Statistics and Its Interface Volume 9 (2016) 233–238

An exponential-squared estimator in the autoregressive model with heavy-tailed errors

Yunlu Jiang

In this paper, an exponential-squared estimator is introduced in the autoregressive model with heavy-tailed errors. Under some conditions, the $\sqrt{n}$-consistency of the proposed estimator is established. Since the exponential-squared estimator involves a tuning parameter λ, we select λ via a five-fold cross validation procedure. Simulation studies illustrate that the finite sample performance of proposed method performs better than that of a self-weighted composite quantile regression (SWCQR) method and self-weighted least absolute deviation (SWLAD) method in terms of Sd and MSE when the error follows a heavy-tailed distribution and there are outliers in the dataset. Finally, we apply the proposed methodology to analyze the Recruitment series.

AMS 2000 subject classifications: Primary 62G20; secondary 62G35.
Keywords and phrases: Autoregressive model, Exponential squared loss, Heavy-tailed distribution.

1. INTRODUCTION

In the time series area, the autoregressive (AR) models have played a very important role in modern Statistics, and have successfully applied to a variety of research fields, such as social sciences, economics, finance, business, demographics, and meteorology. The estimation of the order is an outstanding statistical problem in an autoregressive model. However, heavy-tailed time series data is often encountered in economics and finance [6], hydrology [1], and teletraffic data [12]. In the last several decades, there were many robust estimators proposed and widely studied for the AR models in the literature. For instance, [3] introduced M-estimation for autoregressions with infinite variance. [8] studied a SWLAD estimator. [7] proposed an empirical likelihood method to estimate the unknown parameters of infinite variance autoregressive models. [2] considered a weighted quantile regression for AR models to tackle with infinite variance errors. [10] developed a SWCQR estimation procedure to estimate unknown parameter in an infinite variance autoregressive model. [5] studied the least tail-trimmed squares estimator for infinite variance autoregressions.

Based on the exponential squared loss function [4, 11], [11] introduced a robust variable selection in a linear regression model, and studied their robustness. [11] showed that the proposed estimation approach not only had very good robustness when there were outliers in the dataset or the error followed a heavy-tailed distribution, but also was as asymptotically efficient as the least squares method under normal error. To obtain a robust and efficient estimator for AR models, we propose an exponential-squared estimator based on the exponential squared loss function for the autoregressive models. We show that the proposed estimator possesses $\sqrt{n}$-consistency. Since the exponential squared loss function has a tuning parameter, we use a 5-fold cross validation to select the tuning parameter. Simulation studies and a real data analysis demonstrate that the proposed estimator can obtain high efficiency and robust when the error follows a heavy-tailed distribution.

The rest of this paper is organized as follows. In Section 2, we introduce an exponential-squared estimator in an AR model, and give the $\sqrt{n}$-consistency of the proposed estimator. Meanwhile, we use a 5-fold cross validation to select the tuning parameter. In Section 3, simulation studies and a real data analysis are conducted to compare the finite-sample performance of the existing and proposed methods. We conclude with a few remarks in Section 4. A proof is given in the Appendix.

2. METHODOLOGY AND MAIN RESULTS

2.1 An exponential-squared estimator

Suppose that $\{Y_1, \cdots, Y_{n+p}\}$ satisfying a following $AR(p)$ model,

$$\begin{aligned} Y_i &= \phi_1 Y_{i-1} + \phi_2 Y_{i-2} + \cdots + \phi_p Y_{i-p} + \epsilon_i \\ &= \mathbf{X}_i^T \boldsymbol{\phi} + \epsilon_i, i = p+1, \cdots, p+n, \end{aligned} \tag{2.1}$$

where $\boldsymbol{\phi} = (\phi_1, \phi_2, \cdots, \phi_p)^T$ is the unknown autoregressive coefficient, $\mathbf{X}_i = (Y_{i-1}, \cdots, Y_{i-p})^T$, and $\{\epsilon_i, i = p+1, \cdots, p+n\}$ is a sequence of independent and identically distributed random variables with mean 0. We assume that there is a correct model with the true autoregression coefficient $\boldsymbol{\phi}_0$.

Ideally, if the error term follows a normal distribution, the ordinary least squares (OLS) estimator is reputed to be an efficient estimator. Unfortunately, one seldom knows the true density function of the error term. Specially, if the error ϵ_t follows a heavy-tailed distribution and there are outliers in the dataset, the least squares estimator is not very robust.

Recently, [10] proposed a SWCQR method, and showed that the resulting SWCQR estimator was more robust and efficient than the OLS estimator when the error followed a heavy-tailed distribution. For a given positive integer q, let $\tau_k = k/(q+1), k = 1, \cdots, q$. Note that the conditional τ_k quantile of Y_i under given $\mathbf{X}_i$ is $\mathbf{X}_i^T\boldsymbol{\phi} + b_k^*$, where b_k^* is τ_k quantile of ϵ_i. The SWCQR estimator of $\boldsymbol{\phi}$ can be obtained by minimizing the following function with respect to $\boldsymbol{\phi}$ and $b_1, \cdots, b_q$,

$$\sum_{k=1}^{q}\sum_{t=p+1}^{n+p}\omega_t\rho_{\tau_k}\left(Y_t - b_k - \sum_{j=1}^{p}\phi_j Y_{t-j}\right), \tag{2.2}$$

where $\rho_\tau(u) = \tau u - uI(u<0)$, and ω_t is given in [10].

Next, we will introduce an exponential-squared estimator by using the exponential squared loss function. The exponential-squared estimator is defined as

$$\hat{\boldsymbol{\phi}}_n = arg\max_{\boldsymbol{\phi}}\sum_{t=p+1}^{n+p}\exp\left\{-(Y_t - \sum_{j=1}^{p}\phi_j Y_{t-j})^2/\lambda\right\}, \tag{2.3}$$

where λ is a tuning parameter, and controls the degree of robust and efficiency. We will select the tuning parameter by a 5-fold cross validation in Section 2.2. For convenience, we call our proposed estimator as the ESL estimator. Next, we study the asymptotic properties of the ESL estimator. Before presenting asymptotic properties, we list some conditions:

(C1) The characteristic polynomial $1 - \phi_1 t - \cdots - \phi_p t^p$ has all roots outside the unit circle.
(C2) Let $\mathbf{X}_t = (Y_{t-1}, \cdots, Y_{t-p})^T$, $\epsilon_t = Y_t - \mathbf{X}_t^T\boldsymbol{\phi}_0$, $I(\boldsymbol{\phi}_0)$ is negative definite, where

$$I(\boldsymbol{\phi}_0) = \frac{2}{\lambda}E\left[\mathbf{X}_t\mathbf{X}_t^T e^{-\epsilon_t^2/\lambda}\left(\frac{2\epsilon_t^2}{\lambda} - 1\right)\right].$$

Theorem 2.1. *Assume that conditions (C1)and (C2) hold. There exists a local maximum $\hat{\boldsymbol{\phi}}_n$ satisfying $\sqrt{n}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) = O_p(1)$.*

2.2 The choice of tuning parameter

From (2.3), we can see that the proposed estimator depends on the tuning parameter λ. Therefore, before we obtain the ESL estimator, we first should select the tuning parameter. There are many methods to select λ, such as cross-validation (CV), generalized cross-validation (GCV), AIC, and BIC. In this paper, we use a 5-fold cross validation procedure to select λ. Denote the full dataset by $\mathbf{D}$, and let the training and test set be $\mathbf{D} - \mathbf{D}^\nu$ and $\mathbf{D}^\nu$, respectively. For each λ and ν, we obtain the ESL estimator $\hat{\boldsymbol{\phi}}^{(\nu)}(\lambda)$ of $\boldsymbol{\phi}$ using the training set $\mathbf{D} - \mathbf{D}^\nu$. The 5-fold cross-validation criterion is defined as

$$CV(\lambda) = \sum_{\nu=1}^{5} mad\{Y_k - \hat{\boldsymbol{\phi}}^{(\nu)}(\lambda)^T\mathbf{X}_k, (Y_k, \mathbf{X}_k) \in \mathbf{D}^\nu\},$$

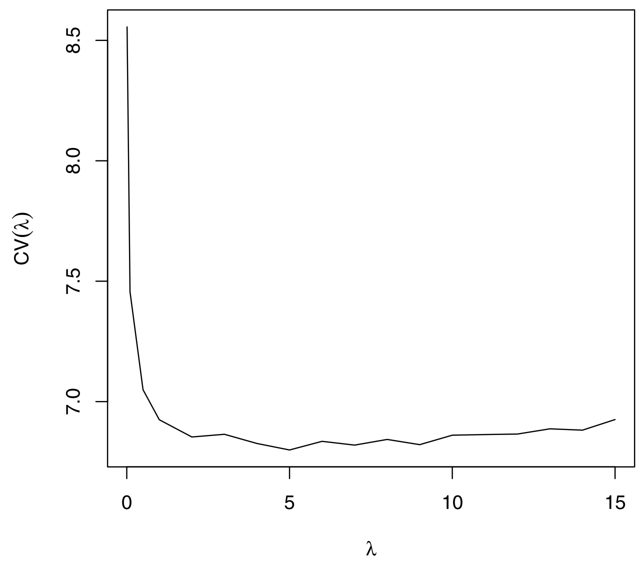

Figure 1. $CV(\lambda)$ against λ.

where $\mathbf{X}_k = (Y_{k-1}, \cdots, Y_{k-p})^T$, $mad\{\mathbf{V}_n\}$ is the median absolute deviation (MAD) estimator based on the dataset $\mathbf{V}_n$. We select λ that minimizes $CV(\lambda)$.

3. SIMULATION STUDY

In this section, we conduct simulation studies to examine the finite sample performance of the proposed estimator, and then demonstrate the proposed methodology by a real data analysis. We first explain how to select the tuning parameter λ. We choose $n = 500$, and the dataset is generated according to the following autoregressive model

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \epsilon_t, \tag{3.1}$$

where $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)^T = (0.5, 0, -0.7)^T$, and the error term ϵ_t follows a standard Cauchy distribution. The original 500 sample is randomly partitioned into 5 equal size subsamples. We plot $CV(\lambda)$ against the tuning parameter λ as depicted in Figure 1. From Figure 1, we obtain $\lambda = 5$ by minimizing the $CV(\lambda)$.

In the following, we evaluate the performance of various methods with different sample sizes. We simulate 500 data sets from the autoregressive model (2.1) with sample sizes of $n = 100, 200, 500$. In this simulation, we choose $AR(2)$ with $\boldsymbol{\phi} = (\phi_1, \phi_2)^T = (0.8, -0.3)^T$, and $AR(3)$ with $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)^T = (0.5, 0, -0.7)^T$. The dataset are generated by the following three mechanisms:

(1) the error term ϵ_t follows a standard normal distribution, $N(0, 1)$;

(2) the error term ϵ_t follows a t-distribution with degree of freedom 2, t_2; However, the the first 5% dataset of $\{Y_1, \cdots, Y_{n+p}\}$ are replaced by 100;

(3) we take the same setting as (2), except that the error term follows a Cauchy distribution.

Table 1. The mean, standard deviation (Sd), and MSE for the estimators of ϕ under normal error

n	Estimator	MSE	$\phi_1 = 0.8$ Mean(Sd)	$\phi_2 = -0.3$ Mean(Sd)
	ESL	0.0217	0.8113(0.1064)	−0.2909(0.1024)
100	SWCQR	0.0193	0.8018(0.0915)	−0.3043(0.1052)
	SWLAD	0.0283	0.7923(0.1104)	−0.2977(0.1276)
	LS	0.0163	0.8101(0.0873)	−0.2994(0.0934)
	ESL	0.0190	0.7715(0.1053)	−0.2779(0.0822)
200	SWCQR	0.0116	0.7770(0.0787)	−0.2865(0.0695)
	SWLAD	0.0169	0.7733(0.0998)	−0.2833(0.0783)
	LS	0.0103	0.7803(0.0744)	−0.2831(0.0645)
	ESL	0.0048	0.7952(0.0430)	−0.3111(0.0541)
500	SWCQR	0.0042	0.7975(0.0433)	−0.3054(0.0492)
	SWLAD	0.0059	0.7975(0.0520)	−0.3072(0.0566)
	LS	0.0039	0.7970(0.0397)	−0.2996(0.0486)

Table 2. The mean, standard deviation (Sd), and MSE for the estimators of ϕ under t_2 error

n	Estimator	MSE	$\phi_1 = 0.8$ Mean(Sd)	$\phi_2 = -0.3$ Mean(Sd)
	ESL	0.0089	0.7942(0.0669)	−0.3006(0.0668)
100	SWCQR	0.0110	0.7856(0.0702)	−0.3026(0.0770)
	SWLAD	0.0128	0.7867(0.0794)	−0.3021(0.0834)
	LS	0.0904	0.7787(0.1012)	−0.0308(0.0862)
	ESL	0.0034	0.8042(0.0424)	−0.3040(0.0409)
200	SWCQR	0.0077	0.7920(0.0612)	−0.3051(0.0629)
	SWLAD	0.0089	0.7945(0.0645)	−0.3043(0.0694)
	LS	0.0943	0.9075(0.0755)	−0.0324(0.0744)
	ESL	0.0018	0.8004(0.0307)	−0.2987(0.0302)
500	SWCQR	0.0050	0.8038(0.0479)	−0.3147(0.0507)
	SWLAD	0.0062	0.8008(0.0554)	−0.3128(0.0647)
	LS	0.0296	0.8699(0.0853)	−0.1989(0.0860)

We compare our proposed method (ESL) with SWCQR proposed by [10], SWLAD introduced by [8], and the least squares (LS) method. For the ESL estimator, we choose the tuning parameter λ based on another 500 independent samples for each simulation. In order to evaluate the finite-sample performance, we calculate the mean, standard deviation (Sd) as well as the mean-squared errors (MSE) for the estimators of ϕ over 500 simulations, respectively.

The results are summarized in Table 1–6. From Table 1–6, we can find that the Sd and MSE of ESL, SWCQR, and SWLAD in all three settings are decreased as the sample size n increases. Meanwhile, according to Table 1 and Table 4, the Sd and MSE of ESL, SWCQR, and SWLAD are very close and are slightly higher than those of LS when the error has a standard normal distribution. However, when the error follows a heavy-tailed distribution and there are outliers in the dataset, the Sd and MSE of our proposed method are smaller than those of SWCQR and SWLAD method. This illustrates that our proposed method is robust to outliers in the dataset, and is more efficient than SWCQR

Table 3. The mean, standard deviation (Sd), and MSE for the estimators of ϕ under Cauchy error

n	Estimator	MSE	$\phi_1 = 0.8$ Mean(Sd)	$\phi_2 = -0.3$ Mean(Sd)
	ESL	0.0037	0.7967(0.0403)	−0.2989(0.0459)
100	SWCQR	0.0097	0.7935(0.0724)	−0.2957(0.0673)
	SWLAD	0.0142	0.7884(0.0875)	−0.2969(0.0810)
	LS	0.1046	0.8193(0.2187)	−0.1565(0.1915)
	ESL	0.0006	0.7990(0.0184)	−0.2973(0.0162)
200	SWCQR	0.0061	0.8011(0.0945)	−0.2831(0.1126)
	SWLAD	0.0091	0.7960(0.0608)	−0.2819(0.0716)
	LS	0.0916	0.8526(0.2145)	−0.1576(0.1522)
	ESL	0.0001	0.7988(0.0079)	−0.2996(0.0067)
500	SWCQR	0.0035	0.8152(0.0354)	−0.2772(0.0399)
	SWLAD	0.0042	0.8151(0.0405)	−0.2824(0.0456)
	LS	0.0296	0.8699(0.0853)	−0.1989(0.0860)

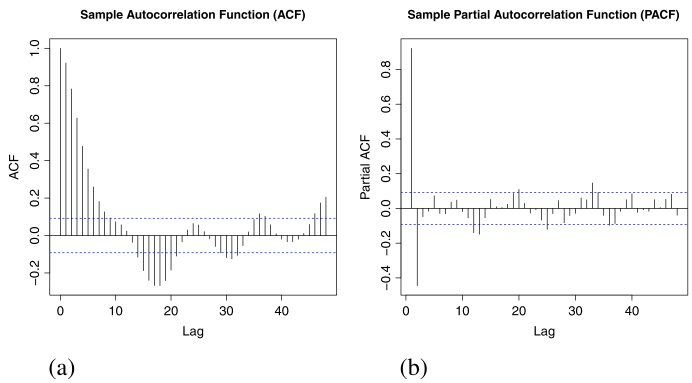

Figure 2. ACF (a) and PACF (b) of the Recruitment series.

when the error is heavy tailed and there are outliers in the dataset.

3.1 Real data application

As an illustration, we apply the proposed methodology to model the Recruitment series [9] (number of new fish). There are 453 months of observed recruitment ranging over the years 1950–1987. The autocorrelation function (ACF) and the partial autocorrelation function (PACF) given in Figure 2 are consistent with the behavior of an AR(2).

We first use the OLS method to analyze the dataset. The estimated results are given in Table 7. Furthermore, a normal Q-Q plot of regression residuals is shown in Figure 3. From Figure 3, we can find that Q-Q plot lacks linearity. Therefore, the error indicates non-normality. We use the Shapiro-Wilk normality test, and obtain that the p-value is 2.843×10^{-7}. This also illustrates that the error is a heavy-tailed distribution.

Next, we apply the SWCQR method and ESL method to tackle this dataset. For the ESL method, we use the proposed method to select the tuning parameter λ, and obtain $\lambda = 0.1$. The results are shown in Table 7. The standard er-

Table 4. The mean, standard deviation (Sd), and MSE for the estimators of ϕ under normal error

			$\phi_1 = 0.5$	$\phi_2 = 0$	$\phi_3 = -0.7$
n	Estimator	MSE	Mean(Sd)	Mean(Sd)	Mean(Sd)
	ESL	0.0334	0.4751(0.0936)	−0.0074(0.1187)	−0.6710(0.0967)
100	SWCQR	0.0306	0.4894(0.0945)	−0.0179(0.1126)	−0.6801(0.0921)
	SWLAD	0.0406	0.4817(0.1135)	−0.0121(0.1298)	−0.6828(0.1025)
	LS	0.0283	0.4912(0.0903)	−0.0158(0.1081)	−0.6701(0.0866)
	ESL	0.0134	0.5060(0.0607)	−0.0112(0.0682)	−0.6766(0.0666)
200	SWCQR	0.0139	0.4994(0.0579)	−0.0073(0.0747)	−0.6850(0.0697)
	SWLAD	0.0206	0.5028(0.0737)	−0.0113(0.0907)	−0.6837(0.0823)
	LS	0.0121	0.5003(0.0538)	−0.0076(0.0707)	−0.6778(0.0619)
	ESL	0.0036	0.4972(0.0284)	−0.0003(0.0378)	−0.6928(0.0367)
500	SWCQR	0.0036	0.5003(0.0299)	−0.0054(0.0367)	−0.6929(0.0363)
	SWLAD	0.0063	0.5007(0.0413)	−0.0060(0.0485)	−0.6922(0.0468)
	LS	0.0032	0.4993(0.0275)	−0.0036(0.0346)	−0.6904(0.0337)

Table 5. The mean, standard deviation (Sd), and MSE for the estimators of ϕ under t_2 error

			$\phi_1 = 0.5$	$\phi_2 = 0$	$\phi_3 = -0.7$
n	Estimator	MSE	Mean(Sd)	Mean(Sd)	Mean(Sd)
	ESL	0.0068	0.5103(0.0407)	−0.0008(0.0480)	−0.6932(0.0531)
100	SWCQR	0.0138	0.5026(0.0675)	−0.0106(0.0736)	−0.6855(0.0604)
	SWLAD	0.0164	0.5015(0.0705)	−0.0101(0.0838)	−0.6895(0.0662)
	LS	0.5183	0.7469(0.0943)	−0.0632(0.1165)	−0.0475(0.0739)
	ESL	0.0035	0.5011(0.0280)	−0.0045(0.0389)	−0.6928(0.0340)
200	SWCQR	0.0051	0.4978(0.0386)	−0.0017(0.0455)	−0.6960(0.0400)
	SWLAD	0.0059	0.4954(0.0410)	0.0042(0.0484)	−0.6985(0.0434)
	LS	0.7354	0.9778(0.0882)	−0.1377(0.1135)	−0.0218(0.0883)
	ESL	0.0010	0.4996(0.0163)	0.0001(0.0204)	−0.6988(0.0182)
500	SWCQR	0.0031	0.5007(0.0288)	−0.0031(0.0361)	−0.6964(0.0318)
	SWLAD	0.0043	0.5034(0.0341)	−0.0067(0.0427)	−0.6931(0.0356)
	LS	1.0499	1.1383(0.0939)	−0.2783(0.1077)	0.0324(0.0922)

Table 6. The mean, standard deviation (Sd), and MSE for the estimators of ϕ under Cauchy error

			$\phi_1 = 0.5$	$\phi_2 = 0$	$\phi_3 = -0.7$
n	Estimator	MSE	Mean(Sd)	Mean(Sd)	Mean(Sd)
	ESL	0.0012	0.5007(0.0176)	0.0002(0.0235)	−0.6990(0.0196)
100	SWCQR	0.0087	0.5175(0.0469)	−0.0199(0.0569)	−0.6817(0.0479)
	SWLAD	0.0198	0.5289(0.0699)	−0.0337(0.0825)	−0.6681(0.0730)
	LS	0.2868	0.7154(0.1456)	−0.1013(0.1383)	−0.3329(0.2366)
	ESL	0.0005	0.4976(0.0112)	0.0021(0.0132)	−0.7017(0.0120)
200	SWCQR	0.0032	0.5114(0.0305)	−0.0112(0.0321)	−0.6874(0.0300)
	SWLAD	0.0148	0.5265(0.0622)	−0.0286(0.0731)	−0.6739(0.0599)
	LS	0.4633	0.8421(0.2245)	−0.1910(0.1951)	−0.3017(0.2532)
	ESL	0.0001	0.5001(0.0046)	0.0002(0.0052)	−0.6999(0.0039)
500	SWCQR	0.0006	0.5059(0.0121)	−0.0072(0.0138)	−0.6915(0.0096)
	SWLAD	0.0025	0.5209(0.0199)	−0.0201(0.0214)	−0.6786(0.0183)
	LS	0.3239	0.7611(0.2324)	−0.1894(0.1970)	−0.4299(0.2358)

rors based on 100 replacement bootstrap samples are given in their corresponding parentheses. From Table 7, we can find that the estimated results of ESL and SWCQR are very different. We utilize the resulting models to make a forecast, and calculate the following relative mean square error (RMSE) to evaluate the predict performance of ESL and SWCQR,

$$RMSE(\hat{\boldsymbol{\phi}}) = \frac{(\mathbf{y} - \hat{\mathbf{y}}_{\hat{\boldsymbol{\phi}}})^T(\mathbf{y} - \hat{\mathbf{y}}_{\hat{\boldsymbol{\phi}}})}{(\mathbf{y} - \hat{\mathbf{y}}_{\hat{\boldsymbol{\phi}}_{LS}})^T(\mathbf{y} - \hat{\mathbf{y}}_{\hat{\boldsymbol{\phi}}_{LS}})}.$$

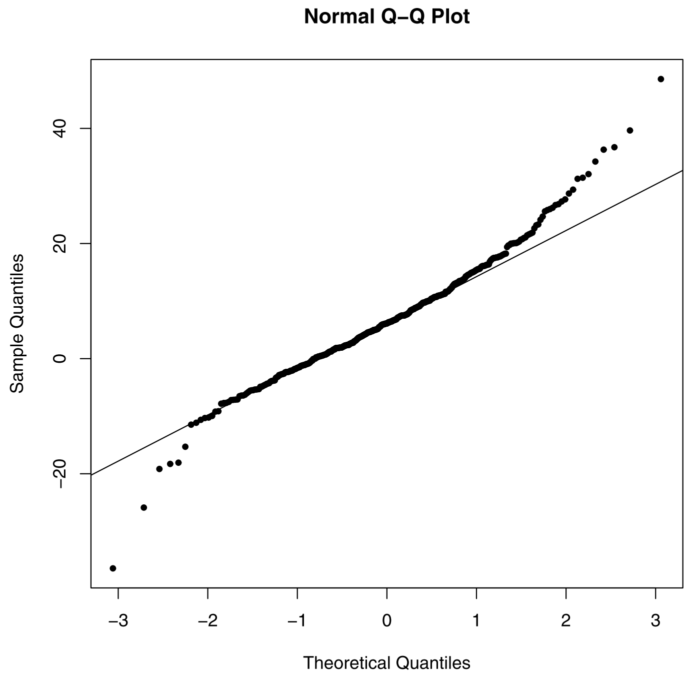

Figure 3. Q-Q plot of regression residuals.

Table 7. Estimated autoregressive parameters for the Recruitment series

Estimator	ϕ_1	ϕ_2
OLS	1.3541(0.0515)	−0.4632(0.0658)
SWCQR	1.3259(0.0466)	−0.4295(0.0526)
ESL	1.8304(0.0100)	−0.8843(0.0394)

By simple calculation, we yield that $RMSE(\hat{\phi}^{ESL}) = 0.8843$ and $RMSE(\hat{\phi}^{SWCQR}) = 0.9670$. Therefore, we suggest the ELS method to tackle with time series data with a heavy-tailed distribution in practice.

4. DISCUSSION

In this paper, we proposed an exponential-squared estimator for the unknown autoregressive parameter in the AR model with heavy-tailed errors. The advantages of the proposed method were illustrated through numerical simulations and a real data analysis. According to simulation studies, ESL method has a smaller MSE than the SWCQR method when the error followed a heavy-tailed distribution. It is very interesting to investigate the asymptotic properties of proposed estimator and the effect of choice of tuning parameter on outcomes, which leaves for further work.

APPENDIX

Proof of Theorem 2.1. Denote $\mathbf{X}_i = (Y_{i-1}, \cdots, Y_{i-p})^T$, and

$$L_n(\boldsymbol{\phi}) = \sum_{i=p+1}^{n+p} \exp\left\{-(Y_i - \boldsymbol{\phi}^T \mathbf{X}_i)^2/\lambda\right\}.$$

By the Taylor expansion, we have

$$\begin{aligned} D_n(\mathbf{u}) &= L_n(\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u}) - L_n(\boldsymbol{\phi}_0) \\ &= n^{-1/2} L_n'(\boldsymbol{\phi}_0)\mathbf{u} - \frac{1}{2}\mathbf{u}^T\left[-I(\boldsymbol{\phi}_0)\right]\mathbf{u}\{1 + o_p(1)\}, \end{aligned} \tag{A.1}$$

where $\mathbf{u}$ is p-dimensional vector such that $\|\mathbf{u}\| = C$, and

$$L_n'(\boldsymbol{\phi}_0) = \sum_{i=p+1}^{n+p}\left[\exp\left\{-\epsilon_i^2/\lambda\right\}\frac{2\epsilon_i}{\lambda}\mathbf{X}_i\right],$$

$$I(\boldsymbol{\phi}_0) = \frac{2}{\lambda}E\left[\mathbf{X}_i\mathbf{X}_i^T e^{-\epsilon_i^2/\lambda}\left(\frac{2\epsilon_i^2}{\lambda} - 1\right)\right].$$

According to the condition (C2), $I(\boldsymbol{\phi}_0)$ is negative definite. By the classical central limit theorem, $n^{-1/2}L_n'(\boldsymbol{\phi}_0) = O_p(1)$. Therefore, for the last equality of Equation (A.1), the second term dominates the first term uniformly in $||\mathbf{u}|| = C$ by choosing a sufficiently large C. Thus, for any given $\epsilon > 0$, we have

$$P\left\{\sup_{||\mathbf{u}||=C} L_n(\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u}) < L_n(\boldsymbol{\phi}_0)\right\} \geq 1 - \epsilon. \tag{A.2}$$

Equation (A.2) implies that with probability at least $1 - \epsilon$ that there exists a local maximum $\hat{\boldsymbol{\phi}}_n$ in the ball $\{\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u} : ||\mathbf{u}|| \leq C\}$. This completes the proof of Theorem 2.1.

ACKNOWLEDGEMENTS

Jiang's research is partially supported by the National Natural Science Foundation of China (No. 11301221), and is supported by the Fundamental Research Funds for the Central Universities (No. 11615455).

Received 14 August 2014

REFERENCES

[1] Castillo, E. (1988). Extreme value theory in engineering. *Statistical Modeling and Decision Science*. Academic Press, New York. MR0958518

[2] Chen, Z., Li, R., and Wu, Y. (2012). Weighted quantile regression for ar model with infinite variance errors. *Journal of Nonparametric Statistics* **24**, 3, 715–731. MR2968898

[3] Davis, R. A., Knight, K., and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Processes and their Applications* **40**, 1, 145–180. MR1145464

[4] Friedman, J., Hastie, T., and Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The Annals of Statistics* **28**, 2, 337–407. MR1790002

[5] Hill, J. B. (2013). Least tail-trimmed squares for infinite variance autoregressions. *Journal of Time Series Analysis* **34**, 2, 168–186. MR3028364

[6] Koedijk, K. G., Schafgans, M., and de Vries, C. G. (1990). The tail index of exchange rate returns. *Journal of international economics* **29**, 1, 93–108.

[7] Li, J., Liang, W., He, S., and Wu, X. (2010). Empirical likelihood for the smoothed lad estimator in infinite variance autoregressive models. *Statistics & Probability Letters* **80**, 17, 1420–1430. MR2669783

[8] Ling, S. (2005). Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **67**, 3, 381–393. MR2155344

[9] Shumway, R. H. and Stoffer, D. S. (2011). *Time Series Analysis and Its Applications: with R Examples*. Springer. MR2721825

[10] Tang, L., Zhou, Z., and Wu, C. (2012). Efficient estimation and variable selection for infinite variance autoregressive models. *Journal of Applied Mathematics and Computing* **40**, 1, 399–413. MR2965339

[11] Wang, X., Jiang, Y., Huang, M., and Zhang, H. (2013). Robust variable selection with exponential squared loss. *Journal of the American Statistical Association* **108**, 502, 632–643. MR3174647

[12] Willinger, W. and Paxson, V. (1997). Discussion: Heavy tail modeling and teletraffic data. *The Annals of Statistics* **25**, 5, 1856–1866. MR1474072

Yunlu Jiang
Department of Statistics
College of Economics, Jinan University
Guangzhou, 510632
China
E-mail address: tjiangyl@jnu.edu.cn