APPROXIMATE BAYES ESTIMATION OF PARAMETERS OF THE NEAR(2) MODEL

Subashan Perera

Center on Aging

University of Kansas Medical Center

Kansas City, KS 66160-7117

Key Words: exponential; autoregressive; time series; random coefficients; two-stage conditional least squares; out-of-range estimates

ABSTRACT

The published method for parameter estimation of the NEAR(2) model suffers from estimates frequently falling outside of the parameter space when the sample size is small and/or the true values of the parameters are close to the boundary of the parameter space. In order to alleviate this problem, while retaining the asymptotic consistency of the published estimators, an approximate Bayes correction is suggested. The approximate Bayes correction treats the asymptotic distribution of the estimators as an approximate likelihood for the data and employs a prior distribution supported in the parameter space. The results of a simulation study are also presented.

1. INTRODUCTION

The existence of time series data that cannot be satisfactorily modelled with the standard Gaussian autoregressive-moving average models has lead to the development of time series models based on the exponential distribution. The most general and flexible of these is the NEAR(2) model proposed by Lawrance (1980) and used to model a series of wind velocity data by Lawrance and Lewis (1985). The NEAR(2) model is defined by:

$$X_t = \begin{cases} \beta_1 X_{t-1} & w.p. \;\; \alpha_1 \\ \beta_2 X_{t-2} & w.p. \;\; \alpha_2 \\ 0 & w.p. \;\; 1-\alpha_1-\alpha_2 \end{cases} + \epsilon_t, \qquad (1)$$

where the residual sequence $\{\epsilon_t\}$ is defined as:

$$\epsilon_t = \begin{cases} b_2 E_t & w.p. \;\; p_2 \\ b_3 E_t & w.p. \;\; p_3 \\ E_t & w.p. \;\; 1-p_2-p_3 \end{cases} .$$

Here, $\{E_t\}$ is an independent and identically distributed sequence of standard exponential variates, $p_2 = \{(\alpha_1\beta_1+\alpha_2\beta_2)b_2-(\alpha_1+\alpha_2)\beta_1\beta_2\}/\{(b_2-b_3)(1-b_3)\}$, $p_3 = \{(\alpha_1+\alpha_2)\beta_1\beta_2-(\alpha_1\beta_1+\alpha_2\beta_2)b_3\}/\{(b_2-b_3)(1-b_3)\}$, $b_2 = \{s+(s^2-4r)^{1/2}\}/2$ and $b_3 = \{s-(s^2-4r)^{1/2}\}/2$, where $s = (1-\alpha_1)\beta_1+(1-\alpha_2)\beta_2$ and $r = (1-\alpha_1-\alpha_2)\beta_1\beta_2$. Chan (1988) showed that necessary and sufficient conditions for the existence of a strictly stationary and ergodic NEAR(2) process, which is Markov, are $0 \le \alpha_1, \alpha_2, \alpha_1+\alpha_2 \le 1$, $0 \le \beta_1, \beta_2 \le 1$ and $\alpha_1\beta_1+\alpha_2\beta_2<1$. These conditions will define the parameter space Ω for the

NEAR(2) model.

Section 2 contains a brief review of estimation methods for the NEAR(2) model available in the literature. In Section 3 a brief overview of the approximate Bayes procedure is provided. How the approximate Bayes procedure improves the existing estimators and the results of a simulation study are discussed in Sections 4 and 5 respectively.

2. PARAMETER ESTIMATION

In principle, maximum likelihood estimation (MLE) is possible since a likelihood function (conditional on X_1 and X_2) can easily be constructed as

$$L(\alpha_1, \alpha_2, \beta_1, \beta_2) = \prod_{t=3}^{n} f(X_t \mid X_{t-1}, X_{t-2}),$$

where, from (1),

$$f(X_t \mid X_{t-1}, X_{t-2}) = \alpha_1 f_\epsilon(X_t - \beta_1 X_{t-1}) + \alpha_2 f_\epsilon(X_t - \beta_2 X_{t-2}) + (1 - \alpha_1 - \alpha_2) f_\epsilon(X_t)$$

and $f_\epsilon(.)$ is the probability density function of a mixed exponential distribution given by $f_\epsilon(\epsilon) = (1 - p_2 - p_3)e^{-\epsilon} + (p_2/b_2)e^{-\epsilon/b_2} + (p_3/b_3)e^{-\epsilon/b_3}$, $\epsilon > 0$. However, in practice, MLE is difficult because each $f(X_t \mid X_{t-1}, X_{t-2})$, $t = 1,2,...$, viewed as a function of $(\alpha_1, \alpha_2, \beta_1, \beta_2)$, has discontinuities at $\beta_1 = X_t / X_{t-1}$ and $\beta_2 = X_t / X_{t-2}$, and the total number of discontinuities increases with the sample size. Smith (1986) attempted to circumvent the problem of discontinuities by smoothing $L(\alpha_1, \alpha_2, \beta_1, \beta_2)$, but the necessity of choosing the smoothing parameter by trial and error makes his approach unsuitable for study by simulation. He also investigated maximizing $L(\alpha_1, \alpha_2, \beta_1, \beta_2)$ on a discrete grid along (β_1, β_2) to obtain the MLE's. However, there is no guarantee that the maximum actually occurs at a point on the grid.

For NEAR(2) parameter estimation, Karlsen and Tjøstheim (1988) used the two-stage conditional least squares method proposed by Nicholls and Quinn (1982) for their random coefficients autoregressive (RCA) processes, since they are a generalization of the NEAR models. This involves the reparameterization $a_i = \alpha_i \beta_i,\ \sigma_{ii} = \beta_i^2 \alpha_i (1-\alpha_i),\ i = 1,2,$ and using conditional least squares (CLS) for the minimization of $Q_1(\theta) = \sum_{t=3}^{n} \{(X_t - 1) - E_\theta(X_t - 1 \mid F_{t-1})\}^2$ to obtain $\hat{a}_1$ and $\hat{a}_2$, where F_{t-1} is the -field generated by $(X_1, \ldots, X_{t-1})$ and $\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2)$ or $\theta = (a_1, a_2, \sigma_{11}, \sigma_{22})$ as appropriate. Then, CLS is again applied to squared residuals by minimizing $Q_2(\theta) = \sum_{t=3}^{n} \{\hat{U}_t^2 - E_\theta(\hat{U}_t^2 \mid F_{t-1})\}^2$ to obtain $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$, where $\hat{U}_t = (X_1 - 1) - \hat{a}_1(X_{t-1} - 1) - \hat{a}_2(X_{t-2} - 1)$. The resulting estimators are

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = \left\{ \sum_{t=3}^{n} \begin{bmatrix} (X_{t-1} - 1)^2 & (X_{t-1} - 1)(X_{t-2} - 1) \\ (X_{t-1} - 1)(X_{t-2} - 1) & (X_{t-2} - 1)^2 \end{bmatrix} \right\}^{-1} \sum_{t=3}^{n} \begin{bmatrix} (X_t - 1)(X_{t-1} - 1) \\ (X_t - 1)(X_{t-2} - 1) \end{bmatrix}$$

and

$$\begin{bmatrix} \hat{\sigma}_{11} \\ \hat{\sigma}_{22} \end{bmatrix} = \left\{ \sum_{t=3}^{n} \begin{bmatrix} (X_{t-1}^2 - 2)^2 & (X_{t-1}^2 - 2)(X_{t-2}^2 - 2) \\ (X_{t-1}^2 - 2)(X_{t-2}^2 - 2) & (X_{t-2}^2 - 2)^2 \end{bmatrix} \right\}^{-1}$$

$$\times \sum_{t=3}^{n} \begin{bmatrix} X_{t-1}^2 - 2 \\ X_{t-2}^2 - 2 \end{bmatrix} \left\{ \hat{U}_t - 1 + (\hat{a}_1 - \hat{a}_2)^2 + 2\hat{a}_1\hat{a}_2 X_{t-1} X_{t-2} \right\}$$

from which the estimators of the original parameters can be recovered by using $\hat{\alpha}_i = \hat{a}_i^2 / (\hat{\sigma}_{ii} + \hat{a}_i^2),\ \hat{\beta}_i = (\hat{\sigma}_{ii} + \hat{a}_i^2)/\hat{a}_i,\ i = 1,2.$ However, in the simulation results presented in Karlsen and Tjøstheim (1988), these estimates lie outside the parameter space a high proportion of times when the sample size n is small. The strong consistency and asymptotic normality of these estimates are stated in a theorem in Karlsen and Tjøstheim (1988), but are not exploited. These asymptotic properties can also be obtained by modifying the proof of Nicholls and Quinn (1982) to account for the smaller number of parameters in the NEAR(2) model than the comparable RCA model. Details and a proof of the Theorem 2.1 are presented in Perera (2000).

Theorem 2.1:

Let $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$ and $\hat{\beta}_2$ be the Karlsen and Tjøstheim (1988) estimators of the parameters of a NEAR(2) model. Then,

$$\sqrt{n}\hat{W}^{-1/2}(\hat{\alpha}_1-\alpha_1,\ \hat{\alpha}_2-\alpha_2,\ \hat{\beta}_1-\beta_1,\ \hat{\beta}_2-\beta_2)' \ \overset{D}{\rightarrow}\ N(0,I_4),$$

where $\hat{W}$ is a strongly consistent estimate of the asymptotic covariance matrix W. See the Appendix for definitions of W and $\hat{W}$.

3. APPROXIMATE BAYES ESTIMATION

An approximate Bayes procedure explored by Ogunyemi, Hutton and Nelson (1993) treats the asymptotic distribution of the parameter estimators of a complex model as an approximate likelihood for the data that can be used to obtain approximate posterior estimators in the presence of a prior distribution. Also see Chapter 2 of Tanner (1996) for a discussion of normal approximations to a likelihood. To apply this approach, suppose that a family of densities in indexed by an unknown parameter $\theta \in \Omega \subseteq \mathbb{R}^p$ where Ω is the parameter space, $\mathbb{R}^p$ is p-dimensional Euclidean space, and an estimator $\hat{\theta}_n$ of θ exists such that

$$\hat{W}_n^{-1/2}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, I_p), \tag{2}$$

where $\hat{W}$ is free of θ and possibly random. Then,

$$\tilde{L}(\theta) = (2\pi\,|\hat{W}_n|)^{-p/2} \exp\{-\tfrac{1}{2}(\hat{\theta}_n - \theta)' \hat{W}_n^{-1}(\hat{\theta}_n - \theta)\}$$

may be viewed as an approximate likelihood. In the presence of a prior density $\pi(\theta)d\theta$, the approximate posterior distribution is given by

$$\pi(d\theta \mid \hat{\theta}_n) = \tilde{L}(\theta)\pi(\theta)d\theta \Big/ \int_\Omega \tilde{L}(\theta)\pi(\theta)d\theta .$$

The approximate Bayes estimator of θ is then given by

$$\tilde{\theta} = \int_\Omega \theta \tilde{L}(\theta)\pi(\theta)d\theta \Big/ \int_\Omega \tilde{L}(\theta)\pi(\theta)d\theta ,$$

and Theorem 3.1, which appeared in Ogunyemi, Hutton and Nelson (1993), establishes its strong consistency.

Theorem 3.1

Suppose that the parameter space Ω is an open subset of $\mathbb{R}^p$ and that the prior distribution $\pi(\theta)$ has a positive density with respect to Lebesgue measure such that the marginal densities $\pi_i(\theta_i)$ satisfy the conditions that $\pi_i(\theta_i)$ and $\theta_i \pi_i(\theta_i)$ are bounded and uniformly continuous on Ω, $i = 1, \ldots, p$. If (2) holds and $\hat{\theta}_n \to \theta$ a.e. on an event E, then $\tilde{\theta}_n \to \theta$ a.e. on E.

4. IMPROVING THE EXISTING ESTIMATORS

The approximate Bayes procedure can be used to improve the existing estimators of Karlsen and Tjøstheim (1988) by reducing the proportion of estimates that fall outside Ω. If a uniform prior density $\pi(\theta) = 1_\Omega(\alpha_1, \alpha_2, \beta_1, \beta_2)$ is used, where $1_\Omega(.)$ is the indicator function over the set Ω and let $\hat{\theta}_n = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)$, the approximate posterior density is

$$\pi(\theta \mid \hat{\theta}_n) = \frac{(2\pi \mid \hat{W}_n \mid)^{-2} \exp\left\{-\frac{1}{2}(\hat{\theta}_n - \theta)' \hat{W}_n^{-1} (\hat{\theta}_n - \theta)\right\} 1_\Omega(\theta)}{\int_\Omega (2\pi \mid \hat{W}_n \mid)^{-2} \exp\left\{-\frac{1}{2}(\hat{\theta}_n - \theta)' \hat{W}_n^{-1} (\hat{\theta}_n - \theta)\right\} 1_\Omega(\theta)}$$

$$= (1/C) \exp\left\{-(\hat{\theta}_n - \theta)' \hat{W}_n^{-1} (\hat{\theta}_n - \theta)/2\right\} 1_\Omega(\theta),$$

where C is a normalizing constant which will make $\pi(\theta \mid \hat{\theta}_n)$ integrate to unity. Note that $\pi(\theta \mid \hat{\theta}_n)$ is a 4-variate normal distribution with "mean" $\hat{\theta}_n$ and "covariance matrix" $\hat{W}_n$ truncated at the boundary of the parameter space. The approximate Bayes estimator

$$\tilde{\theta}_n = (\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\beta}_1, \tilde{\beta}_2) = \int_\Omega \theta \, \pi(\theta \mid \hat{\theta}_n) d\theta$$

will be strongly consistent by Theorem 3.1, and is forced to lie within Ω, thus pulling an estimate $\hat{\theta}_n$ that falls outside Ω inside.

Although computation of $\tilde{\theta}_n$ is possible by (numerical) integration of $\pi(\theta \mid \hat{\theta}_n)$ over Ω, estimating it by using the law of large numbers is less time consuming. A large number of 4-variate gaussian deviates with mean $\hat{\theta}_n$ and covariance matrix $\hat{W}_n$ could be generated, and the arithmetic mean of those that lie within Ω could be taken as $\tilde{\theta}_n$.

5. SIMULATION RESULTS AND DISCUSSION

In order to assess the performance of the approximate Bayes correction in reducing the proportion of out-of-parameter-space estimates, a simulation study was conducted by generating 500 NEAR(2) processes from each of the sample size-paramater combinations considered. For each such process, the parameters were estimated by the method of Karlsen and Tjøstheim (1988), and were improved using an approximate Bayes correction. The averages and root mean square errors (RMSE) of both sets of estimates, and the percentage of estimates that fell within Ω for the parameter combination $\alpha_1 = 0.4,\ \alpha_2 = 0.5,\ \beta_1 = 0.6,$ $\beta_2 = 0.7$ are presented in Table I. Table II contains similar information for the parameter combination $\alpha_1 = \alpha_2 = 0.25,\ \beta_1 = \beta_2 = 0.5$. Results of these two parameter-sample size combinations were chosen to be presented so that comparisons could be made to the results of Smith (1986) and Karlsen and Tjøstheim (1988).

The approximate Bayes procedure has reduced the percentage of Karlsen and Tjøstheim (1988) estimates that fall outside Ω, by as much as 20.6% when $\alpha_1 = 0.4,\ \alpha_2 = 0.5,\ \beta_1 = 0.6,\ \beta_2 = 0.7$ and $n = 200$. When the sample size is smaller ($n \leq 1{,}000$), this reduction in statistically significant based on a McNemar test and a type I error rate of 0.05, as indicated with an * in Tables I-II. As the sample size increases, both sets of estimates seem to converge to the true values

Table I

Averages, root mean square errors (within parenthesis), and percent within parameter space of both sets of estimates for the parameter combination $\alpha_1 = 0.4$, $\alpha_2 = 0.5$, $\beta_1 = 0.6$, $\beta_2 = 0.7$. * indicates a significant difference in percentage of estimates in Ω between estimation methods.

Karlsen and Tjøstheim (1988) Estimates:					
n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	% in Ω
*100	0.393(0.148)	0.464(0.154)	0.665(0.181)	0.711(0.146)	42.6
*200	0.392(0.120)	0.490(0.119)	0.627(0.156)	0.680(0.122)	58.2
*500	0.406(0.089)	0.494(0.097)	0.608(0.119)	0.698(0.095)	78.6
*1,000	0.406(0.071)	0.504(0.068)	0.599(0.092)	0.689(0.073)	89.2
10,000	0.401(0.026)	0.501(0.024)	0.602(0.037)	0.697(0.027)	100.0
30,000	0.400(0.014)	0.501(0.013)	0.601(0.021)	0.700(0.015)	100.0
Approximate Bayes Estimates:					
n	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	% in Ω
*100	0.333(0.213)	0.390(0.248)	0.522(0.295)	0.577(0.311)	63.2
*200	0.371(0.165)	0.445(0.183)	0.567(0.222)	0.622(0.228)	71.2
*500	0.392(0.116)	0.475(0.135)	0.575(0.165)	0.667(0.167)	85.0
*1,000	0.399(0.089)	0.494(0.097)	0.586(0.124)	0.674(0.121)	92.4
10,000	0.401(0.026)	0.501(0.024)	0.602(0.037)	0.697(0.027)	100.0
30,000	0.400(0.014)	0.501(0.013)	0.601(0.021)	0.700(0.015)	100.0

of the parameters, as expected. However, the problem of estimates falling outside and Tjøstheim (1988) estimates that fall outside Ω, by as much as 20.6% when $\alpha_1 = 0.4$, $\alpha_2 = 0.5$, $\beta_1 = 0.6$, $\beta_2 = 0.7$ and $n = 200$. When the sample size is smaller ($n \leq 1{,}000$), this reduction in statistically significant based on a McNemar test and a type I error rate of 0.05, as indicated with an * in Tables I-II. As the sample size increases, both sets of estimates seem to converge to the true values of the parameters, as expected. However, the problem of estimates falling outside

Table II

Averages, root mean square errors (within parenthesis), and percent within parameter space of both sets of estimates for the parameter combination combination $\alpha_1 = \alpha_2 = 0.25$, $\beta_1 = \beta_2 = 0.5$. * indicates a significant difference in percentage of estimates in Ω between estimation methods.

Karlsen and Tjøstheim (1988) Estimates:					
n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	% in Ω
*100	0.350(0.206)	0.301(0.176)	0.517(0.203)	0.594(0.230)	27.6
*200	0.337(0.196)	0.302(0.163)	0.495(0.189)	0.528(0.185)	41.4
*500	0.301(0.139)	0.294(0.142)	0.472(0.181)	0.483(0.183)	74.8
*1,000	0.287(0.106)	0.294(0.114)	0.464(0.161)	0.461(0.160)	93.4
10,000	0.253(0.030)	0.258(0.033)	0.497(0.059)	0.492(0.058)	100.0
30,000	0.251(0.020)	0.251(0.018)	0.499(0.040)	0.499(0.035)	100.0
Approximate Bayes Estimates:					
n	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	% in Ω
*100	0.302(0.216)	0.284(0.198)	0.444(0.278)	0.465(0.283)	44.2
*200	0.312(0.214)	0.279(0.186)	0.426(0.254)	0.450(0.249)	56.4
*500	0.290(0.145)	0.289(0.152)	0.444(0.210)	0.456(0.210)	82.0
*1,000	0.285(0.112)	0.294(0.122)	0.456(0.174)	0.453(0.170)	96.0
10,000	0.253(0.030)	0.258(0.033)	0.497(0.059)	0.493(0.058)	100.0
30,000	0.251(0.020)	0.251(0.018)	0.499(0.040)	0.499(0.035)	100.0

Ω and therefore the need for the approximate Bayes procedure is little in larger (n>1,000) samples. The smaller RMSE's of Karlsen and Tjøstheim (1988) estimates in smaller samples is due to the RMSE being based on a smaller number of presumably better behaved data sets. Similar results are obtained for other parameter combinations with the observation that the performance of both methods deteriorate near the boundaries of Ω where α_1 or α_2 is small.

In principle, the percentage of estimates within Ω must be 100% for the approximate Bayes procedure since the support of the approximate posterior distribution is Ω . However, lack of information in some data sets cause some matrices needed for the computation of $\hat{W}$ to be less than full rank, thereby not allowing the implementation of the approximate Bayes procedure.

APPENDIX

Let $v_{11} = E\{(X_{t-1} - 1)^2\}$, $v_{12} = E\{(X_{t-1} - 1)(X_{t-2} - 1)\}$, $v_{22} = E\{(X_{t-2} - 1)^2\}$,

$r_{11} = E\{(X_{t-1}^2 - 2)^2\}$, $r_{12} = E\{(X_{t-1}^2 - 2)(X_{t-2}^2 - 2)\}$, $r_{22} = E\{(X_{t-2}^2 - 2)^2\}$,

$U_t = X_t - 1 - a_1(X_{t-1} - 1) - a_2(X_{t-2} - 1)$ and

$\xi_t = X_t^2 - 1 + (a_1 - a_2)^2 + 2a_1a_2X_{t-1}X_{t-2} - \sigma_{11}(X_{t-1}^2 - 2) - \sigma_{22}(X_{t-2}^2 - 2)$.

Then the elements of asymptotic covariance matrix S of $\hat{a}_1$, $\hat{a}_2$, $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$ are given by

$$s_{11} = E[\{v_{11}(X_{t-1} - 1) + v_{12}(X_{t-2} - 1)\}^2 U_t^2],$$

$$s_{12} = E[\{v_{11}(X_{t-1} - 1) + v_{12}(X_{t-2} - 1)\}\{v_{12}(X_{t-1} - 1) + v_{22}(X_{t-2} - 1)\}U_t^2],$$

$$s_{13} = E[\{v_{11}(X_{t-1} - 1) + v_{12}(X_{t-2} - 1)\}\{r_{11}(X_{t-1}^2 - 2) + r_{12}(X_{t-2}^2 - 2)\}U_t\xi_t],$$

$$s_{14} = E[\{v_{11}(X_{t-1} - 1) + v_{12}(X_{t-2} - 1)\}\{r_{12}(X_{t-1}^2 - 2) + r_{22}(X_{t-2}^2 - 2)\}U_t\xi_t],$$

$$s_{22} = E[\{v_{12}(X_{t-1} - 1) + v_{22}(X_{t-2} - 1)\}^2 U_t^2],$$

$$s_{23} = E[\{v_{12}(X_{t-1} - 1) + v_{22}(X_{t-2} - 1)\}\{r_{11}(X_{t-1}^2 - 2) + r_{12}(X_{t-2}^2 - 2)\}U_t\xi_t],$$

$$s_{24} = E[\{v_{12}(X_{t-1} - 1) + v_{22}(X_{t-2} - 1)\}\{r_{12}(X_{t-1}^2 - 2) + r_{22}(X_{t-2}^2 - 2)\}U_t\xi_t],$$

$$s_{33} = E[\{r_{11}(X_{t-1}^2 - 2) + r_{12}(X_{t-2}^2 - 2)\}\xi_t^2],$$

$$s_{34} = E[\{r_{11}(X_{t-1}^2 - 2) + r_{12}(X_{t-2}^2 - 2)\}\{r_{12}(X_{t-1}^2 - 2) + r_{22}(X_{t-2}^2 - 2)\}\xi_t^2], \text{ and}$$

$$s_{44} = E[\{r_{12}(X_{t-1}^2 - 2) + r_{22}(X_{t-2}^2 - 2)\}\xi_t^2].$$

By the delta method, the asymptotic covariance matrix of $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$ and $\hat{\beta}_2$ is then given by $W = G' S G$ where G is the partial derivative matrix

$$G = \begin{bmatrix} 2a_1\sigma_{11}/(\sigma_{11}+a_1^2)^2 & 0 & -a_1^2/(\sigma_{11}+a_1^2)^2 & 0 \\ 0 & 2a_2\sigma_{22}/(\sigma_{22}+a_2^2)^2 & 0 & -a_2^2/(\sigma_{22}+a_2^2)^2 \\ 1-\sigma_{11}/a_1^2 & 0 & 1/a_1 & 0 \\ 0 & 1-\sigma_{22}/a_2^2 & 0 & 1/a_2 \end{bmatrix}.$$

By the ergodic theorem, v_{11} can be strongly consistently estimated by $\hat{v}_{11} = \frac{1}{n}\sum_{t=3}^{n}(X_{t-2}-1)^2$. By estimating other quanities similarly, and using the strong consistency of $\hat{a}_1$, $\hat{a}_2$, $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$, $\hat{W}$ can easily be obtained.

ACKNOWLEDGEMENTS

This work is supported in part by NSF-SCREMS grant DMS-9628643. I wish to thank Professor Paul I. Nelson for his guidance.

BIBLIOGRAPHY

Chan, K. (1988). “On the existence of the stationary and ergodic NEAR(p) model,” *Journal of Time Series Analysis*, 9, 319-328.

Karlsen, H. and Tjøstheim, D. (1988). “Consistent estimates for the NEAR(2) and NLAR(2) time series models,” *Journal of the Royal Statistical Society-Series B*, 50, 313-320.

Lawrance, A. J. (1980). “Some autoregressive models for point processes,” *Point*

Processes and Queueing Problems (P. Bartfai and J. Tomko, Eds.), 257-275. Colloquia Mathematica Societis Janos Bolyai, 24, Amsterdam: North Holland.

Lawrance, A. J. and Lewis, P. A. W. (1985). "Modelling and residual analysis of nonlinear autoregressive time series in exponential variables," *Journal of the Royal Statistical Society-Series B*, 47, 165-202.

Nicholls, D. F. and Quinn, B. G. (1982). *Random Coefficient Autoregressive Models: an Introduction*, Heidelberg: Springer-Verlag.

Ogunyemi, O. T., Hutton, J. E. and Nelson, P. I. (1993). "Approximate Bayes estimators for stochastic processes," *Journal of Statistical Computing and Simulation*, 45, 219-231.

Perera, S. (2000). Ph. D. Dissertation, *Autoregressive Time Series Models with Exponential Marginal Distributions*, Kansas State University, Manhattan.

Smith, R. L. (1986). "Maximum likelihood estimation for the NEAR(2) model," *Journal of the Royal Statistical Society-Series A*, 48, 251-257.

Tanner, M. A. (1996). *Tools for Statistical Inference*, Third Edition. New York: Springer-Verlag.