A C T A U N I V E R S I T A T I S L O D Z I E N S I S
FOLIA OECONOMICA 255, 2011

*Daniel Kosiorowski**

DEPTH BASED STRATEGIES TO ROBUST ESTIMATION OF ARIMA PARAMETERS

Abstract. In this paper we propose two strategies for robust estimation of ARMA and GARCH models. The propositions are based on two statistical depth functions namely famous regression depth introduced by Rousseeuw and Hubert (1998) and general band depth function introduced by Lopez-Pintado and Romo (2006). We study a performance of the propositions on various time series simulated from ARMA(1,1) and GARCH(1,1) models containing additive outliers.

Key words: depth function, robust estimation, ARMA, GARCH.

I. OUTLIERS IN TIME SERIES

Often data considered in a broad range of economic applications contains one or more atypical observations called outliers. This refers to observations that are well separated from the majority or a center of the data cloud, or in some way deviate from the general pattern of the data. Outliers in financial or macro-economical time series are more complex than in the other situations, where there is no temporal dependence in the data (for details see Marona et al. (2006)). Time series outliers can have an arbitrarily negative influence on parameter estimates for time series models, and the nature of this influence depends on the type of outlier. In the time series setting we encounter several different types of outliers. From a model – based point of view we have among others additive outliers (AO), replacement outliers (RO) and innovation outliers (IO). The AO model is a special case of the RO model, IO outliers refers to a special type of a process ex. ARMA process with a heavy – tailed distribution of the innovations (for details see Marona et al. (2006). Further on we use a probability model for time series outliers including additive outliers (AO). Let x_t be a wide – sense stationary 'core' process of interest, and let v_t be a stationaryoutlier process which is a non – zero fraction ε of time i.e. $P(v_t = 0) = 1 - \varepsilon$. Under an AO model, instead of x_t one actually observes

* Ph.D. Department of Statistics, Cracow University of Economics.

$$y_t = x_t + v_t , \qquad (1)$$

where the processes x_t and v_t are assumed to be independent of one another.

The AO model will generate mostly isolated outliers if v_t is i.i.d. process, with scale much larger than that of x_t. In the presence of the AO outliers in economical time series classical estimators of ARMA and GARCH models generally became useless.

II. DEPTH FUNCTION IN ROBUST TIME SERIES ANALYSIS

A statistical depth function expresses the centrality or "outlyingness" of an observation within a set of data (or with respect to a probability distribution) and provides a criterion to order observations from center - outwards. For a detailed overview see Serfling (2006) and references therein. For other applications of the statistical depth functions in a robust economical analysis see e.g. Kosiorowski (2007) or Kosiorowski (2008).

We use the notion introduced by Rousseeuw and Hubert (1998) notion of regression depth in order to propose a robust procedure for the ARMA(p,q) parameters estimation. The regression depth measures the quality of any candidate fit in a linear regression setting. This fits with higher regression depth fit the data better than does fits with lower regression depth. Hence, the regression depth ranks all possible fits from worst (depth=0) to best (maximal depth) case. The errors in the underlying regression model are assumed to be independent, each having zero median. These are very weak conditions, e.g. the error distribution does not have to be symmetrical, nor does it have to stay the same across different values of predictors. A maximal depth estimator (MDE) is a fit which maximizes regression depth. This is one of the best robust estimators for linear regression (for details see Van Aelst and Rousseeuw (2000)).

We also use a generalized band depth function introduced by Lopez-Pintado and Romo (2006). Lopez-Pintado and Romo (2006) has extended the notion of statistical depth function to deal with functional observations. They proposed robust graph – based methods for supervised classification of curves. We incorporate their concepts in robust estimation for ARMA and GARCH models.

III. ROBUST ESTIMATION FOR ARMA MODELS

An important class of models for describing the single time series $\{z_t\}$ is the class of autoregressive - moving average models referred to as ARMA(p,q) models

$$(z_t - \mu) = \phi_1(z_{t-1} - \mu) + \ldots + \phi_p(z_t - \mu) + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}\,, \qquad (2)$$

where z_t is a stationary time series with a fixed mean μ, a_t is a random residual series, $\phi_1,\ldots,\phi_p$, $\theta_1,\ldots,\theta_q$, μ are parameters to be estimated from the data.

ARMA models may be fitted to data, using an iterative cycle of identification, estimation and checking. Classical statistical procedures for estimating ARMA models based on maximum likelihood, least squares or autocovariance estimates are not robust in the presence of AOs or ROs. Proposed diagnostic procedures generally suffer from the masking problem. Estimates based on regular residuals (M-, S- estimates) are not very robust. This is due to the fact that an outlier in one period, does not only affect the residual corresponding to this period, but it may also affect all the subsequent residuals (for details see Maronna et al. (2006)). Recently Muller et al. (2009) proposed a generalization of the MM- estimates introduced by Yohai for regression which are robust when the series contain outliers. They showed also several asymptotic properties of the propositions.

In this paper we propose an alternative procedure for robust estimation ARMA models. Our regression depth based proposition is user-friendly and performs well due to the very good statistical properties of the regression depth.

PROPOSITION 1: Let $\mathbb{X}_T = \{y_1,\ldots,y_T\}$, $2 < T$ denote a time series under consideration. We obtain estimates of the parameters of ARMA(p,q) in a two step procedure :

STEP 1: We calculate MDE estimates of the AR(p) part of the underlying process by choosing $\phi_1,\ldots,\phi_p$ as the maximal regression depth estimates applied to a data set $\mathbb{Y} = \{y_1,\ldots,y_{T-p}\}$, $\mathbb{X}_1 = \{y_2,\ldots,y_{T-p+1}\}$, $\ldots$, $\mathbb{X}_p = \{y_p,\ldots,y_T\}$ obtained from $\mathbb{X}_T$.

STEP 2: We add the MA(q) part to the estimated in the step 1 AR(p) part by minimizing a robust measure of a dispersion between observed and generated by the model values e.g. MAD (median absolute deviation).

IV. ROBUST ESTIMATION FOR GARCH MODELS

Many time series display time – varying dispersion, or uncertainty, in the sense that large (small) absolute innovations tend to be followed by other large (small) absolute innovations. Let y_t denote the observable univariate discrete-time stochastic process of interest. Denote the corresponding innovation process by ε_t, where $\varepsilon_t = y_t - E_{t-1}(y_t)$, and $E_{t-1}(\cdot)$ refers to the expectation conditional on time (t-1) information. A general specification for the innovation process that takes account of the time-varying uncertainty would then be given by

$$\varepsilon_t = z_t \sigma_t , \tag{3}$$

where z_t is an i.i.d. mean - zero, unit - variance stochastic process, and σ_t represents the time-t latent volatility; i.e. $E(\varepsilon_t^2 \mid \sigma_t) = \sigma_t^2$. In the GARCH(p,q) model, the conditional variance is parameterized as a distributed lag of past squared innovations and past conditional variances,

$$\sigma_t^2 = c + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \sigma_{t-j}^2 \tag{4}$$

$$\equiv \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2 , \tag{5}$$

where B denotes the backshift (lag) operator, i.e., $B^i y_t \equiv y_{t-i}$.

This model is usually estimated by maximum likelihood (QML) assuming that the distribution of one observation conditionally to the past is normal. The QML estimate based on a normal likelihood is very sensitive to the presence of a few outliers in the sample. Several authors proposed robust estimates for GARCH(p,q) models (see Muller and Yohai (2007)). The main part of the propositions however is based on the predictors of the conditional variance which are very sensitive to large outliers. We propose a strategy based on two statistical depth functions and standard ARMA-based method of identification of the GARCH(p,q) model. We can use our proposition in the case of analysing several time series generated by the same model. Some part of the time series may be outliers and/or each time series may contain AO outliers. In our opinion our simple depth based proposition could be an alternative to latent variables approaches, which generally need very long time series or to a BM-estimators proposed by Muller and Yohai (2007). First it is worth noticing that rearranging the terms in (5), we obtain

$$[1-\alpha(B)-\beta(B)]\varepsilon_t^2 = \omega + [1-\beta(B)]\nu_t \tag{6}$$

where $\nu_t \equiv \varepsilon_t^2 - \sigma_t^2$. Since $E_{t-1}(\nu_t) = 0$, and the GARCH(p,q) formulation in (6) we can estimate process as an ARMA(max{p,q},p) model for the squared innovation process $\{\varepsilon_t^2\}$.

PROPOSITION 2: Let
$\mathbf{y}_1(t) = \{y_1^{(1)},...,y_n^{(1)}\}$, $\mathbf{y}_2(t) = \{y_1^{(2)},...,y_n^{(2)}\}$, ..., $\mathbf{y}_k(t) = \{y_1^{(k)},...,y_n^{(k)}\}$,
denote k- time series generated by the same GARCH(p,q) process. We obtain an estimate of the underlying process in a two-step procedure.

STEP 1: We choose the deepest time series from the k- considered time series as a sample median induced by Lopez-Pintado and Romo generalized band depth.

STEP 2: We apply the first proposition to the standard ARMA-based identification of the GARCH(p,q) process.

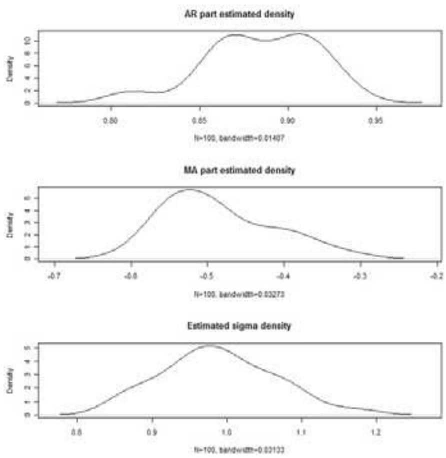

Fig 1. Kernel density estimation of the proposed parameters estimators of the ARMA(1,1) with $\phi_1 = 0.7$, $\theta_1 = -0.5$, $\sigma = 1$. Each of the simulated trajectories contained 5\% of the additive outliers.

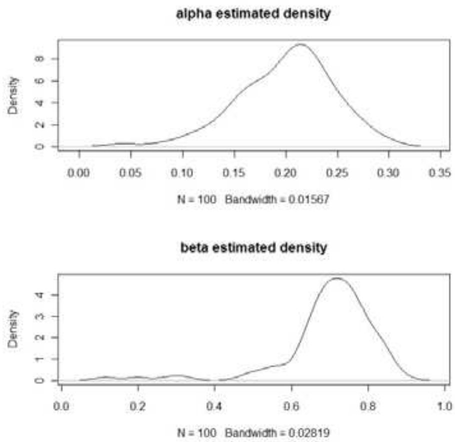

Fig. 2: Kernel density estimation of the proposed parameters estimators of the GARCH(1,1) with $\alpha_1 = 0.2$, $\beta_1 = 0.7$. Two of each five of the simulated trajectories contained 10% of the additive outliers.

V. SIMULATION EVIDENCE

To investigate the behavior of our first proposition we ran simulations with 100 samples of sizes 200 generated by a stationary normal one-dimensional AR(1) with $\phi_1 = 0.7$ and two-dimensional AR(1,1) with $\phi_1 = 0.7$, $\theta_2 = -0.5$ models, both with $\sigma_u = 1$ and $\gamma = 0$, (scale of the innovations and the intercept). We considered situations when instead of data point one actually observes (AO model) $y_t = x_t + v_t$, where the processes x_t and v_t are assumed to be independent of one another. We assumed v_t has a normal mixture distribution $v_t \sim (1-\varepsilon)N(0,1) + \varepsilon N(0,100)$ where $\varepsilon = 0\%,5\%,10\%,20\%$. Table 1 shows the results for the maximal regression depth regression estimate for AR(1) parameters where $\varepsilon = 0\%,5\%,10\%,20\%$. of the AO outliers. Table 2 shows the results the maximal regression depth estimate for AR(1,1) parameters with $\varepsilon = =0\%,5\%,10\%,20\%$. of the AO outliers. It is easily seen that the conditional least squares (similarly LAD, LTS) estimate is much affected by AO outliers in both situations. Note that the proposed estimate performs better than classical M-estimate for AR parameters. We also ran also simulations of sizes n=500 from ARMA(1,1) model with $\phi_1 = 0.9$, $\theta_1 = -0.5$, $\mu = 0$, $\sigma_u = 1$. Figure 1 shows kernel density estimation of the proposed parameters estimators of the ARMA(1,1). Each of the simulated trajectories contained 5% of the additive outliers. These results seem to be rather promising. In order to examine the performance of the second proposition we ran simulations of five time series generated by GARCH(1,1) model with $\alpha = 0.2$, $\beta = 0.7$. Each of the simulated five time series was of size n=500 observations. Two of the simulated five time series contained 10% of the AO outliers. Figure 2 shows results of kernel estimation of the densities of parameters proposed estimators. The results show very good properties of the proposition in terms of robustness to the AO outliers. Note that the first part of the second proposition (Pintado-Lopez & Romo median) also performs well also in cases of several ARMA time series or real data examples.

Tab 1. Percent of the additive outliers in a simulated trajectories from AR(1) model with $\phi_1 = 0.7$ vs. mean of the first proposition estimates

OUTLIERS %	$\hat{\phi}_1$
0	0.692
5	0.615
10	0.575
15	0.478

Tab 1. Percent of the additive outliers in a simulated trajectories from AR(2) model with $\phi_1 = 0.7$, $\phi_2 = 0.2$ vs. mean of the first proposition estimates

OUTLIERS %	$\hat{\phi}_1$	$\hat{\phi}_2$
0	0.776	–0.492
5	0.631	–0.335
10	0.489	–0.234
15	0.268	–0.075

VI. CONCLUSIONS

In our opinion the proposed strategy for ARMA(p,q) model estimation is an attractive approach to robust estimation of the real economic processes' parameters. Simulation studies show that our approach is not only more robust than conditional least squares or least absolute deviations estimators but also than some new promising propositions as M- or S- estimators (see Marona et al. (2006), and procedures based on robust filters. Note that our proposition performs well also in situations where data does not contain outliers.

REFERENCES

Kosiorowski D. (2007), *Nonparametric Equity of Two Shapes Test Based on Multivariate Quantile Functional*, Bulletin of the ISI 56th Session.

Kosiorowski D. (2008), *Robust Classification and Clustering Based on the Projection Depth Function*, [in:] Brito P. (Eds.), Proceedings in Computational Statistics 2008 (COMPSTAT 2008), vol. II, p. 209–216, Physica-Verlag.

Lopez-Pintado S. and Romo J. (2006), *Depth-based classification for functional data*, [in:] Liu R.Y., Serfling R., Souvaine D. L. (Eds.), Series in Discrete Mathematics and Theoretical Computer Science, AMS, vol. 72, 103–119.

Marona R. A., Martin R. D., Yohai V. J. (2006), *Robust Statistics – Theory and Methods*. John Wiley & Sons, Chichester.

Muler N., Pena D. and Yohai V. J. (2009), *Robust estimation for ARMA models*, Annals of Statistics 37 (2),816-840.

Muler N.,Yohai V. J. (2007), *Robust estimates for GARCH models*, Technical Report Instituto de Calculo Facultaded de Cinecias Exactas y Naturales Universidad de Buenos Aires.

Rousseeuw P. J., hubert M. (1998), *Regression Depth*, Journal of the American Statistical Association, 94, 388–433}.

Serfling R. (2006), *Depth Functions in Nonparametric Multivariate Inference*, [in:] Liu R.Y., Serfling R., Souvaine D. L. (Eds.), Series in Discrete Mathematics and Theoretical Computer Science, AMS, vol. 72, 1–15.

Van Aelst S., Rousseeuw P. J. (2000), *Robustness Properties of Deepest Regression*, J. Multiv. Analysis, 73, 82–106.

Daniel Kosiorowski

STRATEGIE ODPORNEJ ESTYMACJI PARAMETRÓW MODELU ARIMA

W pracy badamy kilka strategii odpornej estymacji parametrów modeli ARIMA i GARCH. Porównujemy między innymi podejścia wykorzystujące statystyczne funkcje głębi: wykorzystujące koncepcję głębi odnoszącą się do funkcji a zaproponowaną przez Lopez-Pintado i Romo (2005) oraz własne propozycje wykorzystujące głębie regresyjną.