IJST (2014) 38A3: 239-241
Iranian Journal of Science & Technology

http://ijsts.shirazu.ac.ir

On L^2 convergence of the maximum weighted pairwise likelihood estimators in the AR(1) models

M. R. Kazemi[1]* and A. R. Nematollahi[2]

[1]*Department of Statistics, College of Sciences, Fasa University, Fasa, Iran*
[2]*Department of Statistics, College of Sciences, Shiraz University, Shiraz, Iran*
E-mail: kazemi@fasau.ac.ir & nematollahi@susc.ac.ir

Abstract

Recently, the strong consistency and asymptotic distribution for the maximum consecutive pairwise likelihood estimators (MCPLE) have been established in the linear time series models. In this paper, the weak convergence of the maximum weighted pairwise likelihood estimator (MWPLE) of the parameters of the AR(1) models is established by using the concept of L^2 convergence (convergence in mean square).

***Keywords**:* Pairwise likelihood; composite likelihood; autoregressive process; L^2 convergence

1. Introduction

The pairwise likelihood (PL) is a special case of the composite likelihood proposed in Lindsay (1988) as a pseudo-likelihood. In PL, the pseudo-likelihood is constructed by the product of the bivariate likelihood of all possible pairs of observations. Detailed accounts of PL can be found in Cox and Reid (2004). For an excellent review on the composite likelihood methods, see Varin (2008). A general recent discussion on theoretical aspects and possible applied contexts are also considered in Varin et al. (2011).

Recently, Davis and Yau (2011) have established the consistency and asymptotic distribution for the MCPLE in the linear time series models. In particular, they showed that the asymptotic relative efficiency of the MCPLE to the MLE is one for all values of the AR(1) parameter. Formally, let X_t follows the invertible stationary AR(1) models as,

$$X_t = \phi X_{t-1} + Z_t,\ \ |\phi| < 1, \tag{1}$$

where $Z_t \sim N(0, \sigma^2)$. Let $\mathbf{X}_i = (X_i, X_{i+1})^T$. It is easy to show that

$$\mathbf{X}_i \sim N_2(\mathbf{0}\ , \mathbf{\Sigma}),$$

where

$$\mathbf{\Sigma} = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix}.$$

Now, consider the WPL, $L_{wpl}(\mathbf{x}; \phi, \sigma^2)$, given by

$$L_{wpl}(\mathbf{x}; \phi, \sigma^2) = \prod_i \prod_j \left(f_{(X_i, X_j)}(x_i, x_j; \phi, \sigma^2)\right)^{\pi_{ij}}, \tag{2}$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ is the vector of the observations, and

$$\pi_{ij} = \begin{cases} 1, & j = i+1 \\ 0, & \text{otherwise}, \end{cases}$$

are the corresponding weights according to the autoregressive property of the observations. The weighted pairwise log-likelihood is then given by

$$l_{wpl}(\mathbf{x}; \phi, \sigma^2) \coloneqq \ln\left(L_{wpl}(\mathbf{x}; \phi, \sigma^2)\right)$$
$$\propto -\frac{1}{2}\sum_{i=1}^{n-1} \{\ln\mathbf{\Sigma} + \mathbf{x}_i^T \mathbf{\Sigma}^{-1}\mathbf{x}_i\}$$
$$= -\tfrac{1}{2}[(n-1)\ln(1-\phi^2) - 2(n-1)\ln(\sigma^2) - \sigma^{-2}\textstyle\sum_{i=1}^{n-1} (x_i^2 + x_{i+1}^2) + 2\phi\sigma^{-2}\sum_{i=1}^{n-1} x_i x_{i+1}].$$

So the weighted pairwise score function is

$$\mathbf{S}_{wpl}(\mathbf{x}; \phi, \sigma^2) = \left(\frac{\partial l_{wpl}(\mathbf{x}; \phi, \sigma^2)}{\partial \phi}, \frac{\partial l_{wpl}(\mathbf{x}; \phi, \sigma^2)}{\partial \sigma^2}\right)$$
$$= \left(-\frac{(n-1)\phi}{1-\phi^2} + \frac{\sum_{i=1}^{n-1} x_i x_{i+1}}{\sigma^2}, -\frac{n-1}{\sigma^2} + \frac{\frac{1}{2}\sum_{i=1}^{n-1} (x_i^2 + x_{i+1}^2) - \phi\sum_{i=1}^{n-1} x_i x_{i+1}}{\sigma^4}\right).$$

Set $\mathbf{S}_{wpl}(\mathbf{X}; \phi, \sigma^2)$ equal to zero and solve the equations with respect to ϕ, σ^2, then the MWPLE of ϕ and σ^2 are given by

$$\tilde{\phi} = \frac{2\sum_{i=1}^{n-1} X_i X_{i+1}}{\sum_{i=1}^{n-1} (X_i^2 + X_{i+1}^2)},$$

*Corresponding author
Received: 20 June 2013 / Accepted: 16 April 2014

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n-1} (X_i^2 + X_{i+1}^2)}{2(n-1)} - \frac{\tilde{\phi}\sum_{i=1}^{n-1} X_i X_{i+1}}{n-1},$$

respectively. These estimators are also derived by Davis and Yau (2011). In this paper we study the limiting behavior of these estimators by using the limiting behavior of the statistics

$$T_n = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i X_{i+1},$$
$$T_n' = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i^2,$$
$$T_n'' = \frac{1}{n-1}\sum_{i=1}^{n-1} X_{i+1}^2. \quad (3)$$

However, Davis and Yau (2011) have established the strong convergence of these estimators and the strong convergence always implies the weak convergence, but in this paper we use a simple method to establish the L^2 convergence of these estimators which in turn establishes their weak convergence of them. The weak convergence of the WPLE$\tilde{\phi}$ is then an immediate consequence of the Slustky's theorem.

2.Main Result

The following lemmas are necessary to find the limiting distribution of T_n, T_n' and T_n''.

Lemma 2.1. Suppose that X_t follows the first-order autoregressive process defined by (1), and let T_n, T_n' and T_n'' be as in (3). Then

$$E(T_n) = \frac{\phi\sigma^2}{1-\phi^2},\ E(T_n') = \frac{\sigma^2}{1-\phi^2} \text{ and } E(T_n'') = \frac{\sigma^2}{1-\phi^2}.$$

Proof: The proof is easily done by using the facts that $E(X_i X_{i+1}) = \gamma(1) = \frac{\phi\sigma^2}{1-\phi^2}$ and $E(X_i^2) = \gamma(0) = \frac{\sigma^2}{1-\phi^2}$ for $\forall i = 1,2,\dots,n$, where $\gamma(.)$ is the autocovariance function of the model.

We say that the sequence Y_n converges in the r-th mean(or in the L^r-norm) to Y, for some $r \geq 1$, if

$$\lim_{n\to\infty} E(|\mathrm{Y}_n - Y|^{\mathrm{r}}) = 0,$$

and it is often denoted by $Y_n \xrightarrow{L^r} Y$. The L^2 convergence is established for $r = 2$, where we say that $\mathrm{Y_n}$ converges in L^2 or mean square to Y, denoted by $\mathrm{Y}_n \xrightarrow{L^2} Y$.

Lemma 2.2. Let $E(Z_i) = \alpha$ and $\delta_{ij} = cov(Z_i, Z_j)$ such that $\lim_{|j-i|\to\infty} \delta_{ij} = 0$. Then $\bar{Z} \xrightarrow{L^2} \alpha$, where $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$, in the sense that $\lim_{n\to\infty} E(\bar{Z}_n - \alpha)^2 = 0$.

Proof: Note that

$$E(\bar{Z}_n - \alpha)^2 = E\left(\frac{1}{n}\sum_{i=1}^{n} Z_i - \alpha\right)^2 = \frac{1}{n^2}E\left(\sum_{i=1}^{n} (Z_i - \alpha)\right)^2.$$

Let $Y_i = Z_i - \alpha$, then $E(Y_i) = 0$ and $E(Y_i Y_j) = \delta_{ij}$. Then

$$E(\bar{Z}_n - \alpha)^2 = \frac{1}{n^2}E(\textstyle\sum_{i=1}^{n} Y_i)^2 = \frac{1}{n^2}E\left(\sum_{i=1}^{n}\sum_{j=1}^{n} Y_i Y_j\right) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}.$$

Moreover,

$$\left|\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}\right| \le \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} |\delta_{ij}| \le \frac{1}{n^2}\left[\sum_{i=1}^{n} |\delta_{ii}| + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} |\delta_{ij}|\right] = \frac{1}{n^2}\left[\sum_{i=1}^{n} |\delta_{ii}| + 2\left\{\sum_{i=1}^{n-1}\sum_{j=i+1}^{i+M} |\delta_{ij}| + \sum_{i=1}^{n-1}\sum_{j=i+M+1}^{n} |\delta_{ij}|\right\}\right] = o(1) + o(1) + \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+M+1}^{n} |\delta_{ij}|.$$

The last term in the above equation vanishes, whenever n goes to infinity, by the assumption of $\lim_{|j-i|\to\infty} \delta_{ij} = 0$. The proof of this lemma is completed.

Searle (1971) obtains a well-known relation to compute the covariance between two completely different quadratic forms in a general context. Let $\boldsymbol{Y}_i \sim N(\mathbf{0}, \mathbf{C}_{ii})$ and $\mathbf{C}_{ij} = E(\mathbf{Y}_i^T \mathbf{Y}_j)$, then

$$cov(\mathbf{Y}_1\mathbf{A}_{12}\mathbf{Y}_2, \mathbf{Y}_3\mathbf{A}_{34}\mathbf{Y}_4) = trace(\mathbf{A}_{12}\mathbf{C}_{23}\mathbf{A}_{34}\mathbf{C}_{41} + \mathbf{A}_{12}\mathbf{C}_{24}\mathbf{A}_{34}\mathbf{C}_{31}), \quad (4)$$

see Searle (1971, page 64-65) for more details.

Lemma 2.3. Suppose that X_t follows the first-order autoregressive process defined by (1). Then

$$\lim_{|j-i|\to\infty} cov(X_i X_{i+1}, X_j X_{j+1}) = 0.$$

Proof: Since $E(X_n) = 0$ for $\forall n \in N$,and by using relation (4), take $X_i = \mathbf{Y}_1$, $X_{i+1} = \mathbf{Y}_2, \mathbf{A}_{12} = \mathbf{A}_{34} = 1$, $X_j = \mathbf{Y}_3$and $X_{j+1} = \mathbf{Y}_4$. So, we can write

$$\begin{aligned} cov(X_i X_{i+1}, X_j X_{j+1}) &= trace\{cov(X_{i+1}, X_j)cov(X_{j+1}, X_i) + cov(X_{i+1}, X_{j+1})cov(X_j, X_i)\} \\ &= \gamma(j-i-1)\gamma(i-j-1) + \gamma(j-i)\gamma(i-j) \\ &= \left(\frac{\sigma^2}{1-\phi^2}\right)^2 \left[\phi^{|j-i-1|}\phi^{|i-j-1|} + \phi^{|j-i|}\phi^{|i-j|}\right], \end{aligned}$$

which converges to zero as $|j-i| \to \infty$, because of the stationarity of the model (1).

Lemma 2.4. Suppose that X_t follows the first-order autoregressive process defined by (1). Then

$$\lim_{|j-i|\to\infty} cov(X_i^2, X_j^2) = 0.$$

Proof: The proof is easily followed by noting that $cov(X_i^2, X_j^2) = 2\left(\frac{\sigma^2}{1-\phi^2}\right)^2 \phi^{2|j-i|}$.

Now, we are in a position to state the main result of this paper.

Theorem 2.1. Suppose that X_t is the strictly stationary solution of (1), and let T_n, T_n' and T_n'' are as in (3). Then $T_n \xrightarrow{L^2} \frac{\phi\sigma^2}{1-\phi^2}, T_n' \xrightarrow{L^2} \frac{\sigma^2}{1-\phi^2}$ and $T_n'' \xrightarrow{L^2} \frac{\sigma^2}{1-\phi^2}$.

Proof: First note that T_n, T_n' and T_n'' are unbiased estimators for $\frac{\phi\sigma^2}{1-\phi^2}$, $\frac{\sigma^2}{1-\phi^2}$ and $\frac{\sigma^2}{1-\phi^2}$ respectively, by using Lemma 2.1. The proof is now completed by applying the Lemmas 2.2 and 2.3.

The convergence in L^2 immediately implies the convergence in probability (the weak convergence), and so T_n, T_n' and T_n'' are consistent estimators for parameters $\frac{\phi\sigma^2}{1-\phi^2}$, $\frac{\sigma^2}{1-\phi^2}$ and $\frac{\sigma^2}{1-\phi^2}$, respectively. Now, the consistency of the WPLE, $\tilde{\phi}$ can be easily derived, by using the Slutsky's theorem.

3.Conclusion

This note is concerned with the asymptotic L^2 properties of PL procedures for the parameter of the AR(1) models. We have applied a simple method to establish the L^2 convergence of the estimators proposed by Davis and Yau (2011) which in turn establish their weak convergence. The weak convergence of the WPLE is also studied.

Acknowledgement

The authors would like to thank the Editor and and the reviewers for many constructive suggestions.

References

Cox, D. R., & Ried, N. (2004). A note on pseudo likelihood constructed from marginal densities. *Biometrika*, 91(3), 729–737.

Davis, R. A., & Yau, C. Y. (2011). Comments on pairwise likelihood in time series models. *Statistica Sinica*, 21(1), 255–277.

Lindsay, B. G. (1988). Composite likelihood methods. *Contemporary Mathematics*, 80, 221–240.

Searle, S. R. (1971). *Linear models*. John-Wiley, New York.

Varin, C. (2008). On composite marginal likelihoods. *Advances in Statistical Analysis*, 92(1), 1–28.

Varin, C., Reid, N., & Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, 21(1), 5–42.