The Korean Communications
in Statistics Vol. 7, No. 3, 2000
pp. 837-843

A Note on Bootstrapping M-estimators in TAR Models.[1]

Sahmyeong Kim[2]

Abstract

Kreiss and Franke(1992) and Allen and Datta(1999) proposed bootstrapping the M-estimators in ARMA models. In this paper, we introduce the robust estimating function and investigate the bootstrap approximations of the M-estimators which are solutions of the estimating equations in TAR models. A number of simulation results are presented to estimate the sampling distribution of the M-estimators, and asymptotic validity of the bootstrap for the M-estimators is established.

Keywords : TAR models; Bootstrap; Estimating Function; M-estimators; Asymptotic validity

1.Introduction

The Threshold Autoregressive (TAR) models are defined as

$$X_t = \sum_{i=1}^{k} \sum_{j=1}^{p} \theta_{ij} X_{t-j} I_{ij} + \varepsilon_t, \quad t=1,\ 2,\ \ldots\ldots \tag{1.1}$$

where I_{ij} is the indicator function of the event ($X_{t-j}\ \varepsilon\ R_i$), R_i, I=1,...,k are k disjoint regions with R = $\bigcup_{i=1}^{k} R_i$. and { ε_t} are i.i.d random variables with zero mean and variance σ^2. In general, it is not clear whether (1.1) admits a stationary solution. However in various special cases, sufficient conditions for the same can be obtained. As for example, Petrucelli and Woolford(1984) constructed the following special case of the TAR(1) model;

$$X_t = \theta_1 X_{t-1}^{+} + \theta_2 X_{t-1}^{-} + \varepsilon_t, t=1,2,\ldots \tag{1.2}$$

where { ε_t} is a sequence of i.i.d random variables which has strictly positive density on the real line, has zero mean, finite variance σ^2, $X_{t-1}^{+} = X_{t-1} I_{(X_{t-1} > 0)}$ and $X_{t-1}^{-} = X_{t-1} I_{(X_{t-1} \le 0)}$. They showed that the necessary and sufficient conditions for the existence of a stationary and ergodic solution of the process in (1.2) are

$$\theta_1 < 1,\ \ \theta_2 < 1 \ \text{and} \ \ \theta_1\theta_2 < 1. \tag{1.3}$$

1) This work was partially supported by a research fund of Chung-Ang University,2000
2) Assistant Professor, Department of Applied Statistics, Chung-Ang University, Seoul, 156-756, Korea. E-mail : sahm@cau.ac.kr

Consider the following estimating equation.

$$\sum_{t=1}^{n} \psi\, [\, X_t - H(\, \underline{X}_{t-1};\theta)]\frac{d(H(\, \underline{X}_{t-1};\theta))}{d\theta} = 0 \qquad (1.4)$$

where ψ is a specified function. Typically, ψ is a bounded function in the context of robust inference. Kreiss and Franke(1992) and Allen and Datta(1999) used the estimating equation for the bootstrap approximations. But, Basawa et. al. (1985) pointed out that the estimating equation in (1.4) is not fully robust because the part $\frac{d(H(\underline{X}_{t-1};\theta))}{d\theta}$ is not robustified. For this reason, we need to consider a different type of robust estimating equation which was suggested by Basawa et. al.(1985).

In chapter 2, we introduce a different type of robust estimating function and we establish the asymptotic validity of the estimators which are consistent solutions for the new estimating equation in the TAR models. Percentiles for the studentized data and the bootstrap data will be presented in chapter 3 to compare empirically the validity of the boostrap estimators in the TAR models.

2. Bootstrap Approximations

2.1 Robust Estimating Functions for TAR Models

In this section, we will investigate the asymptotic properties of robust estimators of parameters in TAR models by using maximum likelihood type estimators (M-estimators). For simplicity, we use the TAR(1) model from now on.

Consider the robust estimating function

$$\begin{aligned} S_n(\theta) &= \sum_{t=1}^{n} \psi_1(\varepsilon_t(\theta))\,\psi_2\!\left(\frac{d\mu_t(\,\underline{X}_{t-1};\theta)}{d\theta}\right) \\ &= \sum_{t=1}^{n} \begin{pmatrix} g_1(X_t;\theta) \\ \vdots \\ g_p(X_t;\theta) \end{pmatrix} \end{aligned} \qquad (2.1)$$

where

$$\mu_t(\,\underline{X}_{t-1};\theta) = E(X_t|F_{t-1}),\;\; F_{t-1} = \sigma(X_{t-1},\ldots,X_{t-p}\,),$$

$$\varepsilon_t(\theta) = X_t - (\,\theta_1 X_{t-1}^{+} - \theta_2 X_{t-1}^{-}).$$

Note that

$$\psi_2\!\left(\frac{d\mu_t(\,\underline{X}_{t-1};\theta)}{d\theta}\right) = \begin{pmatrix} \psi_2\!\left(\frac{d\mu_t(\,\underline{X}_{t-1};\theta)}{d\theta_1}\right) \\ \vdots \\ \psi_2\!\left(\frac{d\mu_t(\,\underline{X}_{t-1};\theta)}{d\theta_p}\right) \end{pmatrix}.$$

Next, we will establish the asymptotic normality of the estimator satisfying the estimating equation $S_n^*(\theta) = 0$ in (2.1). Consider the regularity conditions.

1. $\{X_t, t \geq 1\}$ is stationary and ergodic with $EX_t^2 < \infty$.
2. ψ_1 and ψ_2 are bounded and differentiable in θ.
3. $E(\psi_1(\varepsilon_1)) = 0$.
4. For all θ^* such that $\theta^* \xrightarrow{p} \theta$, $n^{-1}[\sum_{t=1}^{n} \frac{d}{d\theta} g(X_t;\theta) - \frac{d}{d\theta} g(X_t,\theta)|\theta=\theta^*] \xrightarrow{p} 0$.
5. The p×p matrix A= $E_\theta(\frac{dg_i}{d\theta_j})$ exists for all i, j = 1, ..., p and A is a non-singular matrix.
6 The p×p matrix B= $E_\theta(g_i g_j)$ exists for all i, j = 1,....,p and B is a positive definite matrix.

The following theorem establishes the asymptotic normality of the estimator which satisfies the estimating equation $S_n(\theta) = 0$ in (2.1).

Theorem 2.1 Suppose that $\hat{\theta}_n$ is a $\sqrt{n}$-consistent solution of $S_n(\theta) = 0$ in (2.1). Then under the conditions 1-6, we have

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, V) \tag{2.2}$$

where

$$V = (A^T B^{-1} A)^{-1}.$$

proof: This proof can be done as in Kim(1998)

2.2 Asymptotic validity

The bootstrap samples from an TAR(1) model are made like Kreiss and Franke(1992) and Allen and Datta(1999). In order to make bootstrap samples, we need to calculate a preliminary parameter estimator $\hat{\theta}_M = (\hat{\theta}_{1M}, \hat{\theta}_{2M})$ which is the solution of $S_n(\theta) = 0$ in (2.1) by using the data $X = (X_1, X_2, \ldots, X_n)^T$. We also need to calculate the residuals which is to estimate the errors.

First we can construct the residuals through

$$e_t = \sum_{t=1}^{n}(X_t - \hat{\theta}_{1M} X_{t-1}^+ - \hat{\theta}_{2M} X_{t-1}^-),$$

where $\hat{\theta} = (\hat{\theta}_{1M}, \hat{\theta}_{2M})^T$ is the solution of $S_n(\theta) = 0$ in (2.1).

Based on the residuals, we can define the following centered empirical distribution

$$\hat{F}_n(x) = \frac{1}{n}\sum_{t=1}^{n} I_{[e_t-\bar{e}]}(x), \quad x \in R,$$

where $\bar{e} = \frac{1}{n}\sum_{t=1}^{n} e_t$.

Next we resample i.i.d variables e_t^*, $1 \le t \le n$, from $\hat{F}_n$ and make the bootstrapped data

$$X^* = (X^*_1, \ldots, X^*_n)^T, \; X_t^* = \hat{\theta}_{1M} X_{t-1}^{*+} + \hat{\theta}_{2M} X_{t-1}^{*-}, \quad t = 1, \ldots, n$$

with the initial value $X_0^* = 0$, $\varepsilon_0^* = 0$.

Next, we need to establish the robust bootstrap estimating function

$$S_n^* = \sum_{t=1}^{n} \psi_1(e_t^*)\psi_2(\underline{X}_{t-1}{}^*), \tag{2.3}$$

where $\underline{X}_{t-1}{}^* = (X_{t-1}{}^{*+}, X_{t-1}{}^{*-})$, ψ_1 and ψ_2 are bounded and differentiable functions.

Assume that $\hat{\theta}^*$ is the M-estimator which satisfies the following

$$S_n^*(\hat{\theta}^*) = o_{p*(1)} \quad \text{a.s.} \tag{2.4}$$

Allen and Datta(1999) assume the following

$$n^{1/2}(\hat{\theta}^* - \hat{\theta}_M) = O_{p*(1)} \quad \text{a.s.} \tag{2.5}$$

Theorem 3.1: Under the conditions in Theorem 2.1 and (2.4) and (2.5),
we have $\sqrt{n}(\hat{\theta}^* - \hat{\theta}) \xrightarrow{d} N(0,V)$, almost surely as $n \to \infty$,
where $V = (A^T B^{-1} A)^{-1}$, $A = E[\psi_1'(\varepsilon_t)\ \psi_2\ \psi_2^T]$, $B = E[\psi_1^2(\varepsilon_t)\ \psi_2\ \psi_2^T]$.

Remarks: Allen and Datta(1999) assume $E|X_t|^3 < \infty$ whereas we assume $E|X_t|^2 < \infty$. This is due to the robustifying bootstrap estimating function in (2.1). The existence of strongly consistent estimator of θ is also assumed by Kreiss and Franke(1992).

proof: This proof can be done by using theorem 3.1 of Allen and Datta(1999) with slight modifications. So, we omit the proof.

3. Simulation Results For Bootstrap

We present a simulation study to compare the bootstrap distributions with the true sampling distributions of the M-estimators with true sampling distributions of the M-estimators in the TAR(1) model. We have three choices for the error distribution, standard normal, standard double exponential and mixture of two normals, i.e, 0.9* N(0,1) + 0.1*N(0,5) are selected. Sample size of n=200 is used. We compute the percentiles of the studentized M-estimator

$T=\sqrt{n}(\hat{\theta}_M-\theta)/\hat{\sigma}^2$ where $\hat{\sigma}^2=\hat{V}(\hat{\theta}_M)$ and V is in (2.2).

For example, In case of TAR(1) model,

$$\hat{\sigma}^2=\frac{\frac{1}{n}\sum_{t=1}^{n}\psi_1^2(e_t)\cdot\frac{1}{n}\sum_{t=1}^{n}\psi_2^2(X_{t-1})}{\frac{1}{n}\sum_{t=1}^{n}[\psi_1'(e_t)]^2\cdot\frac{1}{n}\sum_{t=1}^{n}[X_{t-1}\psi_2(X_{t-1})]^2} \tag{3.1}$$

where $e_t = X_t-\hat{\theta}_{1M}X_{t-1}^{+}-\hat{\theta}_{2M}X_{t-1}^{-}$ and $\hat{\theta}_M=(\hat{\theta}_{1M},\hat{\theta}_{2M})^T$ is a solution of $S_n(\theta)=0$ in (2.1). The number of Monte Carlo simulation is 2000 and we get 2000 values of T. We repeated the whole procedure 30 times and we computed the average and the standard deviation of each percentile. The average percentiles of T are reported in the first rows of the following tables. Standard deviations of the percentiles are in parentheses. For the TAR(1) model the parameter value is θ = (0.3, -0.3).

Next, we compute the percentiles of the bootstrap approximations T^* for a give sample where $T^*=\sqrt{n}(\hat{\theta}_n-\theta^*)/s^*$ and

$$s^{*2}=\frac{\frac{1}{n}\sum_{t=1}^{n}\psi_1^2(e_t^*)\cdot\frac{1}{n}\sum_{t=1}^{n}\psi_2^2(X_{t-1}^*)}{\frac{1}{n}\sum_{t=1}^{n}[\psi_1'(e_t^*)]^2\cdot\frac{1}{n}\sum_{t=1}^{n}[X_{t-1}^*\psi_2(X_{t-1}^*)]^2}$$

Sample size of n=200 was used for the bootstrap procedure. Under the three distributions, the bootstrap approximations work reasonably. They work quite well under the heavy-tailed error distributions. In case of the contaminated normal and the double exponential distributions, the bootstrap estimator has a little high variability. We expect that the situation would be improved for larger sample sizes.

Table 3.1. Percentiles of the studentized M-estimator of a TAR(1) model with Standard Normal error.

	Percentile	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
θ_1=0.3	True	-1.666 (0.048)	-1.275 (0.037)	-0.799 (0.032)	-0.456 (0.028)	0.100 (0.028)	0.634 (0.027)	0.956 (0.033)	1.392 (0.039)	1.746 (0.041)
	Bootstrap	-1.681 (0.068)	-1.287 (0.067)	-0.814 (0.065)	-0.477 (0.063)	0.074 (0.068)	0.615 (0.064)	0.940 (0.069)	1.378 (0.075)	1.731 (0.088)
θ_2=-.3	True	-1.760 (0.037)	-1.412 (0.030)	-0.983 (0.032)	-0.668 (0.033)	-0.127 (0.030)	0.434 (0.033)	0.770 (0.036)	1.234 (0.032)	1.622 (0.046)
	Bootstrap	-1.761 (0.065)	-1.428 (0.056)	-0.995 (0.058)	-0.673 (0.061)	-0.127 (0.067)	0.424 (0.066)	0.764 (0.067)	1.241 (0.072)	1.625 (0.072)

Table 3.2. Percentiles of the studentized M-estimator of a TAR(1) model with Contaminated Normal error for σ^2=5.

	Percentile	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
θ_1=0.3	True	-1.666 (0.0501)	-1.278 (0.048)	-0.811 (0.036)	-0.479 (0.030)	0.075 (0.027)	0.625 (0.022)	0.949 (0.026)	1.387 (0.036)	1.735 (0.047)
	Bootstrap	-1.604 (0.198)	-1.193 (0.185)	-0.708 (0.174)	-0.356 (0.173)	0.203 (0.176)	0.744 (0.168)	1.061 (0.173)	1.489 (0.168)	1.838 (0.168)
θ_2=-.3	True	-1.772 (0.040)	-1.417 (0.032)	-0.984 (0.027)	-0.665 (0.025)	-0.124 (0.027)	0.424 (0.036)	0.766 (0.031)	1.240 (0.038)	1.642 (0.046)
	Bootstrap	-1.807 (0.107)	-1.473 (0.121)	-0.052 (0.123)	-0.738 (0.122)	-0.199 (0.135)	0.363 (0.146)	0.711 (0.158)	1.198 (0.159)	1.604 (0.167)

Table 3.3. Percentiles of the studentized M-estimator of a TAR(1) model with with the standard Double Exponential error.

	Percentile	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
θ_1=0.3	True	-1.327 (0.063)	-0.885 (0.045)	-0.381 (0.040)	-0.032 (0.038)	0.542 (0.028)	1.067 (0.028)	1.374 (0.033)	1.786 (0.040)	2.107 (0.042)
	Bootstrap	-1.438 (0.232)	-0.999 (0.214)	-0.469 (0.203)	-0.106 (0.199)	0.457 (0.192)	0.994 (0.187)	1.302 (0.183)	1.710 (0.193)	2.048 (0.199)
θ_2=-.3	True	-1.943 (0.033)	-1.644 (0.030)	-1.258 (0.023)	-0.962 (0.026)	-0.451 (0.028)	0.111 (0.035)	0.460 (0.038)	0.971 (0.040)	1.401 (0.056)
	Bootstrap	-1.908 (0.123)	-1.610 (0.123)	-1.222 (0.131)	-0.924 (0.133)	-0.394 (0.143)	0.165 (0.149)	0.533 (0.160)	1.049 (0.179)	1.496 (0.193)

REFERENCES

[1] Allen, M. and Datta,S.(1999). A note on bootstrapping M-estimators in ARMA models. *Journal of Time Series Analysis* 20, 365-379.

[2] Basawa, I.V., Huggins, R.M. and Staudte, R.G.(1985). Robust tests for time series

with an application to first-order autoregressive processes. *Biometrika* 72, 559-571.

[3] Godambe, V.P.(1985). The foundations of finite sample estimation in stochastic processes. *Biometrika* 72, 419-428.

[4] Kim, S.(1998). Inference for nonlinear time series models via estimating functions. Ph. D. dissertation, Univ. of Georgia.

[5] Kreiss, J.P. and Franke,J.(1992). Bootstrapping stationary autoregressive moving average models. *Journal of Time Series Analysis* 13, 297-317.

[6] Petrucelli, J.D. and Woolford,S.W.(1984). A threshold AR(1) model. *Journal of Applied Probability* 21, 270-286.