ARMA models with time dependent coefficients

Guy Mélard [1]

[1]ECARES, Université libre de Bruxelles, and ITSE sprl, Belgium
(gmelard@ulb.ac.be)

12 to 14 March 2019
NTTS 2019
New Techniques and Technologies for Statistics 2019
Charlemagne Conference Centre, Brussels, Belgium

Outline

1. Introduction
   - Objectives
   - Status
2. Body
   - Submitted text
   - The tdARMA Model
   - Example: tdARMA(1,1) model
   - U.S. Industrial Production series
3. Conclusions, Acknowledgements & References
   - Conclusions
   - Acknowledgements
   - References

Objectives

- Several methods to deal with time-varying time series models (Dahlhaus, 1996, 1997), rarely used for official statistics (Van Bellegem & von Sachs, 2004)
- An attempt by [AM2017] in NTTS 2017 for time-dependent AutoRegressive (tdAR) models with a small study on about 20 Belgian series
- Big dataset of U.S. industrial production series
- Question: do they benefit from tdARIMA models?
- Main idea: replace constant parameters by deterministic functions of time t
- First approximation: linear functions, hence 2 parameters (intercept + slope) instead of 1
- Main advantage: there is an asymptotic theory and empirical experiments to justify tests that slopes are 0

Status

- At the end of the NTTS 2017 presentation, a first study was presented on the U.S. industrial production dataset
- A large proportion of series showed significant slopes
- Comparison between constant ARIMA (cARIMA) and tdARIMA models
- Main criticisms
  - Many parameters were not statistically significant
  - BIC criterion was often worse for td models
  - Forecasts were generally worse also
  - Since the series had many outliers, a doubt remained

Submitted text

In the submitted text we presented new results on more recent data and promised improved results

- Longer series than in 2017
- Fitting the series with Tramo to obtain initial models but also ...
- ... working on linearised series instead of raw series
  - series corrected for outliers
  - and calendar and Easter effects
- Using a global test on the slopes in addition to the Wald tests on each slope
- Considering a stepwise procedure to eliminate non significant slopes
- Using rolling forecasts, not only forecasts with a fixed origin

The tdARMA Model

Definition of tdARMA(p, q), time dependent ARMA processes

- $(x_t, t \in \mathbb{N})$ solution of
  $x_t = \sum_{k=1}^{p} \phi_{tk} x_{t-k} + e_t - \sum_{k=1}^{q} \theta_{tk} e_{t-k}$
- $(e_t, t \in \mathbb{N})$ are independent random variables with mean 0 and constant standard deviation σ
- $\phi_{tk} = \phi_{tk}(\beta)$, $k = 1, ..., p$, $\theta_{tk} = \theta_{tk}(\beta)$, $k = 1, ..., q$: coefficients that are deterministic functions of time t and β (and possibly series length n)
- Vector β ($r \times 1$): all parameters to be estimated except σ
- True value $\beta = \beta^0$
- Initial values $x_t, e_t, t < 1$, supposed to be equal to 0 $^{(*)}$
- Under some assumptions, the Gaussian QML estimator is consistent and asymptotically normal [AM2006, AM2019]

Example: tdARMA(1,1) case

$$x_t^{(n)} = \phi_t^{(n)}(\beta) x_{t-1}^{(n)} + e_t^{(n)} - \theta_t^{(n)}(\beta) e_{t-1}^{(n)},$$

$$\phi_t^{(n)}(\beta) = \phi + \frac{t - \frac{n+1}{2}}{n-1}\phi', \quad \theta_t^{(n)}(\beta) = \theta + \frac{t - \frac{n+1}{2}}{n-1}\theta'$$

- Parameters: $\beta = (\phi, \phi', \theta, \theta')^T$ (with due conditions)
- Term $(n+1)/2$ to achieve orthogonality
- Factor and factor $1/(n-1)$ (or $1/n$) just for the asymptotics (to restrain the coefficient to a finite interval)
- Not a random sequence x_t but well random array $x_t^{(n)}$
- Classical stationary case with $\phi' = \theta' = 0$

ULB

Data & Procedure

Study on time series in a dataset of U.S. industrial production (Proietti-Lütkepohl, 2013), data from Jan1986

`http://www.federalreserve.gov/releases/g17/ipdisk/ip_nsa.txt`

- New release with fewer series (293) but some much longer
- Series fitted till Dec2016 leaving 2017-2018 to forecasts
- Finding constant ARIMA (cARIMA) models in an automated way, using Gómez & Maravall (2001) Tramo
- We use the linearised series produced
- Then we replaced the constant coefficients by linear functions of t for order ≤ 13, hence (tdARIMA) models
- We fit the cARIMA and tdARIMA models with the same specialized software package
- For the tdARIMA models, we omit non-significant slopes one by one

Statistics

Specification	# series	%	note
Levels vs logs	70/223	24/ 76	
Regular diff. vs none	280/13	96/ 4	4% with 2 differences
Seasonal diff. vs none	279/14	95/ 5	0% with 2 differences
Stationary vs nonstationary	0/293	0/100	
Airline model vs other	84/209	29/ 71	
Outliers vs none		79/ 21	0-20, on average 2.63
Trading day effect vs none		47/ 53	
Easter effect vs none		40/ 60	
ARMA parameters vs none		100/ 0	1-7, on average 3.12

cARIMA vs tdARIMA comparison

We compare the results of tdARIMA vs cARIMA models using the following criteria

- Is highest $|t|$ statistic of the td parameters > 1.96 *
- Is tdARIMA SBIC $<$ cARIMA SBIC ?
- Is tdARIMA residual standard deviation $<$ cARIMA one ?
- Is tdARIMA P-value of the Ljung-Box statistic for residual autocorrelation $>$ cARIMA one ?
- Is tdARIMA mean absolute percentage error (MAPE) in % $<$ cARIMA one ? (for year 2017)
- Same for rolling forecasts at horizons $h = 1, 3, 6, 12$?

We count the percentage over the 293 series.

*There is also a global test

Results

Criteria		%	% if td sign.
Highest $|t|$ statistic of td parameters > 1.96	(*)	46.42%	100.00%
tdARIMA SBIC $<$ cARIMA SBIC	(**)	20.82%	44.85%
tdARIMA residual std dev $<$ cARIMA		37.20%	77.94%
tdARIMA LB P-value $>$ cARIMA		26.28%	55.15%
tdARIMA forecasting MAPE $<$ cARIMA		24.23%	50.00%
tdARIMA $h = 1$ rolling forecasts MAPE $<$ cARIMA		18.09%	38.24%
tdARIMA $h = 12$ rolling forecasts MAPE $<$ cARIMA		17.41%	36.03%

(*) Statistically significant slope parameters at the 5 % level.

(**) In this version, the non-significant parameters were omitted. Previously the percentage was less than 10%.

Conclusions

- The results are better than in the previous version
- Significant slopes were not a consequence of presence of outliers, as feared
- They show that about one half of the series can be better represented by tdARIMA models rather than by cARIMA models
- For that half, the forecasts are not better, nor worse
- We can now take profit from the few longer series available, possibly with other submodels for the coefficients than linear dependency
- Considering a (marginal) heteroscedasticity is also possible
- The exercise can be reproduced on EU data

Acknowledgements

- Agustín Maravall for his help in order to produce the linearised time series in Tramo
- Rajae Azrak for her contribution to the theory (AM2006 & AM 2019)
- Ahmed Ben Amara for a very first version of a part of the program chain (Microsoft Excel Visual Basic modules)
- Dario Buono from Eurostat for exchanges and suggestions on the 2017 version

Thank you for your attention

[AM2006] Azrak R. and Mélard, G. (2006). Asymptotic properties of quasi-maximum likelihood estimators for ARMA models with time-dependent coefficients, *Statist. Inference for Stoch. Processes* **9**, No 3, 279-330.

[AM2017] Azrak R. and Mélard, G. (2017). AR models with time-dependent coefficients with a macro-economic example, NTTS 2017 Conference on New Techniques and Technologies for Statistics, Eurostat, Brussels, 14-16 March, 2017.

[AM2019] Azrak, R. and Mélard, G. (2019) Asymptotic properties of conditional least-squares for array time series, in preparation.

Box, G. E. P., Jenkins, G. M., Reinsel G. C. & Ljung, G. M. (2015). *Time series analysis, forecasting and control*, 5th edn. Wiley, New York.

Dahlhaus, R. (1996a) Maximum likelihood estimation and model selection for locally stationary processes, *J. Nonparametr. Statist.* **6**, 171-191.

Dahlhaus, R. (1997) Fitting time series models to nonstationary processes, *Ann. Statist.* **25**, 1-37.

Gomez, V., and Maravall, A. (2001) Automatic modelling methods for univariate series. Chapter 7 in Peña, D. Tiao, G. C. and Tsay, R.S., *A Course in Time Series Analysis*. Wiley, New York, 2001, pp. 171-201.

[M1982] Mélard, G. (1982) The likelihood function of a time-dependent ARMA model, in O. D. Anderson and M. R. Perryman (eds), *Applied Time Series Analysis*, North-Holland, Amsterdam, 229-239.

[M1985] Mélard, G. (1985) *Analyse de données chronologiques*, Coll. Séminaire de mathématiques supérieures - Séminaire scientifique OTAN (NATO Advanced Study Institute) **89**, Presses de l'Université de Montréal, Montréal. Reprint `http://homepages.ulb.ac.be/~gmelard/Montreal1985_2012.pdf`

Proietti, T., and Lütkepohl, H. (2013) Does the BoxCox transformation help in forecasting macroeconomic time series? *International Journal of Forecasting*, **29**, 88-99.

Van Bellegem, S. and von Sachs, R. (2004), Forecasting economic time series with unconditional time-varying variance, International Journal of Forecasting 20, 611-627.