IMPORTANCE OF ROBUST METHODS FOR PARAMETER ESTIMATING IN AR(*p*)

Samuel Flimmel – Jan Fojtík - Ivana Malá - Jiří Procházka

Abstract

Recently, there has been a growth of demand for robust parameter estimation methods. One of the reasons for this growth is an extensive usage of big data, since with a growing number of observations, the probability of outlier presence also rises. With an outlier presence, it is highly recommended to work with robust methods because standard methods are not able to deal correctly with outliers, and, consequently, standard estimates are usually biased. Autoregressive process AR(p) is well known and widely used in statistics and economical modelling. It is very important to estimate parameters of this model correctly, and we show the suitability of robust methods for this task. We present several robust methods and compare them with a standard method using a simulation study. Additive outlier (AO) model and innovative outlier (IO) model are used in the simulations to contaminate data with outliers. For the simulation study, we use the R statistical software.

Key words: parameter estimating, robust methods, autoregressive process

JEL Code: C02, C22, G10

Introduction

AR(*p*) process is well-known and widely used as one of the process which can explain the residue of randomness in a random process. Often, Box-Jenkins methodology (Box, 1970) is used to identify an appropriate process which we could use to represent our data or residues. The Box-Jenkins methodology has several steps. Firstly, we need to solve the seasonality and the stationarity of the process. Secondly, we estimate ARMA orders and then we can estimate the parameters. The second step, meaning the ARMA order determination, was the main point of interest in our paper (Flimmel, 2017) at the last conference. This year, we focus on the final step – parameter estimation.

Currently, when we face the big data problems, the importance of using robust methods is growing. Robust methods are usually more insensitive to outliers and they give better estimation in the case of outlier presence, as it was already shown by (Chan,1992).

A comprehensive overview of the most important robust methods for ACF estimation was made by (Dürre, 2015). For a more detailed description of the methods you can see (Maronna, 2006), (Ma, 2000) or others.

In Section 1, we establish some notation that we work with in this paper. In Section 2, we briefly introduce two robust methods and a standard method that we use in our comparison. In Section 3, we show results from our simulation study used to compare the methods.

1 Definitions and notation

Let us define Gaussian white noise, which is a zero-mean mutually uncorrelated time series $\{\varepsilon_n, n \in N_0\}$ with an unknown constant variance $\sigma_\varepsilon^2 > 0$.

We define an autoregressive process AR(p) by the equation

$$X_n = \varphi_1 X_{n-1} + \varphi_2 X_{n-2} + \cdots + \varphi_p X_{n-p} + \varepsilon_n, \tag{1}$$

where $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_p) \in R^p$ is a vector of parameters, $\{\varepsilon_n, n \in N_0\}$ is the white noise and $\varphi_p \neq 0$.

We define an autocovariance function of the lag k *R(k)* of the stationary process $\{X_n, n \in N_0\}$ as

$$R(k) = E(X_k - \mu)(X_0 - \mu), \tag{2}$$

where μ is the expected value of the process.

Let us define an autocorrelation function (ACF) of the lag k $\rho(k)$ of the stationary process $\{X_n, n \in N_0\}$ as

$$\rho(k) = \frac{R(k)}{\sigma_X^2}, \tag{3}$$

where σ_X^2 is the variance of the process.

2 Estimation methods

Let us briefly introduce all methods that we use in our simulation study. Firstly, we need to estimate an autocorrelation function of the process. We have m+1 observations $X_0, X_1, \ldots, X_m$, from which we estimate the ACF.

Let us start with a standard method, e.g. according to (Hamilton, 1994):

$$\hat{\rho}_S(k) = \frac{\sum_{i=0}^{m-k}\left(X_{i+k} - \bar{X}\right)\left(X_i - \bar{X}\right)}{\sum_{i=0}^{m}\left(X_i - \bar{X}\right)^2}, \tag{4}$$

where $\bar{X}$ is the average of $X_0, X_1, \ldots, X_m$.

We introduce two robust methods: a method based on the Gnanadesikan-Kettenring approach and a method based on robust filtering.

The method based on the Gnanadesikan-Kettenring approach, which was introduced by (Gnanadesikan and Kettenring, 1972), is defined as

$$\hat{\rho}_{GK}(k) = \frac{Q_{m-k}^2(u+v) - Q_{m-k}^2(u-v)}{Q_{m-k}^2(u+v) + Q_{m-k}^2(u-v)}, \tag{5}$$

where u is the vector $(X_{m-k}, X_{m-k+1}, \ldots, X_m)$, v is the vector $(X_0, X_1, \ldots, X_k)$ and Q_m is a robust estimator of the scale. It was proposed by (Croux, 1992) and it is defined as:

$$Q_m = c\left[\left|X_i - X_j\right|, i < j\right]_l, \tag{6}$$

where $[\cdot]_l$ is the lth order statistic, and l is defined as

$$l = \left\lfloor \frac{\binom{m}{2} + 2}{4} \right\rfloor + 1, \tag{7}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. The factor c is for consistency, for the Gaussian distribution c = 2.2191. The method of this robust ACF estimator was presented by (Ma, 2000).

The robust filtering approach takes the time series structure into account. The idea is to have robust filtered values instead of original observations and calculate the ACF from these filtered values. Practically, we replace outliers by some reasonable values.

Firstly, we estimate a "long" AR process, which we use for robust filtering. Consequently, we obtain fitted values using the robustly filtered τ-scale estimate and, finally, we calculate the autocorrelation function. The method of this robust ACF estimator was presented by (Maronna, 2006).

When we already have the estimation of the ACF, we are able to estimate parameters of the AR(p) model using the moment method:

$$\hat{\boldsymbol{\varphi}} = \begin{pmatrix} 1 & \hat{\rho}(1) & \cdots & \hat{\rho}(p-1) \\ \hat{\rho}(1) & 1 & \cdots & \hat{\rho}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}(p-1) & \hat{\rho}(p-2) & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(p) \end{pmatrix}. \tag{8}$$

The moment method can be used for each ACF estimator. Therefore, we have 3 different parameter estimator that we compare in a simulation study.

3 Simulation study

The simulation study was designed in the *R* software and we use the *R* package *robts*. However, the package is still not approved by CRAN, thus a number of functions was coded by authors of this paper to validate correctness of the package. After the validation, we used functions from the package to obtain estimations in the simulation study.

To evaluate estimation accuracy we use *mean absolute error* (MAE) and *mean absolute percentage error* (MAPE). The mean absolute error is given by

$$MAE = \frac{\sum_{i=1}^{s} \left| {}_i\hat{\varphi}_j - \varphi_j \right|}{s}, \tag{9}$$

where ${}_i\hat{\varphi}_j$ is an estimation of the *i-th* simulation and *j-th* component of the vector of parameters, and *s* is the number of simulations.

The mean absolute percentage error is defined by

$$MAPE = \frac{100}{s} \sum_{i=1}^{s} \left| \frac{{}_i\hat{\varphi}_j - \varphi_j}{\varphi_j} \right|, \tag{10}$$

where ${}_i\hat{\varphi}_j$ is an estimation of the *i-th* simulation and *j-th* component of the vector of parameters, and *s* is the number of simulations.

Firstly, we have the 3 simplest models: AR(1), AR(2) and AR(3). For the probability of outliers being present in one simulation (ε), we choose 3 cases: ε = 0%, ε = 1% and ε = 5%. Therefore, because of the 3 models and 3 outlier probabilities, we have 9 different cases. For every case, we run 5000 simulations with 1000 observations.

In each case, we estimate all parameters of the model using all 3 described methods. The accuracy of the method is evaluated by 2 mentioned criterion: MAE and MAPE.

Model parameters are chosen randomly. Absolute values of the parameters are generated with a uniform distribution, i.e. $\varphi_i \sim U((0.2,1.0))$. Values close to zero are not taken into account because they are difficult to observe. The sign of the parameters is generated

randomly using Bernoulli's distribution with the probability of success π= 0.5. Subsequently, we check whether these parameters give a stationary process, and, if necessary, we repeat the procedure.

Secondly, we work with the models AR(p) for p = 1, ..., 5. Once again, we have the same 3 probabilities of outliers being present in a simulation, and for every probability value we run 5000 simulations with 1000 observations. Subsequently, all estimated parameters are evaluated, using a box plot to have a graphical representation of the results.

We use an additive outlier model and an innovative outlier model (see e.g. Maronna, 2006) in the simulation study, which are described in the following 2 subsections.

1.1 Additive outlier model

We work with the additive outlier (AO) model in this subsection. Firstly, we show Table 1, where we see a comparison of the 3 above described methods.

Tab. 1: Comparison of 3 methods using the AO model

model	φ_k	criteria	standard method			GK approach			robust filtering approach		
			0%	1%	5%	0%	1%	5%	0%	1%	5%
AR(1)	φ_1	MAE	.0187	.1720	.3793	.0204	.0211	.0270	.0205	.0205	.0209
		MAPE	4.2%	33.5%	68.8%	4.6%	4.7%	5.6%	4.6%	4.7%	4.7%
AR(2)	φ_1	MAE	.0200	.1964	.3861	.0243	.0259	.0483	.0543	.0560	.0778
		MAPE	4.5%	34.4%	68.5%	5.3%	5.4%	8.8%	13.8%	13.6%	17.1%
	φ_2	MAE	.0201	.1902	.3737	.0238	.0256	.0465	.0615	.0650	.0945
		MAPE	4.8%	39.3%	71.8%	5.7%	5.9%	9.2%	12.1%	12.6%	17.4%
AR(3)	φ_1	MAE	.0227	.2174	.3981	.0461	.0477	.0770	.0284	.0291	.0320
		MAPE	4.9%	39.7%	71.5%	8.9%	10.2%	14.4%	6.1%	6.2%	6.6%
	φ_2	MAE	.0228	.2034	.3679	.0464	.0458	.0787	.0325	.0342	.0405
		MAPE	5.3%	40.5%	71.3%	10.5%	10.4%	16.3%	7.8%	8.2%	9.7%
	φ_3	MAE	.0234	.2197	.3781	.0473	.0492	.0871	.0322	.0336	.0405
		MAPE	5.6%	48.6%	78.2%	11.4%	11.5%	19.6%	7.4%	7.7%	8.9%

Source: Authors' own calculations

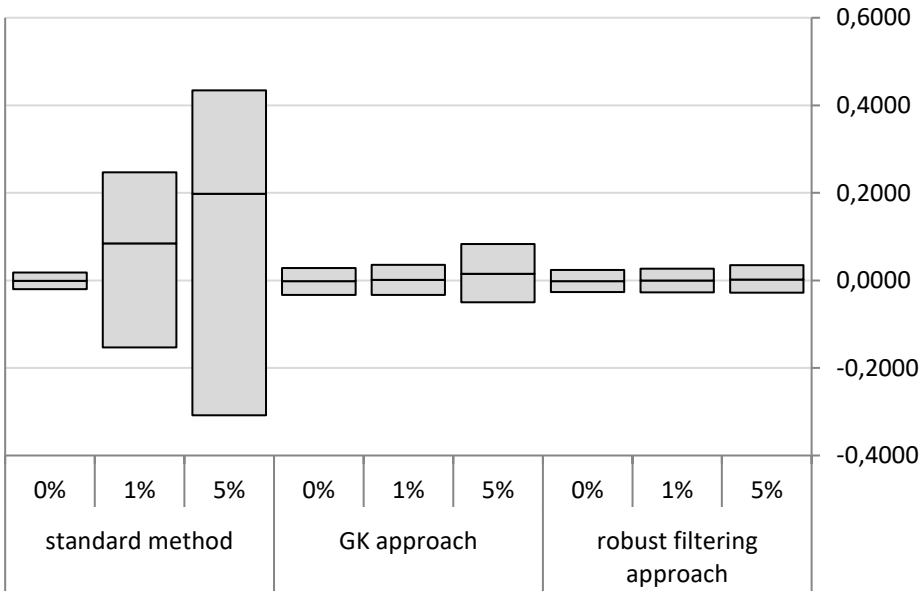
We can see that the standard method is not able to process the outliers. By increasing the probability of outlier presence, which means more outliers present in the observations, the

accuracy of parameter estimation decreases drastically. Naturally, the standard method gives the best results in the case of no outliers in the observations. However, the differences between the standard method and the robust methods are quite small.

The robust filtering approach is more sophisticated and should give better results, but it is not true in the case of AR(2). It was quite surprising and, consequently, additional simulations for this special case were made. Additional simulations confirmed worse results for AR(2) in comparison with the GK approach. The reason why the robust filtering approach is better in AR(1) and AR(3), and not in AR(2), is unknown and will be explored. However, the main message of the table is the incapability of the standard method to work with outliers. On the other hand, the robust methods are much less sensitive to outliers.

Secondly, we show Figure 1 with boxplots of errors (simple difference between the estimates and the real parameters), where we see a dramatic increase of boxplot size in the case of the standard method. The parameter estimation is much more volatile and inaccurate in comparison with the robust methods. The robust filtering approach looks a little better (it is slightly less volatile) than the GK approach because of a smaller difference between the first and the third quartile. We should also mention that the GK approach gave a few extreme estimation errors ($\sim \pm 10$).

Fig. 1: Boxplots of 3 methods using AO model



Source: Authors' own calculations

1.2 Innovative outlier model

We work with the innovative outlier model in this subsection. Firstly, we show Table 2, where we see a comparison of the 3 above described methods.

Tab. 2: Comparison of 3 methods using IO model

model	φ_k	criteria	standard method			GK approach			robust filtering approach		
			0%	1%	5%	0%	1%	5%	0%	1%	5%
AR(1)	φ_1	MAE	.0189	.0179	.0177	.0204	.0225	.0460	.0207	.0203	.0228
		MAPE	4.3%	4.0%	4.0%	4.6%	4.8%	8.4%	4.7%	4.6%	4.9%
AR(2)	φ_1	MAE	.0201	.0190	.0189	.0242	.0270	.0540	.0539	.0604	.1155
		MAPE	4.5%	4.2%	4.2%	5.2%	5.7%	10.8%	13.6%	14.3%	25.3%
	φ_2	MAE	.0202	.0193	.0194	.0238	.0273	.0579	.0611	.0656	.1014
		MAPE	4.9%	4.6%	4.6%	5.7%	6.3%	13.1%	12.2%	12.9%	21.0%
AR(3)	φ_1	MAE	.0233	.0216	.0208	.0543	.0636	.1450	.0298	.0401	.0987
		MAPE	4.9%	4.6%	4.4%	10.9%	12.3%	27.0%	6.4%	8.5%	21.9%
	φ_2	MAE	.0230	.0220	.0214	.0517	.0592	.1253	.0327	.0489	.1127
		MAPE	5.4%	5.1%	5.0%	11.3%	13.1%	26.5%	8.0%	11.7%	28.4%
	φ_3	MAE	.0239	.0232	.0218	.0549	.0644	.1365	.0325	.0451	.0993
		MAPE	5.7%	5.5%	5.2%	12.9%	15.1%	34.0%	7.5%	10.2%	22.5%

Source: Authors' own calculations

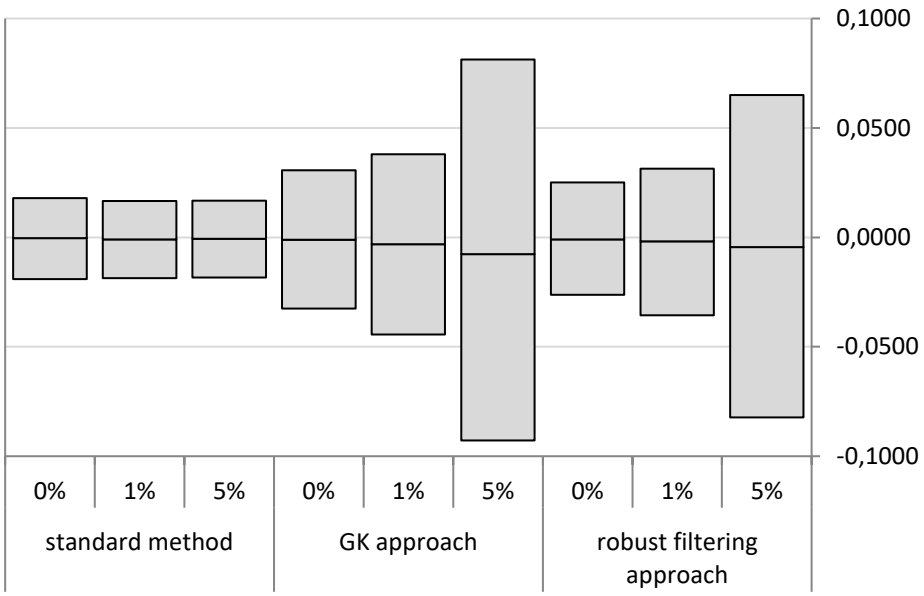
We can see that innovative outliers affect the standard estimation minimally. MAPE results are almost constant in the sense of outlier probabilities. Actually, they slightly decrease with increasing outlier probability. Overall, it seems that the standard non-robust method is able to work with this type of outliers.

On the contrary, both robust methods show worse results in comparison with the standard approach. For AR(1), the robust filtering approach gives at least similar results as the standard approach, but AR(2) and AR(3) cases do not confirm this behavior.

We can see again worse results of the robust filtering approach for the AR(2) case. It shows a certain confirmation of the phenomenon which was pointed out for the AO model. Despite this abnormality, the robust filtering approach gives better results for the remaining models AR(1) and AR(3), in comparison with the GK approach.

Secondly, we show Figure 2 with boxplots of errors (simple difference between the estimates and the real parameters).

Fig. 2: Boxplots of 3 methods using AO model



Source: Authors' own calculations

Fig. 2 confirms the numbers from Table 2, since the size of the boxplots is almost constant for the standard approach. However, for the both robust methods, the size of the boxplots grows. For the GK approach, it grows slightly more. We should remind that the scale of Figure 2 is more detailed in comparison with Figure 1, where the difference was even more dramatic.

Conclusion

We briefly introduced a standard method and two robust methods for ACF estimation. Using ACF estimators, we presented an estimation of the parameters for AR(*p*) model.

We provided a simulation study where we compared the methods. The AOs have a strong impact on the standard method and we should not use this method in such situations. Both robust methods gave better results than the standard method. The method based on robust filtering looked even slightly better than the method based on the GK approach, except for the AR(2) model, which should be studied in more detail.

On the other hand, the IOs have no impact on the standard method, however, the robust methods are affected by them. We have to admit that it seems preferable to work with

the standard approach in the case of the IOs. Even in our last paper (Flimmel, 2017), it was shown that current robust methods have problems with this type of outliers.

In conclusion, we would recommend to check the outlier presence at first. Then we should try to detect the nature of outliers. If we detect innovative outliers, we should use the standard method. But if we detect additive outliers, we should definitely use one of the robust methods, otherwise we risk to estimate the parameters incorrectly.

Acknowledgment

This study was supported by the grant F4/17/2017 (Robustnosť v úmyselnom useknutí časového radu), which has been provided by the Internal grant agency of the University of Economics in Prague.

References

Box, G. E., & Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.

Chan, W. (1992). A note on time series model specification in the presence of outliers. *Journal of Applied Statistics*, 19(1), 117-124. doi:10.1080/02664769200000010

Croux, C., & Rousseeuw, P. J. (1992). *Explicit scale estimators with high breakdown point.* Berkeley, CA: Math. Sciences Research Inst.

Dürre, A., Fried, R., & Liboschik, T. (2015). Robust estimation of (partial) autocorrelation. *Wiley Interdisciplinary Reviews: Computational Statistics, 7*(3), 205-222. doi:10.1002/wics.1351

Flimmel, S., Čamaj, M., Malá, I., & Procházka, J. (2017). Importance of robust methods for ARMA order estimating. In Loster, T., Pavelka, T. (Eds.), *The 11th International Days of Statistics and Economics*, (pp. 374-383). Retrieved from https://msed.vse.cz/msed_2017/article/137-Flimmel-Samuel-paper.pdf

Gnanadesikan, R., & Kettenring, J. R. (1972). Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data. *Biometrics*, *28*(1), 81. doi:10.2307/2528963

Hamilton, J. D. (1994). *Time series analysis.* Princeton, NJ: Princeton University Press.

Ma, Y., & Genton, M. G. (2000). Highly Robust Estimation of the Autocovariance Function. *Journal of Time Series Analysis*, *21*(6), 663-684. doi:10.1111/1467-9892.00203

Maronna, R. A., Martin, R. D., & Yohai, V. J. (2006). *Robust statistics: theory and methods.* Chichester: Wiley.

Contact

Samuel Flimmel
University of Economics, Prague
W. Churchill Sq. 1938/4, Prague, Czech Republic
samuel.flimmel@vse.cz

Jan Fojtík
University of Economics, Prague
W. Churchill Sq. 1938/4, Prague, Czech Republic
xfojj00@vse.cz

Ivana Malá
University of Economics, Prague
W. Churchill Sq. 1938/4, Prague, Czech Republic
malai@vse.cz

Jiří Procházka
University of Economics, Prague
W. Churchill Sq. 1938/4, Prague, Czech Republic
xproj16@vse.cz