Journal of the Nigerian Association of Mathematical Physics
***Volume* 13 (November, 2008), 309 - 314**
© J. of NAMP

The performance of inverse autocorrelation function in model order determination especially with outliers.

N. P. Olewuezi
Department of Mathematics and Computer Science
Federal University of Technology Owerri, Nigeria

Abstract

Noting that the effect of outliers in model order determination could be serious, this paper is concerned with the problem of examining the performance of inverse autocorrelation function in model order determination especially with outliers. Tsay* (1986*) iterative procedure was used to identify the outliers, remove their effects and then specify a tentative model using the inverse autocorrelation function. An example is also presented.

***Keywords*:** Inverse autocorrelation function; outliers; model order determination.

1.0 Introduction

The autocovariance function of a stationary time series $\{X_t\}$of zero mean is defined by $\gamma_k = E(X_t, X_{t-k})$ and the autocorrelation function by $\rho_k = \frac{\gamma_k}{\gamma_0}$

The autocovariance generating function of $\{X_t\}$ is defined by

$$\gamma(z) = \sum_{k-\infty}^{\infty} \gamma_k z^k \tag{1.1}$$

The inverse autocovariances of $\{X_t\}$ are defined in such a way that their generating function is the reciprocal of equation (1). Thus the inverse autocovariance generating function $\gamma_i(z)$ is defined by $\gamma(z)\gamma_i(z) = 1$ and $\gamma_i(k)$ is the inverse autocovariance function at lag k which is the coefficient of Z^k in the expansion of $\gamma_i(z)$ in positive and negative powers of Z.

The inverse autocorrelation function $\rho_i(k)$ is then defined by

$$\rho_i(k) = \frac{\gamma_i(k)}{\gamma_i(0)} \tag{1.2}$$

Cleveland (1972) introduced the concept of the inverse autocorrelation function (IACF). He defined inverse autocovariances as the autocovariances associated with the inverse of the spectral density of the series which Pazen (1974) called the inverse spectral density. That is, let for the discrete stationary process $\{X_t\}$, $fi(\omega) = \{f(\omega)\}^{-1}$ be integrable on the interval (0, 1). The inverse autocovariances of $\{X_t\}$ are defined by $\gamma_i(k) = \int_{-\pi}^{\pi} e^{jk\omega} f_i(\omega) d\omega$ and $\gamma_i(k) = \gamma_i(-k), \Lambda$, k = 0, 1, 2, … .

e-mail:ngolewe@yahoo.com, Phone: 08034933133

Let us consider the following ARMA (*p,q*)

$$X_t - \sum_{j=1}^{p} \alpha_j \; X_{t-j} = \varepsilon_t + \sum_{l=1}^{q} B_l \varepsilon_{t-l} \qquad (1.3)$$

where $\{\varepsilon_t\}$ is a white wise process with variance σ_ε^2. Neglecting the term $\sigma_\varepsilon^2 / 2 \wedge$ in the definition of ARMA process, the spectral density of equation (1.3) is

$$f(\omega) = \frac{\left|\beta\left(e^{-i\omega}\right)\right|^2}{\left|\alpha\left(e^{-j\omega}\right)\right|^2}$$

where $\beta(z) = 1 + \beta_1 z + \beta_2 z^2 + \ldots + \beta_\rho z^\rho$, $\alpha(z) = 1 + \alpha_1 z + \alpha_2 z^2 + \ldots + \alpha_q z^q$

Denoting the autocorrelations and inverse autocorrelations of equation (1.3) by $\rho(k:\alpha,\beta)$ and $\rho_i(k:\alpha,\beta)$,

respectively we have $$\rho_i(k:\alpha,\beta) = \frac{\int_{-\pi}^{\pi} e^{jk\omega} f_i(\omega) d\omega}{\int_{-\pi}^{\pi} f_i(\omega) d\omega} = \frac{\int_{-\pi}^{\pi} e^{jk\omega} \frac{\left|\alpha\left(e^{-j\omega}\right)\right|^2}{\left|\beta\left(e^{-j\omega}\right)\right|^2} d\omega}{\int_{-\pi}^{\pi} \frac{\left|\alpha\left(e^{-j\omega}\right)\right|^2}{\left|\beta\left(e^{-j\omega}\right)\right|^2} d\omega} = \rho(k:\beta,\alpha).$$

Hence, the IACF of an ARMA (*p*,*q*) is the autocorrelation function of the inverse ARMA (*q*,*p*) (that is, with the AR and MA operators interchanged). Therefore, for an AR (p) process, the IACF is identical to the autocorrelation function of the MA(p) with the same parameters in the same order. That is, for a pth order MA process, $\rho_k = 0$ for $k > p$ while for a *qth* order autoregressive process $\rho_i(k) = 0$ where $k > q$.

In practice, the autocorrelations and inverse autocorrelations of an observed time series will not be known exactly but must be estimated from the data. For AR(*p*) models, the IACF cuts off at lag p. This property makes it complete with ACF as a tool for determination of autoregression order.

2.0 Estimates of inverse autocorrelations

As suggested by Cleveland (1972), the two methods of estimating IACF Sten from either the autoregressive method or the window method. The first method of estimating the spectral density function is to fit autoregressive model using a high enough order to give a good fit. The problem with this method is that we have to impose a model on the series. The second method which involves the window estimate involves smoothing the periodogram, *I*(w) given by

$$I(\omega) = \frac{1}{2\pi N} \left| \sum_{i=1}^{N} \left(X_t - \overline{X}\right) e^{j\omega t} \right|^2 = \frac{1}{2\pi} \left[C(0) + 2 \sum_{k=1}^{N-1} c(k) Cos\,\omega k \right]$$

where $c(k)$ denotes the sample estimate of the autocovariance of lag *k*. Although the periodogram, $I(\omega)$ is asymptotically unbiased for the spectral density function $f(\omega)$, its variance does not decrease as *N* increases. There is need to smooth the periodogram $I(\omega)$ by applying some weighting function to $I(\omega)$. There are a number of weight functions usually referred to as windows that are commonly used.

Using the Daniell weight function we estimate the inverse autocorrelations as follows:

Let $\omega_l = \frac{2\pi_t}{N}, l = 0,1,.\Lambda\ , N-1$. Calculating $I(\omega_l) = \frac{1}{\pi}\left[C(0) + 2\sum_{k=0}^{N-1} c(k) Cos\,\omega_{lk}\right]$ and obtaining an estimate of the spectral density by $\hat{f}(\omega_l) = \frac{1}{2p+1}\sum_{k=-p}^{p} I(\omega_{l+k})$ where p is a suitably chosen positive integer. Inverse autocovariances are then estimated by $c_i(k) = \sum_{l=0}^{N-1}\left[\frac{e^{jk\omega_l}}{\hat{f}(\omega_l)}\right]$ and estimates of

inverse autocorrelations are then obtained by $r_i(k) = {c_i(k)}/{c_i(0)}$.

The problem with this method is that there is no accurate way of choosing p. The choice of p has been purely a subjective process most commonly done by plotting the smoothed periodogram for different values of p and choosing that p which gives the smoothest picture, without losing any characteristic feature of the spectrum. The subjectivity introduced in choosing p poses problems. Hipel et al (1977) have suggested trying four values of p between 10 and 40 $\left(\text{where } p \le {n}/{4}\right)$ and choosing the value p which gives the most representative graph of the resultant inverse autocorrelation estimate $r_i(k)$ against lag k. Chatfield (1979) advices against the use of automatic criteria like the AIC, BIC, etc in the determination of p as interest is not in optimal parametric parsimony. He however warns that p must not be so low not to show the form of the sample IACF, nor so large as to make the variance of the estimates too high; the choice of p must be such that for high lags the estimates of the inverse autocorrelations approach zero. He suggests some form of trial and error until the foregoing criteria are met. Hosking (1980) suggests that p should vary with the sample size n.

3.0 Order determination

Stationary Autoregressive (AR) processes have theoretical autocorrelation functions that decay toward zero rather than cut off to zero. The autocorrelation coefficients may alternate in sign frequently, or show a wavelike pattern, but in all cases they tail off toward zero. By contrast AR processes have theoretical partial autocorrelation functions that cut off to zero after lag p, the AR order of the process.

The IACF of a time series are useful at the identification stage of model building. In practice, this quantity must be estimated from the data. The IACF of an AR process cuts off at lag p. It turns out that it has similar properties to the Partial Autocorrelation function (PACF) in that it cuts off at lag p. The theoretical ACF of a MA(*q)* process has a very simple form in that it cuts off at lag q and so the analysts should look for the lag beyond which the values of r_k are close to zero. However, their PACF's tail off toward zero and hence the IACF.

Mixed processes have theoretical autocorrelation functions with both AR and MA characteristics. The ACF tails off towards zero after the first q-p lags with either exponential decay or a damped Sine wave. The PACF and hence the IACF tails off to zero after the first $p - q$ lags. In practice, p and q are usually not larger than two in a mixed model for nonseasonal data.

4.0 Model estimation

After identifying a tentative model, the next step is to estimate the parameters in the model. For an AR (*p*), with *p* estimated, the unknown parameters, $\mu, \phi_1, \phi_2, \ldots, \phi_p$ can be estimated. The classical approach to the problem of estimation for the linear regression case is by the least square method in which the residual sum of squares is minimized with respect to the parameters to be estimated. An approximate least squares method is the Yule-Walker method which involves the recursive Levinson – Durbin Algorithm. The algorithm gives an iterative method of obtaining AR coefficients without going through the inversion routine.

$$\hat{\phi}_{k+1,i} = \hat{\phi}_{k,i} - \hat{\phi}_{k+1}\hat{\phi}_{k,k+1-i}; i = 1,2,\ldots,p$$

$$\hat{\phi}_{k+1} = \hat{\phi}_{k+1,k+1}\frac{1}{\sigma_k^2}\left[C_{k+1} - \sum_{i=1}^{k}\hat{\phi}_{k,i}C_{k+1-i}\right]$$

and $$\sigma_{k+1}^2 = \sigma_k^2\left(1-\phi_{k+1}^2\right)$$

where $\hat{\phi}_{k,i}$ is the estimate of $\phi_{k,i}$, the ith coefficient of an AR model and $\{c_k\}$ is the sample autocovariance function. Parameter estimation for MA and ARMA models is a complex procedure. Estimation problems are more difficult for an MA process than an AR process, because efficient explicit estimators cannot be found. Instead, some form of numerical iteration must be performed. The estimation problems for an ARMA model are similar to those for a MA model in that an iterative procedure has to be used. The residual sum of squares can be calculated at every point in a suitable grid of the parameter values and the values which give the minimum sum of squares may then be assessed. Alternatively, some sort of optimization procedure may be used.

5.0 The effects of outliers

Time series data often contain outliers which have an effect on parameter estimates and more importantly lead to inaccurate forecasts. Outliers in time series depending on their nature may have a moderate to significant impact on the effectiveness of the standard methodology for time series analysis with respect to model identification, estimation and forecasting and also have drastic effects on estimates for such quantities as correlation coefficients, regression coefficients and spectral density estimates.

Outliers can take several forms in time series. Fox (1972) proposed the formal definitions and a classification of time series outliers to two types – additive and innovational outliers. For a properly deduced stationary process, let X_t be the observed series and Z_t be the outlier–free (OF) series. Consider a familiar time series model

$$\Pi(B)z_t = a_t$$
$$\Pi(B) = 1-\Pi_1 B - \Pi_2 B^2 \Lambda ,$$

$\{a_t\}$ is a sequence of identically, independently distributed random variables with zero mean and variance σ^2.

The function $\Pi(B)$ is often expressed as a ratio of $\phi(B)/\theta(B)$ where $\phi(B) = 1-\phi_1 B - ... - \phi_p B^p$ and $\theta(B) = 1-\theta_1 B - ... - \theta_q B^q$ are stationary and invertible operators sharing no common factors.

The models commonly employed on the OF time series Z_t are the Additive Outlier (AO) and innovational outlier (IO) which are defined respectively of a single outlier for a simple case as

$$X_t = Z_t + D\xi_t^{(T)} \tag{5.1}$$

and $$X_t = Z_t + \left(\frac{\theta(B)}{\phi(B)}\right)D\xi_t^{(T)} \tag{5.2}$$

where X_t is the observed series, D is the magnitude of the outlier and $D\xi_t^{(T)} = 1$,

If $t = T$ and O otherwise, which is the time indicator signifying the time occurrence of the outlier.

The *AO* affects the level of the Tth observation whereas an *IO* affects all observations X_T, X_{T+1},…, beyond time through the memory of the system described by $\frac{\theta(B)}{\phi(B)}$.

The *AO* can be regarded as a gross error model. In general, the presence of more than one outlier of various types in a model, is specified by

$$X_t = \sum_{k=1}^{n} W_k(B) D_k \xi_t^{(T)} + Z_t$$

where
$$Z_t = \theta(B)\phi^{-1}(B)a_t$$
$$W_k(B) = 1$$

for AO model

$$W_k(B) = \theta(B) \big/ \phi(B)$$

for IO model at time $t = T_k$ and n is the number of outliers.

Outliers affect the autocorrelation structure of a time series and may also bias the autocorrelation function (ACF), PACF and the IACF. These biases can be severe and they depend on, besides the obvious attributes like the number, type, magnitude and position of the outliers, also the underlying model and its autocorrelation. Deutsch et al (1990), have some results on the effects of outliers on ARMA model identification. They observed that in series of short to moderate length, often the presence of a single outlier will result in a true AR model being falsely identified as an MA or an ARMA model, and the identified lengths (p and q) will also be wrong.

Masarotto (1987) presents a method for robust estimation of the ACF and the PACF.

Research on the effects of outliers has been the most thorough in regression models. The effects of outliers on estimation are well known, and there also exist several methods of detecting these effects. Cook and Weisberg (1982), Barnett and Lewis (1994) examined these effects.

A common approach to deal with outliers in a time series is to identify the locations and the types of outliers and then use intervention models discussed in Box and Tiao (1975) to accommodate the outlier effects. This approach requires iterations between stages of outlier detection and estimation of an intervention model.

6.0 Empirical illustration

Following Tsay (1986), the proposed method uses only the least squares method, that is, the linear regression techniques, to obtain parameter estimates. To examine the usefulness of IACF in model order determination especially with outliers, an illustration was carried out. The data considered was the first word – Gessel adaptive score [Mickey et al (1967)]. Olewuezi (2007) proposed an algorithm which was used for the computation of inverse autocovariance function (IACVF). An outlier input series was assumed and to observe the timings of the outliers, Tsay (1986) detection technique was followed. The fitted model estimates for the series was shown in Table I which gives the corresponding summary statistics for the series.

The models fitted were

$$Y_t = (1 - 0.3199B)^{-1}(1 - 0.7211B)X_t + (1 + 0.552B)a_t$$

with $\sigma_{at}^2 = 15.19$ for the outlier free series

and

$$Y_t = (1 - 0.1694B)^{-1}(1 - 0.8901B)X_t + (1 + 0.3211B)a_t$$

with $\sigma_{at}^2 = 23.62$ for the outlier contaminated series.

The model residual variance with the outlier contaminated is about 35.69% multiple of that with the outlier free series. The test criterion fails to suggest any model discrepancy.

Table 6.1: Model estimate of the series

		X	***Y***	**Error model (*a_t*)**	**Diagnostic Checking**
Outlier – Contaminated Series	AR estimate	0.8901	0.1694	-0.3211	
	Standard Error	0.2621	0.0112	0.0225	
	Model Residual Variance	102.6	261.52	23.62	0.000013
	Timing of Outliers	1,2,4,5,&20	1,2,3,5,19,21		
Outlier-Free	**Number of Outliers**	**5**	**6**		

AR estimate	0.7211	0.3199	0.552	
Standard Error	0.1816	0.0024	0.0170	
Model Residual Variance	64.36	103.25	15.19	0.000001

7.0 Summary and conclusion

The performance of inverse autocorrelation function in model order determination especially with outliers was discussed. The presence of outlying observations or structural changes raises the question of efficiency and adequacy in fitting models to time series data.

From the results, it was clear that the model residual variance was further reduced to 35.69%. Thus, the model selection based on the empirical inverse autocorrelation function can be misleading in the presence of outliers and tends to underestimate model order, thereby resulting in an inappropriate model. Even if the model is appropriately specified, outliers may still produce bias in parameter estimates and hence may seriously affect the efficiency of outlier detection.

Finally, as a suggestion, a comparison is to be made between autocorrelation function and inverse autocorrelation function in model order determination. The types and locations of outliers at different iterations of model estimation may also be considered.

References

[1] V. Barnet and T. Lewis (1994): Outliers in Statistical data. Third edition. John Wiley, Chichester.

G.E.P. Box and G.C. Tiao (1975). Intervention analysis with applications to economic and environmental problems. J. Amer. Sta. Ass. 70, 70 – 79

[2] C. Chatfield (1979). Inverse autocorrelations. J. R. Stat. Soc. A. 142, Part 3, pp 363 – 377

[3] W.S. Cleveland (1972). The inverse autocorrelations of a time series and their applications Technometrics, 14, 277 – 298.

[4] R.D. Cook and S. Weisberg (1982). Residuals and influence in Regression. Chapman and Hall, New York.

[5] S. J. Deutsch et al (1990). Effects of single outlier on ARMA Identification. Communications in statistics, theory and methods, 19, 2207 – 2227.

[6] A. J. Fox (1972). Outliers in Time Series. J. Roy. Stat. Soc. B. 43, 350 – 363.

[7] K.W. Hipel et al (1977). Advances in Box – Jenkins modeling. I. model construction water resource. 13, 567 – 75.

[8] J.R.M. Hosking (1993). Reallocation outliers in time series. Applied Statistics, 42, 301 – 313.

[9] G. Masarotto (1987). Robust identification of ARMA models. Applied Statistics 36, 214 – 220

[10] M.R. Mickey et al (1967). Note on use of stepwise regression in detecting outliers compt. Biomed. Res. 1, 105-111.

[11] N.P. Olewuezi (2007). Some contributions of autocorrelation function and its inverse in the detection of outliers. Ph.D. thesis, Dept. of Statistics, University of Ibadan, Nigeria.

[12] E. Parzen (1974). Some recent advances in time series modeling I.E.E.E. Trans. Aut. Control 723-741.

[13] R.S. Tsay (1986). Time series model specification in the presence of outliers. Jour. Of the Amer. Stat. Assoc. 81, 132-141.