Godsill (2004), to which we refer for further details. In the same paper, it is also shown that α-stable distributions are especially well-suited to model this kind of noise, outperforming other more widespread heavy-tailed distributions such as the Student's t. The histogram and the time-series plot of the data, displayed in figure 8, highlight its heavy-tail features.

In a case with such a large number of observations, the Gibbs sampler of Buckle (1995) would be extremely slow; our approach instead proved to be very fast, requiring only 0.07 seconds per iteration of the chain. After the usual coarse maximum likelihood step, a Metropolis random walk chain was started from the maximum likelihood parameter vector and run for 11000 iterations, with burn-in of 1000. The priors chosen for the experiment were uniforms on

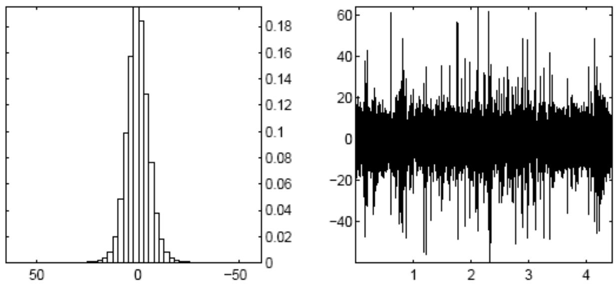

Figure 8: Histogram and pattern of the real audio sample.

Table 3: Descriptive statistics of the posterior distribution of the α-stable parameters for the noise sample.

	Mean	Std. dev.	5% quantile	95% quantile
α	1.8174	0.0069	1.8063	1.8289
β	-0.2306	0.0311	-0.2810	-0.1788
γ	3.5881	0.0153	3.5679	3.6202
δ	0.2179	0.0305	0.1681	0.2661

the whole support for α and β, inverse gamma with parameters $a = 1$ and $b = 1$ for γ and a $\mathcal{N}(0, 3)$ for δ. Descriptive statistics of the posterior distribution are reported in table 3; the evolution of the ergodic mean of the chain is displayed in figure 9.

4 Concluding remarks

In this paper, I have presented a random walk Metropolis MCMC scheme for the parameters of α-stable distributions. Although it is based on an approximated version of the likelihood, this approach was shown to perform remarkably well, being as twice as fast than the Gibbs sampler proposed by Buckle (1995). It is possible to envisage that the availability of a faster MCMC scheme will promote the use of α-stable distributions among practitioners and followers of the Bayesian paradigm. This contribution is a very first step in what appears to be a very promising direction; future research will aim at extending this approach to regression and time series models.

In this paper, we have employed prior distributions on the parameters independent of each other. A few words shall also be spent on this issue, especially for what concerns α and β. It is quite troublesome to analyze the dependency structure of these two parameters, mainly because, as α approaches 2, β tends to become unidentified. Most of the authors have thus preferred to bypass the problem by assuming independency between the two parameters. This choice is however unsatisfactory, and my opinion is that an appropriate joint prior should be employed. This also will be the subject of further research.

Given that they have four parameters instead of two and that they can accommodate asymmetry and heavy tails, there is no surprise that α-stable

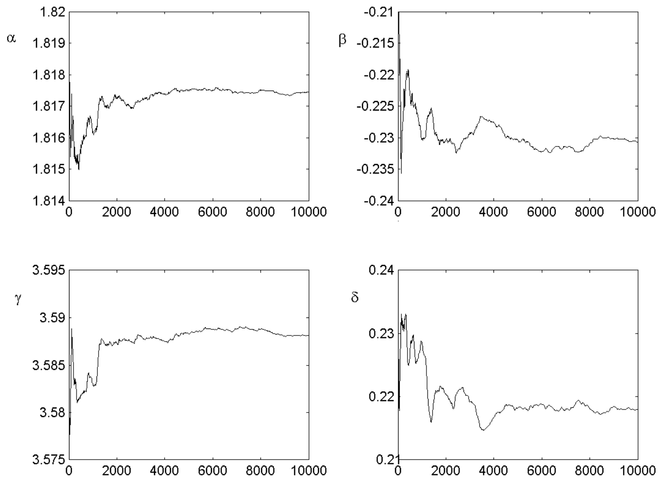

Figure 9: Evolution of the ergodic means for the real audio sample.

distributions should fit data better than a normal. Since in most of the cases the empirical distributions one observes in real world data have a mild degree of leptokurtosis, it is however very important to develop inferential schemes to discern Gaussian from α-stable distributions. A Bayesian approach to the problem could be to construct a reversible-jump Markov chain to obtain the posterior probabilities of both alternative models. This will be the subject of future research.

References

Bergstrøm, H. (1952), 'On some expansions of stable distribution functions', *Arkiv für Mathematik* **2**, 375–378.

Buckle, D. J. (1995), 'Bayesian inference for stable distributions', *Journal of the American Statistical Association* **90**, 605–613.

DuMouchel, W. H. (1973), 'On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution', *Annals of Statistics* **1**, 948–957.

Gnedenko, B. V. & Kolmogorov, A. N. (1954), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading.

Lachmann, R. (1929), *Musik des Orients*, Athenaion, Potsdam.

Lombardi, M. J. & Godsill, S. J. (2004), On-line Bayesian estimation of AR signals in symmetric α-stable noise. Working paper 2004/05, Dipartimento di Statistica "G. Parenti", Università degli studi di Firenze.

McCulloch, J. H. (1986), 'Simple consistent estimators of stable distribution parameters', *Communications in Statistics – Simulation and Computation* **15**, 1109–1136.

Mittnik, S., Rachev, S. T., Doganoglu, T. & Chenyao, D. (1999), 'Maximum likelihood estimation of stable paretian models', *Mathematical and Computer Modelling* **29**, 275–293.

Tsionas, E. G. (1999), 'Monte Carlo inference in econometric models with symmetric stable distributions', *Journal of Econometrics* **88**, 365–401.

Zolotarev, V. M. (1986), *One-dimensional Stable Distributions*, American Mathematical Society, Providence.

Copyright © 2004
Marco J. Lombardi