Journal of the Nigerian Association of Mathematical Physics
Volume **15 (November, 2009), pp 117 – 122**

On full order autoregressive model fitting

Ette Harrison Etuk
Department of Mathematics and Computer Science
Rivers State University of Science and Technology
Port Harcourt, Nigeria.

Abstract

This is a study of the relative performance of nine automatic autoregressive order determination criteria for full-order modeling using the least squares approach. Of the nine: AIC, FPE, CAT_2, S, FPE4, Φ, SIC, BIC and CAT_3, we have found the trio, AIC, FPE and CAT_3, to be the most consistent. The rest underestimate relatively.

1.0 Introduction.

A stationary time series $\{X_t\}$ is said to follow an autoregressive model of order p, if it satisfies the following difference equation

$$X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... + \alpha_p X_{t-p} = \varepsilon_t \quad (1.1)$$

where $\{\varepsilon_t\}$ is a white noise process with variance σ^2. For stationarity of (1.1), it is required that

$$1 + \alpha_1 z + \alpha_2 z^2 + ... + \alpha_p z^p \neq 0, | z | \leq 1$$

Any stationary time series can be expressed as an infinite-order autoregressive (AR) process. In practice, autoregressive modeling amounts to approximating the infinite order auto-regression by the finite one (1.1). The problem is two-fold: the determination of the order p and the estimation of the parameters α_1, …, α_p. The latter can be achieved by the use of any conventional method like those of the least error sum of squares (or the Yule-Walker (Y-W) method), maximum likelihood, maximum entropy, etc. The former has been addressed by many authors some of whom have proposed automatic criteria for it. For example, Akaike [1, 3, 4] has proposed Final Prediction Error (FPE), Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) respectively. Schwarz [13] proposed his Information Criterion (SIC) and [20] proposed Criterion Autoregression Transfer Function (CAT). Order determination is very important since underestimation increases the residual variance and overestimation decreases the reliability of the model.

A comparative study of the criteria has engaged the attention of many researchers such as [5], [11], [12], [16], [18] and [19] to mention a few. In this paper we are interested in the comparative performance of nine criteria AIC, FPE, SIC, S, Φ, CAT_2, CAT_3, FPE4 and BIC for the least squares (Y-W) approach to model estimation. We shall use simulated as well as real series for the study. For the Monte-Carlo method, we shall observe the comparative efficiencies of the criteria with change in sample size. Etuk [12] has compared the same criteria for the maximum entropy method of estimation.

2.0 Order Determination

Consider a realization X_1, X_2, …, X_N of the stationary time series $\{X_t\}$ with $E(X_t) = \mu$. The autocovariance function $\{\gamma_k\}$ and the autocorrelation function $\{\rho_k\}$ defined by

$$\gamma_k = E\{[X_t - \mu][X_{t-k} - \mu]\}, k = 0, \pm 1, \pm 2, \Lambda \text{ and } \rho_k = \frac{\gamma_k}{\gamma_0}, k = 0,1,2,\Lambda$$

respectively, may be estimated by $\{c_k\}$ and $\{r_k\}$ given by

$$c_k = \frac{1}{N} \sum_{t=k+1}^{N} \{[X_t - \overline{X}][X_{t-k} - \overline{X}]\}$$

$$r_k = \frac{c_k}{c_0}, k = 0,1,2,\Lambda$$

respectively, where $\overline{X} = \frac{1}{N} \Sigma X_t$. After specifying a maximum lag L, the order in the range 0, 1, 2, …, L is that for which any of the following criteria is minimum:

$$FPE(p) = \hat{\sigma}_p^2 (1 + \frac{p}{N}), p = 0,1,2,\Lambda \text{ (Akaike, [1])}$$

$$FPE\alpha(p) = \frac{1 + \frac{\alpha p}{N}}{1 - \frac{p}{N}} \hat{\sigma}_p^2, \alpha > 0, p = 0,1,2,\Lambda \text{ (Bhansali and Downham, [6])}$$

$$AIC(p) = N \ln \tilde{\sigma}_p^2 + 2p, p = 0,1,2,\Lambda \text{ (Akaike, [3])}$$

$$BIC(p) = N \ln \tilde{\sigma}_p^2 - (N-p)\ln(1 - \frac{p}{N}) + p \ln N + p \ln\{\frac{1}{p}[\frac{c_0}{\tilde{\sigma}_p^2} - 1]\}, p = 0,1,2,\Lambda \text{ (Akaike, [4])}$$

$$SIC(p) = N \ln \tilde{\sigma}_p^2 + p \ln N, p = 0,1,2,\Lambda \text{ (Schwarz, 13)}$$

$$CAT_2(p) = \frac{1}{N}(\sum_{j=1}^{p} \frac{1}{\hat{\sigma}_j^2}) - \frac{1}{\hat{\sigma}_p^2} = -\left(1 + \frac{1}{N}\right) p = 0, p = 1,2,\Lambda \text{ (Parzen, [2])}$$

$$CAT_3(p) = \frac{1}{N}(\sum_{j=0}^{p} \frac{1}{\hat{\sigma}_j^2}) - \frac{1}{\hat{\sigma}_p^2}, p = 0,1,2,\Lambda \text{ (Tong, [26])}$$

$$\Phi(p) = \ln \tilde{\sigma}_p^2 + \frac{2}{N} pc \ln\ln N, c > 1, p = 0,1,2,\Lambda \text{ (Hannan and Quinn, [15])}$$

$$S_N(p) = (N + 2p)\tilde{\sigma}_p^2, p = 0,1,2,\Lambda \text{ (Shibata, [25])}$$

where $\hat{\sigma}_p^2$ and $\tilde{\sigma}_p^2$ are respectively the least squares and the maximum likelihood estimates of the residual variance.

With the order p estimated, an approximate least squares (Yule-Walker(Y-W)) approach to the estimation of (1.1) is obtained by the application of the recursive formula:

$$\phi_{k+1} = \hat{\alpha}_{k+1,k+1} = \frac{r_{k+1} - \sum_{j=1}^{k} \hat{\alpha}_{rj} r_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\alpha}_{kj} r_j}$$

$$\hat{\alpha}_{k+1,j} = \hat{\alpha}_{kj} - \phi_{k+1}\hat{\alpha}_{k,k-j+1}, \; j = 1,2,\Lambda\ ,k$$

where (1.1) is written more specifically as $X_t + \alpha_{p1} X_{t-1} + \alpha_{p2} X_{t-2} + ... + \alpha_{PP} X_{t-p} = \varepsilon_t$.

3.0 Simulation results

Two AR(2) series with (α_1, α_2) equal to (-0.66, 0.10), (-0.46, 0.08) were simulated twenty independent times each. In the sequel we shall refer to them as series I and II respectively. We used sample sizes of 50, 150 and 250 for each simulated series.

The white noise process of each simulation is a sequence of pseudorandom numbers generated by the use of the RAN function of the FORTRAN 77 language. The sequence is made standard normal. In the sequel we use c = 1.5, an arbitrary choice, for Φ. We observed the frequency with which the correct order of 2 was picked by each order determination criterion for each series. Table 3.1 gives the summary of the results.

Table 3.1: Frequency out of 20 of correct picking of order 2

Series	**N**	**AIC**	**FPE**	**BIC**	**Φ**	**SIC**	**S**	**CAT_2**	**CAT_3**	**FPE4**
I	50	7	7	6	4	4	7	3	7	6
	150	5	5	5	2	2	6	7	5	4
	250	11	11	11	11	9	12	12	11	11
II	50	5	5	4	2	3	4	2	5	3
	150	7	7	3	2	2	5	1	7	2
	250	4	4	3	4	3	3	2	4	3
TOTAL		39	39	32	25	23	37	21	39	29

For all the 120 simulated series i.e. 60 independent series for each model, AIC, FPE and CAT_3 agreed with each other always. AIC and FPE are known to be asymptotically equivalent (e.g. Priestley, [21]). Tong [26] analytically proved that CAT_3 and AIC possess the same local behaviour. Consistent with Shibata's [24] result that asymptotically AIC has a zero probability of underestimation, we observed that the trio tended to overestimate, this tendency increasing with sample size. Akaike [2] also noted this with FPE.

BIC, Φ, SIC and FPE4 recommended the same order most of the time, tending to underestimate. Whereas increase in sample size was observed to enhance the consistency of Φ and SIC, it appeared to decrease that of FPE4 which tended to overestimate relatively.

S showed a tendency to overestimate which increased with sample size but which is less than that of FPE. CAT_2 recommended zero order most often. This was also observed by Tong(1977). Generally, as obvious from Table 3.1, AIC, FPE, CAT_3 performed best, S next, then BIC, FPE4 and so on and CAT_2 least.

4.0 Real series analysis.

Three well analysed series are used. Our method is to compare the models selected by the criteria within an order range of 0 to 30, with earlier ones. Also we shall subject each selected model to the Box-Pierce [8] portmanteau test. Another diagnostic checking method we shall use is comparing the spectra of the models with the estimated raw spectrum. Other diagnostic aids we use are the inverse autocorrelation function (IACF) and the partial autocorrelation function (PACF) (see Etuk [12]).

4.1 Wolfer's sunspot series (1700-1955)

The yearly data are available as from1700 onwards (see Waldmeier, [27]). We used the 256 values of 1700 to 1955 inclusive. FPE, AIC and CAT_3 recommend the model

$$X_t - 1.192X_{t-1} + 0.414X_{t-2} + 0.187X_{t-3} - 0.201X_{t-4} + 0.187X_{t-5} - 0.102X_{t-6} + 0.093X_{t-7} - 0.103X_{t-8} - 0.999X_{t-9} = \varepsilon_t, \hat{\sigma}^2 = 208.13 \quad (4.1)$$

with the Box-Pierce [8] statistic R = 356.846.

FPE4, Φ and S choose the AR(8):

$$X_t - 1.214X_{t-1} + 0.427X_{t-2} + 0.178X_{t-3} - 0.185X_{t-4} + 0.169X_{t-5} - 0.085X_{t-6} + 0.135X_{t-7} - 0.223X_{t-8} = \varepsilon_t, \hat{\sigma}^2 = 210.18, R = 27.17 \quad (4.2)$$

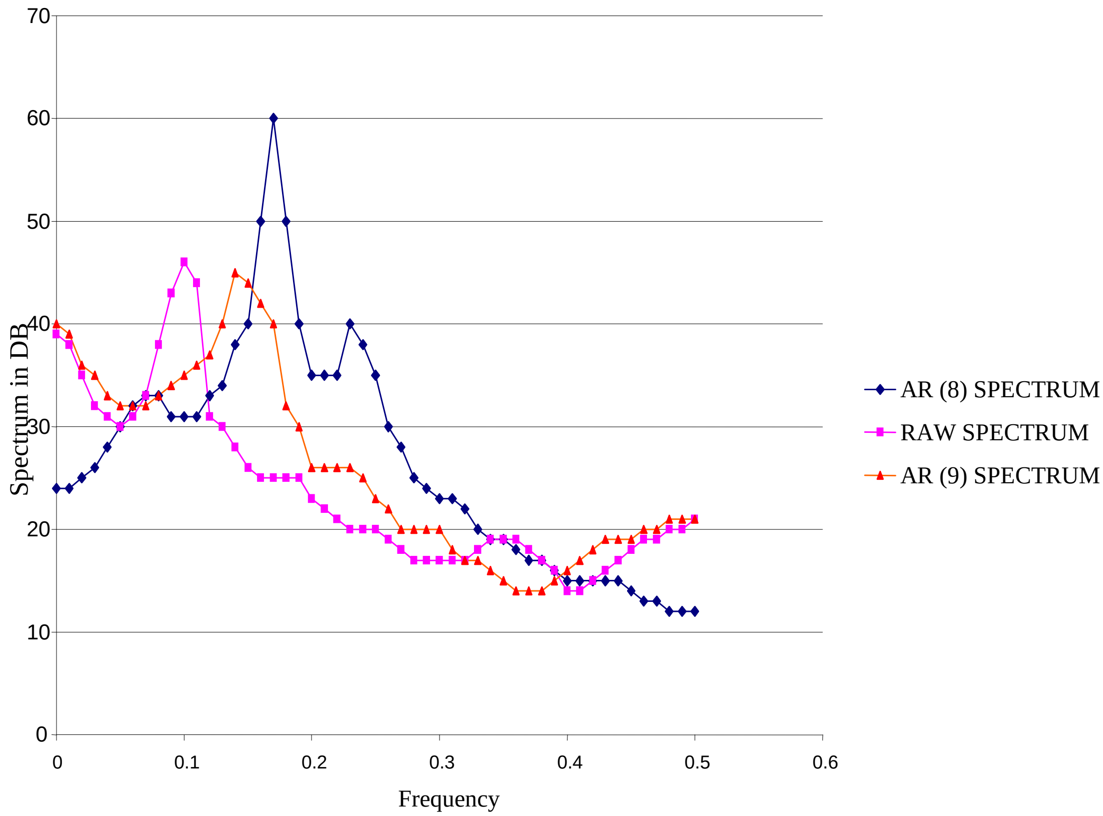

Figure 4.1: Some sunspot series spectra

SIC chooses the AR(3):

$$X_t - 1.248X_{t-1} + 0.452X_{t-2} + 0.153X_{t-3} = \varepsilon_t, \hat{\sigma}^2 = 234.04, R = 48.40 \quad (4.3)$$

BIC chooses the AR(2):

$$X_t - 1.348X_{t-1} + 0.658X_{t-2} = \varepsilon_t, \hat{\sigma}^2 = 239.67, R = 51.61 \quad (4.4)$$

CAT_2 chooses no autoregressive model. IACF suggests an order of 2 and the PACF an order of 3, 8 or 18. The portmanteau test disqualifies (4.1), (4.3) and (4.4) but suggests the adequacy of (4.2). Figure 4.1 is a superimposition of the spectra of (4.1) and (4.2) on the raw spectrum. From this graph we observe a closer agreement of the spectrum of (4.1) to the raw spectrum than that of (4.2). Those of (4.3) and (4.4) are much further from resembling the raw spectrum, [11] This suggests (4.1) as the adequate model. In support [17] and [22] chose an order of 9. Evidence is therefore in favour of FPE, AIC, and CAT_3.

4.2. Series A ([7], pp. 526)

BIC, Φ, SIC, S and FPE4 recommend the model

$$X_t - 0.427X_{t-1} - 0.252X_{t-2} = \varepsilon_t, \hat{\sigma}^2 = 0.1002, R = 100.03 \quad (4.5)$$

On the other hand, AIC, FPE, CAT_2 and CAT_3 pick the model

$$X_t - 0.373X_{t-1} - 0.197X_{t-2} - 0.020X_{t-3} - 0.014X_{t-4} + 0.015X_{t-5} - 0.062X_{t-6} - 0.156X_{t-7} = \varepsilon_t$$

$$\hat{\sigma}^2 = 0.0950, R = 22.44 \quad (4.6)$$

The portmanteau test discredits (4.5) but not (4.6). The IACF and PACF both recommend an order of 7 for the series (see, [12]). Cleveland [10] suggested the order of 7. The spectrum of (4.6) agrees more closely with the raw spectrum than that of (4.5) (see [11]). This implies that (4.6) better fits the data than (4.5).

4.3 Canadian Lynx Numbers (1821-1934) (Campbell and Walker, [9], pp. 430)

We used the well-analysed logarithmic transformation. BIC, Φ and SIC recommend the AR(2):

$$X_t - 1.350X_{t-1} + 0.720X_{t-2} = \varepsilon_t, \hat{\sigma}^2 = 0.3027, R = 40.154 \quad (4.7)$$

The criteria AIC, FPE, CAT_2, CAT_3, FPE4 and S choose the AR(11):

$$X_t - 1.139X_{t-1} + 0.508X_{t-2} - 0.213X_{t-3} + 0.270X_{t-4} - 0.113X_{t-5} + 0.120X_{t-6} - 0.068X_{t-7}$$

$$+ 0.040X_{t-8} - 0.134X_{t-9} - 0.185X_{t-10} + 0.311X_{t-11} = \varepsilon_t, \hat{\sigma}^2 = 0.2263, R = 16.545 \quad (4.8)$$

The IACF recommend an order of 2 whereas PACF recommend 11. Both the R-test and the spectrum test support an order of 11, but not of 2. Etuk [11] found the order 11 adequate. Haggan and Oyetunji [19] and Priestley [21] also chose an AR(11).

5.0 Conclusion.

Etuk [11] has shown that

$$p(BIC) \leq p(SIC) \leq p(\Phi) \leq p(FPE4) \leq p(S) \leq p(FPE) \leq p(AIC)$$

for $N > 44$ where p(BIC) refers to the order p chosen by BIC. Both our Monte Carlo and real series results are consistent with this relationship. Hence, it is advisable to use any of AIC, FPE and CAT_3 in full-order autoregressive modeling with the least squares method. However, other diagnostic checking criteria should also be applied before final model selection.

References.

[1] Akaike, H.(1969). Fitting Autoregressive Models for Prediction. Ann. Inst. Statist. Math., 21, 243 – 247.

[2] Akaike, H. (1970). Statistical Predictor Identification. Ann. Inst. Statist. Math., 22, 203 – 217.

[3] Akaike, H. (1974). A New Look at the Statistical Model Identification . IEEE Trans. Automatic Control AC – 19, 716 – 723.

[4] Akaike, H. (1977). On entropy maximization principle. Proc. Symp. On Applied Statistics (P.R. Krishnaiah, ed.) Amsterdam, North-Holland.

[5] Beamish, N. and Priestley, M. B. (1981). A study of autoregressive and window spectral estimation. Appl. Statist. 30, 41 – 58.
[6] Bhansali, R. J. and Downham, D. Y. (1977). Some properties of the order of an autoregressive model selected by a generalization of Akaike's FPE criterion. Biometrika 64, 547 – 551.
[7] Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis Forecasting and Control. Holden – Day, San Francisco.
[8] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive – integrated moving average time series models. J. Amer. Statist. Ass., 65, 1509 – 1526.
[9] Campbell, M. J. and Walker, A. M. (1977). A Survey of Statistical Work on the Mackenzie River Series of Annual Canadian lynx trappings for the Years 1821 – 1934 and a New Analysis. J. R. Statist. Soc. A 140, 411 – 431.
[10] Cleveland, W. S. (1972). The inverse autocorrelations of a time series and their applications. Technometrics 14, 277 – 297.

[11] Etuk, E. H. (1987). On the selection of Autoregressive-Moving Average Models. Ph.D. Thesis, Dept. of Statistics, Univ. of Ibadan, Nigeria.
[12] Etuk, E. H. (1988). On Autoregressive Model Identification. Jour. Of Official Statistics 4, 113 – 124.
[13] Etuk, E. H. (1998). Comparison of Autoregressive Model Selection Criteria for the Maximum Entropy Approach. Journ. Nig. Stat. Assoc. 12, 75 – 89.
[14] Haggan, V. and Oyetunji, O. B. (1984). On the Selection of Subset Autoregressive Time Series Models. Jour. Time Series Analysis 5, 103 – 113.
[15] Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression. J. R. Statis. Soc. B 41, 190 – 195.
[16] Lutkephol, H. (1985). Comparisons of Criteria for estimating the order of a vector autoregressive process. Jour. Time Series Analysis, 6, 35 – 52.
[17] Morris, J. (1977). Forecasting the Sunspot cycle. J. Roy. Statist. Soc. A 140, 437 – 448.
[18] Oyetunji, O. B. (1983). On the Order determination of Time Series Models. Jour. Nigerian Statist. Ass. 1, 3, 53 – 63.
[19] Pan, W. (2001). Akaike's Information criterion in Generalized Estimation equations. Biometrics 57, 1, 120 – 125.
[20] Parzen, E. (1977). Non-parametric Statistical Data Science: A Unified Approach Based on Density Estimation and testing for "White Noise". Statistical Science Division report No. 47, State University of New York at Buffalo.
[21] Priestley, M. B. (1981).Spectral Analysis and Time Series. Academic Press, London.
[22] Schaerf, M. C. (1964). Estimation of the covariance and regression structure of a stationary time series. Techn. Rep., Dept. of Statistics, Stanford University.
[23] Schwarz, G. (1978). Estimating the Dimension of a model. Annals of Statistics, 6, 461 – 464.
[24] Shibata, R. (1976).Selection of the order of an autoregressive model by Akaike's information criterion. Biometrika 63, 117 – 126.
[25] Shibata, R. (1980). Asymptotically Efficient Selection of the Order of the Model for estimating Parameters of a Linear process. Annals of Statistics, 8, 147 – 164.
[26] Tong, H. (1977). On some alternative approaches to autoregressive order determination. Technical Report 80, Dept. of Maths., University of Manchester Inst. Of Sc. and Tech., U. K.
[27] Waldmeier, M. (1961). The Sun-Spot Activity in the Years 1610 – 1960. Schulthess, Zurich.