Research Journal of Applied Sciences 2 (5): 574-578, 2007

Relationship Between an Innovative Outlier Model and the Intervention Effects Model

[1]I.U. Moffat and [2]E.H. Etuk
[1]Department of Mathematics, Statistics and Computer Science, University of Uyo, Uyo, Nigeria
[2]Department of Mathematics and Computer Science,
River State University of Science and Technology, Port-Harcourt, Nigeria

Abstract: We present a relationship between an alternative estimation procedure of ARMA models based on an innovative outlier framework and the Kalman Filter Smoothers (KFS) estimation of intervention effects for outliers generated by special events. The intervention model describes how the events manifest themselves in the observations. The results are asymptotically equivalent to the estimation of ARMA parameters based on a generalized t-distribution for the residuals with a mixture interpretation relying on the innovative outlier framework. The joint estimation of the estimators, if considered, supports the normality test of the residuals.

Key words: ARMA models, estimation, interventions, kalman filters, regression, state space representation, outliers

INTRODUCTION

As many time series are subject to external influences, they are often characterized by their autocorrelation. In our case, the series that are suspected or known to be under the influence of particular exogenous processes, which could be sudden or unexpected, is of interest. They may take the form of natural disasters, strikes, wars, or an introduction of a new set of rules. Here, the process under consideration is termed intervention and are, by nature hard to characterize in terms of quantities that can be measured. They are often modeled by the introduction of dummy regression variables, which represent the influence that an unquantifiable, external event has on the series.

In time series processes, anomalous observations sometimes occur in patches. In order to account for this and to detect stretches of over influential observation, Bruce and Martin (1989) put forward leave-k-out diagnostics. The parameters of the model fitted to the full data set are compared with those generated by fitting the model to the data when a stretch of k points are taken to be missing. This approach is somewhat limited. There are many types of intervention, such as level shifts and slope changes, which do not fit into the leave-k-out approach. The method also requires a great deal of computationally intensive parameter re-estimation. These problems are, to a certain extent, addressed by Atkinson *et al.* (1995) who use score statistics to approximate the change in hyperparameters when interventions are introduced to a series.

ARMA (Autoregressive Moving Average) models are one of the most widely used approach for stationary time series. Applications of these techniques abound in economics, business and science. The traditional approach to ARMA models has implicitly assumed that the underlying residuals are normally distributed; consequently, minimizing the sum of squares of asymptotically equivalent to maximum likelihood estimators.

A common criticism of the technique of least squares is that it places too much emphasis on outliers. The distributions of the underlying residuals in regression and time series models is unknown and may be characterized by a higher incidence of outlier than is consistent with a normal distribution. These might be due to either thick tailed or contaminated distributions. If the residuals are not normally distributed, estimates based upon a least squares criterion will be neither efficient nor maximum likelihood and may be very sensitive to the occurrence of outliers.

This limitation of the least squares methodology has lead to several alternative approaches. One approach is to specify a more general distribution for the residuals, which may include thick tailed distribution and consider the corresponding maximum likelihood estimators. It should be emphasized that the actual density encouraged in practice need not coincide with the

Corresponding Author: I.U. Moffat, Department of Mathematics, Statistics and Computer Science, University of Uyo, Uyo, Nigeria

assumed distribution in Maximum Likelihood (MLE) estimation. Consequently, the maximum likelihood estimators are sometimes referred to as Quasi Maximum Likelihood (QML) estimators. Robust estimation techniques have provided another approach to estimation, which do not depend upon a particular distribution. However, in each case the selection of the hypothesized density robust of the actual distribution of the unknown and unobserved residuals.

There is a considerable body of literature demonstrating that the properties of least squares estimators in regression models are sensitive to the nature of the random disturbances. Several Monte Carlo studies substantiate the sensitivity of least squares estimators obtained by minimizing the sum absolute deviations (LAD) if the error terms are normally distributed, but this may be reversed if the error terms are generated from a thick tailed distribution such as the double exponential or Cauchy, Oveson (1970), Smith and Hall (1972), Kadiyala and Murthy (1977), Courtsey and Nyquist (1983). Andrews *et al.* (1972) report the result of an extensive analysis of many alternative robust estimators of location.

Much less is known about the properties of alternative estimator of parameter in ARMA models. Martins (1981), Martiin and Yohai (1985a, b) provide excellent distribution of robust estimations techniques in time series models and include many useful references. Bustos and Yohai (1986) proposed two robust ARMA estimators based on residual autocovariances.

Martins (1981) considered robust estimator known as the M and General M (GM) estimator for times series models. The M estimator can be viewed as a generalization of maximum likelihood estimation and was introduced as a estimator for location and regression type problem by Huber (1964, 1973). The M estimators are not qualitatively robust (Martin and Yohai, 1985).

It is well established that the presence of outliers can have a dominating and deleterious effect on standard location model estimators such as the least squares regression, the least absolute deviation regression, or even the generalized M-estimators (Maronna *et al*., 1979; Donoho and Huber, 1983; He, 1991). Furthermore, Sakata and White (1995) show that Quasi-Maximum Likelihood regression (QML) estimators are in general vulnerable to outliers.

The existence of outliers may not be helpful in predicting future returns. However, they may unduly influence the estimation and forecasting of financial time series. Balke and Fomby (1994) and Van Dijk *et al.* (1999) find that neglecting outliers can erroneously suggest misspecification or inadequate modeling.

PARAMETRIC RELATIONSHIP OF THE MODELS

Here, we display the relationship between the model parameters of the innovative outliers and that of simple interventions. We consider the ARMA (p, q) model defined by

$$\phi(B)Y_t = \theta(B)\epsilon_t, \qquad (1)$$

Where $\phi(B) = 1 - \phi_1 B - \phi_2 B - \ldots - \phi_p B$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B - \ldots - \theta_q B$ represent assumed to have a modulus greater than one. Furthermore, the residuals (ϵ_t) satisfy $E(\epsilon_t) = 0$, $var(\epsilon_t) = \sigma$, $corr(\epsilon_t, \epsilon_s) = 0$, $t \neq s$. A consequence of this specification and assumption is that (1) can be equivalently represented as

$$Y_t = \phi(B)\,\theta(B)\,\epsilon_t, \qquad (2)$$

$$\epsilon_t = \theta^{-1}(B)\,\phi(B)Y_t. \qquad (3)$$

The residuals in (1) will be assumed to be distributed as a Generalized T statistic (GT) defined by

$$GT(\epsilon\,;\beta, p, q) = \frac{p}{2\beta q^{1/p} B(1/p, q)(1+|\epsilon|^p/\beta^p q)^{q+1/p}} \quad -\infty < \epsilon < \infty, \qquad (4)$$

Equation 4 defines a very rich family of density, which includes the power exponential, or Box-Tiao (BT)

$$BT(\epsilon;\beta,p) = \frac{pe^{-(|\epsilon|/\beta)^p}}{2\beta\,\Gamma(1/p)}, \infty < \epsilon\ \infty \qquad (5)$$

As a limiting case as $q \to \infty$. The GT distribution is symmetric about the origin and is flexible enough to include the t density ($p = 2$, $\beta = \sigma\sqrt{2}$), the normal ($p = 2$, $\beta = \sigma\sqrt{2}$, $q \to \infty$), double exponential ($p = 1$, $q \to \infty$) and the Cauchy ($\beta = \sigma\sqrt{2}$, $p=2$, $q=1/2$) as special cases. Consequently, it can include thick or thin tailed distributions. The GT can be shown to arise as a result of mixing a BT variate with a stochastic scale parameter, which is distributed as an inverse generalized gamma (Mc-Donald and Butler, 1987). This is a form of an innovative outlier model, which can be associated with thick tails.

In the state space form, an innovative outlier model is obtained by setting $X_i = 0$ and $W_i = H$ in the general specifications

$$y_t = Xt\,\delta + Zt\,\alpha\,t + G_t\,\epsilon_t \qquad (6)$$

$$\alpha_{t+1} = W_t\delta + Tt\,\alpha_t + H_t\epsilon_t \quad (7)$$

yielding $D(i) = Z\,T^{t-I}H = \Sigma\,\pi_{t-i-j}\,\theta_j$, on substitution in the equation

$$D_t(i) = \begin{cases} 0, & t=1,...,I-1 \\ X_i, & t=I, \\ Z_t\ T_{t-1,\,I+1}\,W_t, & t=I+1,...,n \end{cases} \quad (8)$$

Where X_t and W_t model the special events or shock at time t and are referred to as the shock design, D(i) is the intervention shape such that for $k \ge 0$, π_k is the coefficient of B^k in the expansion of $1/\phi(B)$ and $\pi_k = 0$ for $k< 0$. Thus, the signature $D_t(i)$ is made up of coefficients in the polynomial expansion of $\theta(B)/\phi(B)$, that is, the coefficients in the infinite MA representation.

To treat an ARMA model in the state space form, Tsay (1988) considers models of the form

$$y_t = \frac{\phi(B)}{\theta(B)}\xi_t + \delta\omega(B)\xi_t(i) \quad (9)$$

Where $\xi_t(i)$ is an indicator function which takes the value 1 when t=i and is 0 otherwise. The ratio $\theta(B)/\phi(B)$ defines the null model while $\omega(B)\,\xi_t(i)$ defines the signature and δ is the scale parameter. Tsay defines additive outliers, for which $\omega(B) = 1$ and innovative outliers, for which $\omega(B) = \theta(B)/\phi(B)$. These are equivalent to our measurement and state disturbance shocks, respectively. Thus our intervention effects parameter is equivalent to the scale parameter in the innovative outlier model (i.e., $\delta = \beta$).

A DIAGNOSTIC STATISTIC FOR INTERVENTION EFFECTS

To estimate δ having the same magnitude as β, Tsay rewrite the intervention model as

$$\frac{\phi(B)}{\theta(B)}y_t\ \frac{\delta\,\phi(B)}{\theta(B)}\omega(B)\xi_t(i)+\varepsilon_t \quad (10)$$

where the left hand side defines v_t the infinite sample null model innovations. Letting $\lambda_t(i) = \{\phi(B)/\theta(B)\}\,\omega(B)\xi_t(i)$, a quantity which corresponds to the filtered special event signature, yields

$$\tilde{v}_t = \delta\,\lambda_t(i) + \varepsilon_t, \quad (11)$$

Tsay (1988) regresses the v_t's on $\lambda_t(i)$ for each origin i. The values of $\lambda_t(i)$ are determined by the choice of $\omega(B)$. In the innovational outlier case, $\omega(B) = \theta(B)/\phi(B)$ implying that $\lambda_t(i) = \xi_t(i)$. The Generalized Least Squares (GLS) estimate of he intervention parameter δ_i is used to devise statistics to measure the significance of an intervention at t = i. The null model defines the covariance matrix $r^2\Sigma = \text{cov}(y)$. Standard GLS arguments yield.

$$\hat{\delta}_i = \{D\hat{\ }(i)'\Sigma^{-1}D(i)\}^{-1}D(i)'\Sigma^{-1}\Im \quad (12)$$
$$\text{cov}(\delta_i) = \sigma^2\{D(i)'\Sigma^{-1}D(i)\}^{-1}$$

(i) = cov $\{D(i)'\Sigma^{-1}y\}$ Eq. 12 simplifies to

$$\hat{\delta} = S_i^{-1}{}_{Si},\ \text{cov}(\hat{\delta}_i)^2 = \sigma\,S_i^{-1} \quad (13)$$

Where

$$s_i = D(i)'\Sigma^{-1}y,\ S_i = D(i)\Sigma^{-1}D(i) = \sigma^{-2}\ \text{cov}(s_i) \quad (14)$$

The s_i are referred to as the intervention contrasts which are sufficient statistics for estimating the intervention effects. They are the analogue of treatment contrasts for estimating treatment effects in the standard design of experiments. If the data set is uncorrelated, Σ is a diagonal matrix, allowing Σ^{-1} and thus δ_i to be calculated directly from (12). This is equivalent to the usual approach regressing the data y on the intervention effect D(i). In the extreme case where $\Sigma = I$ and D(i) is 0 everywhere except in a single position where it is one the estimate of δ is the observation y_i.

The hypothesis of no special event, $\delta = 0$, is tested with the usual statistic,

$$\hat{\delta}_i\{\text{cov}(\hat{\delta}_i)\}^{-1}\delta_i \qquad = \sigma^{-2}(s'_i S^{-1}s_i) \quad (15)$$

In practice, $\hat{\sigma}^2$ is replaced by the normal biased Maximum Likelihood Estimate (MLE). The MLE of σ^2 assuming $\hat{\delta} = 0$ is given by $\hat{\sigma}^2 = (y'\Sigma^{-1}y)/n$. This estimate is adjusted to give the MLE under the hypothesis that $\delta \ne 0$. Assuming that all other hyper parameter estimates remain fixed; $\hat{\sigma}_i^{-2} = \hat{\sigma}^{-2}\ -n^{-1}s'_i\ S_i^{-1}s_i$ substituting this adjusted estimate into (15) yields the test statistic

$$\tau_i^2 = \hat{\sigma}_i^{-2}\,s'_I S_i^{-2}s_i = \{\hat{\sigma}(si\ S'_i s'_i)^{-1} - n^{-1}\}^{-1} \quad (16)$$

Which has an approximate χ^2_p distribution where p is the rank of S_i. The estimate of the standard error of each component of $\hat{\delta}_i$ is given by the square root of the appropriate diagonal entry in $\sigma^{-2}S_i^{-1}$. Dividing each component by its estimated standard error gives a statistic, which is analogous to the regression t-statistic.

The actual implementation of the methods requires specification of cov (y) = $\sigma^2\Sigma$ and the intervention signature D(i). These are modeled using a state space description of the data y. An efficient method for computing the statistic τ_i^2 or the t-statistic is also crucial. Defining the smoothation vector u = Σ^{-1} y Eq. 14 yields s_i = D(i)'u.

Thus, all of the interventions statistics τ_i^2 can be computed from smoother output. Our method is direct papers and, in addition to computational benefits, provides insights into the nature of intervention process. It allows the simultaneous testing of a wide variety of outlier effects.

Established approaches (Tsay, 1986; Atkinson *et al*., 1997) to computing (14) and estimating δ involve transforming the observations and the signature matrix. By applying the appropriate linear operations, one data can be transformed to the innovations, which are uncorrelated. Applying the same transformation to the regression variables D(i) allows us to generate s_i and S_i and thus calculate the estimate δ and its covariance matrix. Using the lower triangular matrix L to denote the filtering transformation into 0 = Ly where $\nu = (\nu' \ldots \nu'_n)'$ is the vector of Kalman filter innovations that $\Sigma^{-1} = L'F^{-1}L$ so s_i = (LD (i)'$F^{-1}\nu$ where F = σ^{-2}cov(ν) = diag $\{F_1 \ldots F_n\}$. The entries in LD(i) can be calculated by applying the filter to the column of the regression variable D(i). Using this approach, for each origin i and each type of intervention the columns of D(i) must be filtered.

CONCLUSION

Thus the introduction of special events is a powerful tool for analyzing departures from a fitted null model. Many forms of aberrant behavior can be modeled efficiently by special events to the transition equation of a state space representation. We establish that the scale parameter associated with our innovative outlier model is empirically equivalent to the intervention effects estimate. Thus the characteristic features of the generalized t-distribution and that of a state space model are transformable as any number of interventions, can be generated using the output of a single null model KFS run. The theoretical methods in this research can be extended to several other models.

REFERENCES

Andrews, D.F., P.J. Bickel, F.R. Hampel, P.J Huber, W.H. Rogers and J.W. Tukey, 1972. Robust estimates of location (Princeton University Press).

Atkinson, A.C., S.J. Koopman and N. Shephard, 1997. Detecting shocks: Outliers and breaks in time series, J. Econ., 80: 387-422.

Balke, N.S. and T.B. Fomby, 1994. Large shocks, small shocks and economic fluctuations: Outliers in macroeconomic time series. J. Applied Econ., 9: 181-200.

Bruce, A.G. and R.D. Martin, 1989. Leave-k-out diagnostics for time series, J. Royal Stat. Soc. Series, B, 51: 363-424.

Bustos, O.H. and V.J. Yohai, 1986. Robust Estimators for ARMA models. J. Am. Stat. Assoc., 81: 155-168.

Courtesy, D. and N. Nyquist, 1983. On Least Absolute Error Estimation of linear Regression Models with Dependent Residuals. Rev. Econ. Stat., 65: 687-692.

Donoho, D.L. and P.J. Huber, 1983. The Notion of Breakdown Point. In: Bickel, P.J., K.A. Doksum and Jr. J.L. Hodges (Eds.), A Festschrift for Erich L. Lehmann in Honor of his 65th Birthday, Belmont, CA: Wadsworth International Group, pp: 157-184.

He, X.M., 1991. A local breakdown property of robust tests in linear regression. J. Multivariate Anal., 38: 294-305.

Huber, P.J. 1964. Robust Estimation of a Location Parameter. Ann. Mathe. Stat., 35: 73-101.

Huber, P.J., 1973. Robust Regression: Asymptotics, conjectures and Monte Carlo. Ann. Stat., 5: 799-821.

Kadiyala, K.R. and K.S.R. Marshy, 1977. Estimation of Regression Equations with Raunchy Disturbances. Can. J. Stat. Section C: Applications, 5: 111-120.

Maronna, R.A., O.H. Bustos and V.J. Yohai 1979. Bias and Efficiency Robustness of General M-Estimators for Regression with Random Carriers. In Smoothing Techniques for Curve Estimation, T. Gasser and M. Rosenblatt (Eds.), Lecture Notes in Mathematics 757. Berlin: Springer, pp: 91-116.

Martin, R.D. and V.J. Yohai, 1985. Influence Functionals for Time Series," Ann. Stat., 35: 781-818.

Martin, R.D. and V.J. Yohai, 1985. Robustness in Time Series and Estimating Arma Models. Handbook of Statistics. Hannan, E.J., P.R. Krishniah and M.M. Rao (Eds.), 5, Elsevier Science Publishers, pp: 119-155.

Martins, D.R., 1981. Robust Methods for Time Series. Applied Time Series Analysis II, D F. Findley (Eds.) (New York: Academic Press), pp: 683-759.

McDonald, J.B. and R.J. Butler, 1987. Some Generalized Mixture Distributions with an Application to Unemployment Duration. Rev. Econ. Stat., 69: 232-240.

Oveson, Richard, 1970. Regression Parameter Estimation by Minimizing Absolute Deviations. Harvard University, Un-Published Doctoral Dissertation. point with applications to regression model estimation. J. Am. Stat. Assoc. 90: 1099-1106.

Sakata, S. and H. White, 1995. An alternative definition of finite sample breakdown.

Smith, V.K. and T.H. Hall, 1972. A Comparison of Maximum Likelihood Versus BLUE Estimators. Rev. Econ. Stat., pp: 186-190.

Tsay, R.S., 1986. Time series model speci-cation in the presence of outliers, J. Am. Stat. Assoc., 81: 132-141.

Tsay, R.S., 1988. Outliers, level shifts and variances changes in time series. J. Forecasting, 7: 1-20.

van Dijk D., P.H. Franses and A. Lucas, 1999. Testing for smooth transition nonlinearity in the presence of outliers. J. Business Econ. Stat., 17: 217-235.