

Proceeding Paper

Expectation-Maximization Algorithm for Autoregressive Models with Cauchy Innovations [†]

Monika S. Dhull * and Arun Kumar

Department of Mathematics, Indian Institute of Technology Ropar, Rupnagar 140001, India;
arun.kumar@iitrpr.ac.in

* Correspondence: 2018maz0005@iitrpr.ac.in

[†] Presented at the 8th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 27–30 June 2022.

Abstract: In this paper, we study the autoregressive (AR) models with Cauchy distributed innovations. In the AR models, the response variable y_t depends on previous terms and a stochastic term (the innovation). In the classical version, the AR models are based on normal distribution which could not capture the extreme values or asymmetric behavior of data. In this work, we consider the AR model with Cauchy innovations, which is a heavy-tailed distribution. We derive closed forms for the estimates of parameters of the considered model using the expectation-maximization (EM) algorithm. The efficacy of the estimation procedure is shown on the simulated data. The comparison of the proposed EM algorithm is shown with the maximum likelihood (ML) estimation method. Moreover, we also discuss the joint characteristic function of the AR(1) model with Cauchy innovations, which can also be used to estimate the parameters of the model using empirical characteristic function.

Keywords: autoregressive model; Cauchy innovations; EM algorithm



Citation: Dhull, M.S.; Kumar, A. Expectation-Maximization Algorithm for Autoregressive Models with Cauchy Innovations. *Eng. Proc.* **2022**, *18*, 21. <https://doi.org/10.3390/engproc2022018021>

Academic Editors: Ignacio Rojas, Hector Pomares, Olga Valenzuela, Fernando Rojas and Luis Javier Herrera

Published: 21 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Autoregressive (AR) models with stable and heavy-tailed innovations are of great interest in time series modeling. These distributions can easily assimilate the asymmetry, skewness, and outliers present in time series data. The Cauchy distribution is a special case of stable distribution with undefined expected value, variance, and higher order moments. The Cauchy distribution and its mixture has many applications in the field of economics [1], seismology [2], biology [3], and various other fields, but only a few studies have been conducted concerning time series models with Cauchy errors. In [4], the maximum likelihood (ML) estimation of AR(1) model with Cauchy errors is studied.

The standard estimation techniques for the AR(p) model with Cauchy innovations, particularly the Yule–Walker method and conditional least squares method, cannot be used due to the infinite second order moments of the Cauchy distribution. Therefore, it is worthwhile to study and assess the alternate estimation techniques for the AR(p) model with Cauchy innovations. In the literature, several estimation techniques have been proposed to estimate the parameters of AR models with infinite variance errors (see e.g., [5–7]).

In this paper, we propose to use the EM algorithm to estimate the parameters of the distribution and model simultaneously. It is a general iterative algorithm for model parameter estimation which iterates between two steps, namely the expectation step (E-step) and the maximization step (M-step) [8]. It is an alternative to the numerical optimization of the likelihood function which is proven to be numerically stable [9]. We also provide the formula based on the characteristic function (CF) and empirical characteristic function (ECF) of Cauchy distribution for AR(p) model estimation. The idea to use ECF in a time series stable ARMA model has been discussed in [10].

The remainder of the paper is organized as follows. In Section 2, we present a brief overview of the Cauchy AR(p) model, followed by a discussion of estimation techniques, namely the EM algorithm and estimation by CF and ECF. Section 3 checks the efficacy of the estimation procedure on simulated data. We also present the comparative study, where the proposed technique is compared with ML estimation for Cauchy innovations. Section 4 concludes the paper.

2. Cauchy Autoregressive Model

We consider the AR(p) univariate stationary time-series $\{Y_t\}$, $t \in \mathbb{N}$ with Cauchy innovations defined as

$$Y_t = \sum_{i=1}^p \rho_i Y_{t-i} + \varepsilon_t = \rho^T Y_{t-1} + \varepsilon_t, \quad (1)$$

where $\rho = (\rho_1, \rho_2, \dots, \rho_p)^T$ is a p -dimensional column vector, $Y_{t-1} = (Y_{t-1}, Y_{t-2}, \dots, Y_{t-p})^T$ is a vector of p lag terms, and $\{\varepsilon_t\}$, $t \in \mathbb{N}$ are i.i.d. innovations distributed as Cauchy(α, γ). The pdf of Cauchy(α, γ) [11] is

$$f(x; \alpha, \gamma) = \frac{\gamma}{\pi} \left[\frac{1}{\gamma^2 + (x - \alpha)^2} \right], \quad \gamma > 0, \alpha \in \mathbb{R}, x \in \mathbb{R}. \quad (2)$$

The conditional distribution of Y_t given the preceding data $\mathcal{F}_{t-1} = (Y_{t-1}, Y_{t-2}, \dots, Y_1)^T$ is given by [4]

$$p(Y_t | \mathcal{F}_{t-1}) = f(y_t - \rho^T y_{t-1}; \alpha, \gamma) = \frac{\gamma}{\pi} \left[\frac{1}{\gamma^2 + (y_t - \rho^T y_{t-1} - \alpha)^2} \right],$$

where y_{t-1} is the realization of Y_{t-1} . In the next subsection, we propose the methods to estimate the model parameters ρ and innovation parameters α and γ simultaneously.

2.1. Parameter Estimation Using EM Algorithm

We estimate the parameters of the AR(p) model using an EM algorithm which maximizes the likelihood function iteratively. Further, we discuss the time series $\{Y_t\}$ using the characteristic function (CF) and estimation method using CF and ECF. Recently [7], the exponential-squared estimator for AR models with heavy-tailed errors was introduced and proven to be \sqrt{n} -consistent under some regularity conditions; similarly, the self-weighted least absolute deviation estimation method was also studied for the infinite variance AR model [12]. The ML estimation of AR models with Cauchy errors with intercept and with linear trend is studied, and the AR coefficient is shown to be $n^{3/2}$ -consistent under some conditions [4]. For the AR(p) model with Cauchy innovations with n samples, the log likelihood is defined as

$$l(\Theta) = (n - p) \log(\gamma) - (n - p) \log(\pi) - \sum_{t=p+1}^n \log(\gamma^2 + (y_t - \rho^T Y_{t-1} - \alpha)^2),$$

where $\Theta = (\alpha, \gamma, \rho^T)$.

Proposition 1. Consider the AR(p) time-series model given in Equation (1) where error terms follow Cauchy(α, γ). The maximum likelihood estimates of the model parameters using EM algorithm are as follows

$$\hat{\rho}^T = \left(\sum_{t=p+1}^n \frac{(y_t - \alpha) Y_{t-1}^T}{s_t^{(k)}} \right) \left(\sum_{t=p+1}^n \frac{Y_{t-1} Y_{t-1}^T}{s_t^{(k)}} \right)^{-1};$$

$$\hat{\alpha} = \frac{\sum_{t=p+1}^n \frac{(y_t - \rho^T Y_{t-1})}{s_t^{(k)}}}{\sum_{t=p+1}^n \frac{1}{s_t^{(k)}}};$$

and

$$\hat{\gamma} = \sqrt{\frac{n-p}{2 \sum_{t=p+1}^n \frac{1}{s_t^{(k)}}}}$$

where $s_t^{(k)} = (y_t - \rho^{T(k)} Y_{t-1} - \alpha^{(k)})^2 + \gamma^{(k)2}$.

Proof. Consider the AR(p) model

$$Y_t = \rho^T Y_{t-1} + \varepsilon_t, \quad t = p + 1, 2, \dots, n,$$

where ε_t follows Cauchy distribution Cauchy(α, γ). Let (ε_t, V_t) for $t = 1, 2, \dots, n$ denote the complete data for innovations ε . The observed data ε_t are assumed to be from Cauchy(α, γ) and the unobserved data V_t follow inverse gamma IG($1/2, \gamma^2/2$). A random variable $V \sim \text{IG}(a, b)$ if the pdf is given by

$$f_V(v; a, b) = \frac{b^a}{\Gamma(a)} \frac{e^{-b/v}}{v^{a+1}}, \quad a > 0, b > 0, v > 0.$$

We can rewrite ε_t as $\varepsilon_t = Y_t - \rho^T Y_{t-1}$, for $t = 1, 2, \dots, n$. The stochastic relation $\varepsilon = \alpha + \sqrt{V}Z$ with $Z \sim N(0, 1)$ i.e., standard normal and $V \sim \text{IG}(1/2, \gamma^2/2)$ is used to generate Cauchy(α, γ) distribution. Then, the conditional distribution is

$$f(\varepsilon = \varepsilon_t | V = v_t) = \frac{1}{\sqrt{2\pi v_t}} \exp\left(-\frac{1}{2v_t} (\varepsilon_t - \alpha)^2\right).$$

Now, we need to estimate the unknown parameters $\Theta = (\alpha, \gamma, \rho^T)$. To apply the EM algorithm for estimation, we first find the conditional expectation of log-likelihood of complete data (ε, V) with respect to the conditional distribution of V given ε . As the unobserved data are assumed to be from IG($1/2, \gamma^2/2$), the posterior distribution is again an inverse gamma i.e.,

$$V | \varepsilon = e, \Theta \sim \text{IG}\left(1, \frac{(e - \alpha)^2 + \gamma^2}{2}\right).$$

The following conditional inverse first moment and $\mathbb{E}(\log(V) | \varepsilon = e)$ will be used in calculating the conditional expectation of the log-likelihood function:

$$\mathbb{E}(V^{-1} | \varepsilon = e) = \frac{2}{(e - \alpha)^2 + \gamma^2},$$

$$\mathbb{E}(\log(V) | \varepsilon = e) = \log((e - \alpha)^2 + \gamma^2) - \log 2 + 0.5776.$$

The complete data likelihood is given by

$$L(\Theta) = \prod_{t=p+1}^n f(\epsilon_t, v_t) = \prod_{t=p+1}^n f_{\epsilon|V}(\epsilon_t|v_t)f_V(v_t)$$

$$= \prod_{t=p+1}^n \frac{\gamma}{2\pi v_t} \exp\left(-\frac{(\epsilon_t - \alpha)^2 + \gamma^2}{2v_t}\right).$$

The log likelihood function will be

$$l(\Theta) = (n - p) \log(\gamma) - (n - p) \log(2\pi) - 2 \sum_{t=p+1}^n \log(v_t) - \sum_{t=p+1}^n \frac{(\epsilon_t - \alpha)^2 + \gamma^2}{2v_t}.$$

Now, we will use the relation $\epsilon_t = Y_t - \rho^T Y_{t-1}$ in further calculations. In the first step at k th iteration, E-step of EM algorithm, we need to compute the expected value of the complete data log likelihood known as $Q(\Theta|\Theta^{(k)})$, which is expressed as

$$Q(\Theta|\Theta^{(k)}) = \mathbb{E}_{V|\epsilon, \Theta^{(k)}} [\log f(\epsilon, V|\Theta)|\epsilon_t, \Theta^{(k)}] = \mathbb{E}_{V|\epsilon, \Theta^{(k)}} [l(\Theta|\Theta^{(k)})]$$

$$= (n - p) \log \gamma - (n - p) \log 2\pi - \sum_{t=p+1}^n \mathbb{E}(\log v_t|\epsilon_t, \Theta^{(k)})$$

$$- \frac{\gamma^2}{2} \sum_{t=p+1}^n \mathbb{E}(v_t^{-1}|\epsilon_t, \Theta^{(k)}) - \frac{1}{2} \sum_{t=p+1}^n (\epsilon_t - \alpha)^2 \mathbb{E}(v_t^{-1}|\epsilon_t, \Theta^{(k)})$$

$$= (n - p) \log \gamma - (n - p) \log 2\pi - (n - p) \log 2 + 0.5776(n - p) - \sum_{t=p+1}^n \log((\epsilon_t - \alpha^{(k)})^2 + \gamma^{(k)2})$$

$$- \frac{\gamma^2}{2} \sum_{t=p+1}^n \frac{1}{(\epsilon_t - \alpha^{(k)})^2 + \gamma^{(k)2}} - \sum_{t=p+1}^n \frac{(\epsilon_t - \alpha)^2}{(\epsilon_t - \alpha^{(k)})^2 + \gamma^{(k)2}}$$

$$= (n - p) \log \gamma - (n - p) \log 2\pi - (n - p) \log 2 + 0.5776(n - p) - \sum_{t=p+1}^n \log(s_t^{(k)})$$

$$- \frac{\gamma^2}{2} \sum_{t=p+1}^n \frac{1}{s_t^{(k)}} - \sum_{t=p+1}^n \frac{(\epsilon_t - \alpha)^2}{s_t^{(k)}}.$$

where $s_t = (y_t - \rho^T Y_{t-1} - \alpha)^2 + \gamma^2$. In the next M-step, we estimate the parameters α, γ , and ρ^T by maximizing the Q function using the equations below:

$$\frac{\partial Q}{\partial \rho} = 4 \sum_{t=p+1}^n \frac{(y_t - \rho^T Y_{t-1} - \alpha Y_{t-1}^T)}{s_t^{(k)}},$$

$$\frac{\partial Q}{\partial \alpha} = \sum_{t=p+1}^n \frac{(y_t - \rho^T Y_{t-1} - \alpha)}{s_t^{(k)}},$$

$$\frac{\partial Q}{\partial \gamma} = \frac{n - p}{\gamma} - 2\gamma \sum_{t=p+1}^n \frac{1}{s_t^{(k)}}.$$

Solving the above equations at each iteration, we find the following closed form estimates of the parameters at $k + 1$ th iteration:

$$\hat{\rho}^T = \left(\sum_{t=p+1}^n \frac{(y_t - \alpha) Y_{t-1}^T}{s_t^{(k)}} \right) \left(\sum_{t=p+1}^n \frac{Y_{t-1} Y_{t-1}^T}{s_t^{(k)}} \right)^{-1};$$

$$\hat{\alpha} = \frac{\sum_{t=p+1}^n \frac{(y_t - \rho^T Y_{t-1})}{s_t^{(k)}}}{\sum_{t=p+1}^n \frac{1}{s_t^{(k)}}}; \tag{4}$$

and

$$\hat{\gamma} = \sqrt{\frac{n - p}{2 \sum_{t=p+1}^n \frac{1}{s_t^{(k)}}}}$$

where $s_t^{(k)} = (y_t - \rho^T Y_{t-1} - \alpha^{(k)})^2 + \gamma^{(k)2}$. \square

2.2. Characteristic Function for Estimation

Thus far, we have considered the conditional distribution of Y_t given the preceding data \mathcal{F}_{t-1} . Now, we include the dependency of time series $\{Y_t\}$ by defining the variable $d_j = (y_j, \dots, y_{j+p})$ for $j = 1, \dots, n - p$. In each variable $\{d_j\}$, there are p terms the same as adjacent variable. The distribution of $\{d_j\}$ will be multivariate Cauchy with dimension $r = p + 1$.

The CF of each d_j is $c(\Theta, s) = \mathbb{E}(\exp(i s^T d_j))$ and the ECF is $c_n(s) = \frac{1}{n} \sum_{j=1}^{n-p} \exp(i s^T d_j)$ where $s = (s_1, \dots, s_{p+1})^T$.

To estimate the parameters using CF and ECF, we make sure that the joint CF of the AR(p) model has a closed form. In the next result, the closed form expression for the joint CF of the AR(1) model with Cauchy innovations is given.

Proposition 2. *The joint CF of stationary AR(1) model with Cauchy innovations is*

$$c(s_1, s_2; \Theta) = \exp \left[i \alpha (s_1 + s_2) \left(\frac{1}{1 - \rho} \right) \right] \times \exp \left[-\gamma \left(|s_2| + |s_1 + \rho s_2| \left(\frac{1}{1 - |\rho|} \right) \right) \right].$$

Proof. For stationary AR(1) model $y_t = \rho y_{t-1} + \varepsilon_t$, we can rewrite it as

$$y_t = \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \rho^3 \varepsilon_{t-3} + \dots$$

Note that $\{\varepsilon_t\}$ are i.i.d from Cauchy(α, γ) distribution, and the CF of Cauchy(α, γ) is $\mathbb{E}(\exp(i s \varepsilon_t)) = \exp(i \alpha s - \gamma|s|)$ [11]. Then, the joint CF of (y_t, y_{t-1}) is calculated as follows:

$$\begin{aligned}
 c(s_1, s_2; \Theta) &= \mathbb{E}[\exp(i s_1 y_{t-1} + i s_2 y_t)] \\
 &= \mathbb{E}[\exp(i s_1 (\varepsilon_{t-1} + \rho \varepsilon_{t-2} + \rho^2 \varepsilon_{t-3} + \dots)) + i s_2 (\varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots)] \\
 &= \mathbb{E}[\exp(i s_2 \varepsilon_t + i (s_1 + s_2 \rho) \varepsilon_{t-1} + i \rho (s_1 + \rho s_2) \varepsilon_{t-2} + \dots)] \\
 &= \exp(i \alpha s_2 - \gamma |s_2|) \times \exp(i \alpha (s_1 + \rho s_2) - \gamma |s_1 + \rho s_2|) \\
 &\quad \times \exp(i \alpha (\rho s_1 + \rho^2 s_2) - \gamma |\rho s_1 + \rho^2 s_2|) \times \dots \\
 &= \exp(i \alpha (s_1 + s_2) (1 + \rho + \rho^2 + \rho^3 + \dots)) \\
 &\quad \times \exp(-\gamma (|s_2| + |s_1 + \rho s_2| + |\rho| |s_1 + \rho s_2| + |\rho^2| |s_1 + \rho s_2| + \dots)) \\
 &= \exp(i \alpha (s_1 + s_2) \left(\frac{1}{1 - \rho} \right)) \times \exp(-\gamma (|s_2| + |s_1 + \rho s_2| (1 + |\rho| + |\rho^2| + \dots))) \\
 &= \exp \left[i \alpha (s_1 + s_2) \left(\frac{1}{1 - \rho} \right) \right] \times \exp \left[-\gamma \left(|s_2| + |s_1 + \rho s_2| \left(\frac{1}{1 - |\rho|} \right) \right) \right].
 \end{aligned}$$

□

The joint CF for a higher dimension can be obtained in similar manner. Now, the model parameters can be estimated by solving the following integral with CF and ECF as defined in [10]:

$$\int \dots \int w_{\Theta}(s) (c_n(s) - c(s; \Theta)) ds = 0. \tag{5}$$

where optimal weight function

$$w_{\Theta}^*(s) = \frac{1}{2\pi} \int \dots \int \exp(-i s^T d_j) \frac{\partial \log f(y_{j+p} | y_j, \dots, y_{j+p-1})}{\partial \Theta} dy_j \dots dy_{j+p}. \tag{6}$$

Remark 1. For a stationary AR(l) process $\{Y_t\}$ with $p = l$, the ECF estimator defined by Equation (5) with optimal weight function defined in Equation (6) is a conditional ML (CML) estimator and hence asymptotically efficient. The conditional log pdf for Cauchy distribution is:

$$\log f(y_{j+p} | y_j, \dots, y_{j+p-1}) = \log \gamma - \log \pi - \log \left(\gamma^2 + (y_t - \rho^T Y_{t-1} - \alpha^2)^2 \right).$$

The proof is similar to the proof of Proposition 2.1 in [10].

3. Simulation Study

In this section, we assess the proposed model and the introduced estimation technique using a simulated data set. We discuss the estimation procedure for the AR(2) model with Cauchy innovations. The AR(2) model defined in (1) is simulated with ρ_1 and ρ_2 as model parameters. We generate 1000 trajectories, each of size $N = 500$ of Cauchy innovations using the normal variance-mean mixture form $\varepsilon = \alpha + \sqrt{V}Z$ with $Z \sim N(0, 1)$, i.e., standard normal and $V \sim \text{IG}(1/2, \gamma^2/2)$. We then use the following simulation steps to generate the Cauchy innovations:

- step 1: Generate standard normal variate Z ;
- step 2: Generate inverse gamma random variate $\text{IG}(1/2, \gamma^2/2)$ with $\gamma = 2$;
- step 3: Using the relation $\varepsilon = \alpha + \sqrt{V}Z$, we simulate the Cauchy innovations with $\alpha = 1$;
- step 4: The time series data y_t is generated with model parameters $\rho_1 = 0.5$ and $\rho_2 = 0.3$.

The exemplary time series data plot and scatter plot of innovation terms are shown in Figure 1. We apply the discussed EM algorithm to estimate the model parameters and distribution parameters. The relative change in the parameters is considered to terminate the algorithm. The following is the stopping criteria which is commonly used in literature:

$$\max \left\{ \left| \frac{\alpha^{(k+1)} - \alpha^{(k)}}{\alpha^{(k)}} \right|, \left| \frac{\gamma^{(k+1)} - \gamma^{(k)}}{\gamma^{(k)}} \right|, \left| \frac{\rho_1^{(k+1)} - \rho_1^{(k)}}{\rho_1^{(k)}} \right|, \left| \frac{\rho_2^{(k+1)} - \rho_2^{(k)}}{\rho_2^{(k)}} \right| \right\} < 10^{-4}. \quad (7)$$

We compare the estimation results of Cauchy(α, γ) with the EM algorithm and maximum likelihood (ML) estimation. The ML estimates are computed using the inbuilt function “mlcauchy” in R, which uses the exponential transform of the location parameter and performs non-linear minimization by a Newton-type algorithm. The comparison of the estimates of Cauchy(α, γ) are shown in boxplots in Figure 2. From the boxplots, we find that the EM algorithm converges near to the true value of the Cauchy(α, γ) as compared to the ML estimation. There is a possibility of achieving a better result from ML method if a different algorithm or inbuilt function for optimization are used for estimation.

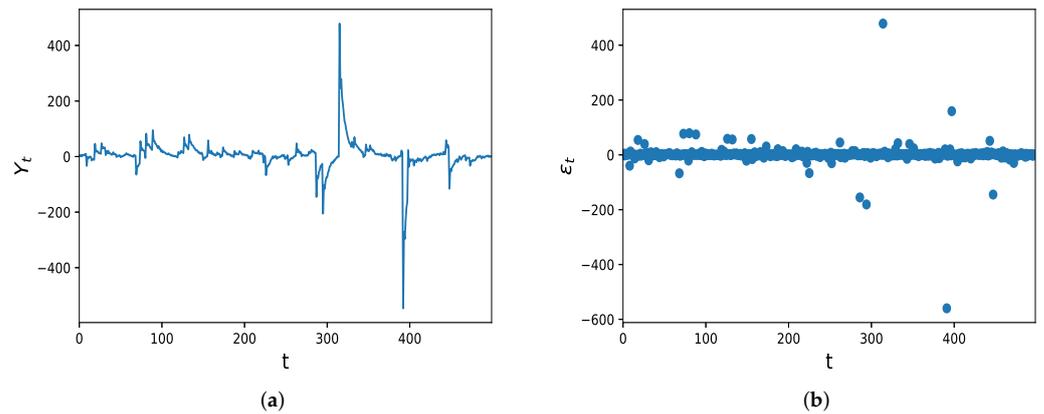


Figure 1. (a) The data plot of exemplary time series of length $N = 500$ and (b) the scatter plot of the corresponding innovation terms of the AR(2) model with Cauchy innovations. The chosen parameters of the model are $\rho_1 = 0.5$, $\rho_2 = 0.3$, $\alpha = 1$, and $\gamma = 2$.

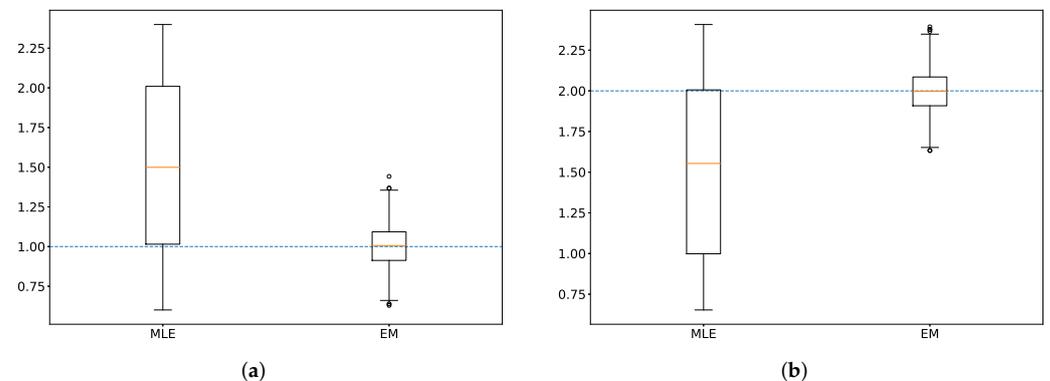


Figure 2. Boxplots of the estimates of the AR(2) model’s parameters with theoretical values (a) $\alpha = 1$ and (b) $\gamma = 2$ represented with blue dotted lines. The boxplots are created using 1000 trajectories, each of length 500.

4. Conclusions and Future Scope

In this work, we derive the closed form of estimates of AR model with Cauchy innovations using an EM algorithm. The performance of the proposed algorithm is compared with the ML method using simulated data. The ML estimation is found using an inbuilt function in R. Another benefit of using the EM algorithm is that it calculates the model as well as the innovation parameters simultaneously. It is evident from the boxplot that the EM algorithm outperforms the ML method. Further, we discuss another approach based on CF to estimate the AR model parameters with stable distribution. In the future, we plan to study and compare the proposed algorithm and ECF based estimation method with the existing techniques in [5–7] for an AR model with infinite variance. Further, the real life phenomena can be studied using the proposed model and methods.

Author Contributions: Conceptualization, A.K. and M.S.D.; Methodology, A.K. and M.S.D.; Writing-Original Draft Preparation, M.S.D. and A.K.; Writing Review & Editing, A.K. and M.S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: M.S.D. would like to thank her institute Indian Institute of Technology Ropar, India and the Ministry of Education (MoE), Government of India, for supporting her research work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, T.; Zhang, P.; Dai, W.S.; Xie, M. An intermediate distribution between Gaussian and Cauchy distributions. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 5411–5421. [[CrossRef](#)]
2. Kagan, Y.Y. Correlations of earthquake focal mechanism. *Geophys. J. Int.* **1992**, *110*, 305–320. [[CrossRef](#)]
3. Bjerkedal, T. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am. J. Hyg.* **1960**, *72*, 130–148. [[PubMed](#)]
4. Choi, J.; Choi, I. Maximum likelihood estimation of autoregressive models with a near unit root and Cauchy errors. *Ann. Inst. Stat. Math.* **2019**, *71*, 1121–1142. [[CrossRef](#)]
5. Jiang, Y. An exponential-squared estimator in the autoregressive model with heavy-tailed errors. *Stat. Interface* **2016**, *9*, 233–238. [[CrossRef](#)]
6. Li, J.; Liang, W.; He, S.; Wu, X. Empirical likelihood for the smoothed LAD estimator in infinite variance autoregressive models. *Statist. Probab. Lett.* **2010**, *80*, 1420–1430. [[CrossRef](#)]
7. Tang, L.; Zhou, Z.; Wu, C. Efficient estimation and variable selection for infinite variance autoregressive models. *J. Appl. Math. Comput.* **2012**, *40*, 399–413. [[CrossRef](#)]
8. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via EM Algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–22. .
9. McLachlan, G.J.; Krishnan, T. *The EM Algorithm and Extensions*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2007.
10. Yu, J.; Knight, J.L. Empirical characteristic function in time series estimation. *Econom. Theory* **2002**, *18*, 691–721.
11. Feller, W. *An Introduction to Probability Theory and Its Applications*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 1991; Volume 2.
12. Ling, S. Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 381–393. [[CrossRef](#)]