Code [will be] available @ <https://github.com/szakrytnoy/financial-econometrics>

# C:\Users\sergeyz\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\91536DAD.tmpIntroduction and scope

This assignment is dedicated to ARMA models – univariate class of econometric models, consisting of autoregressive AR(p) and stochastic MA(q) components. The moving average component represents a linear combination of q lags of white noise.

We are presented with 3 datasets: A, C and E. After graphical exploratory analysis, we assume weak stationarity of the series.

# Methodology

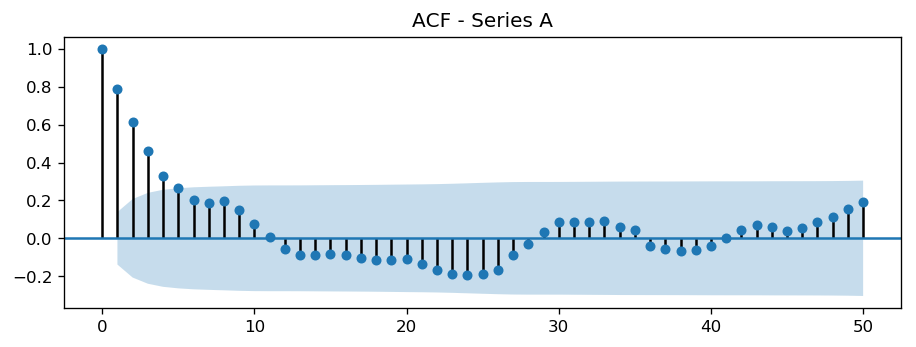
One of the common ways to determine optimal combinations of parameters (p, q) is to plot autocorrelation function (ACF) and partial autocorrelation function (PACF). The latter represents correlation coefficient between the original series and itself with lag = p and excludes the effects of lagged series in-between.

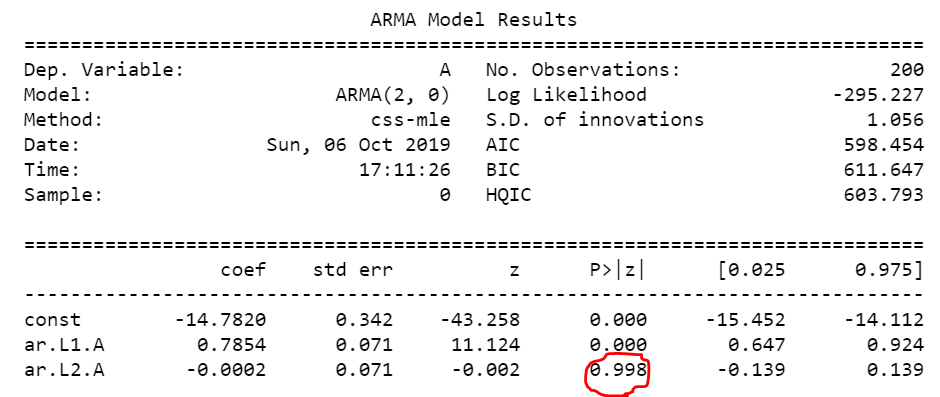
The criteria for optimal model is mirrored in cases of AR and MA processes:

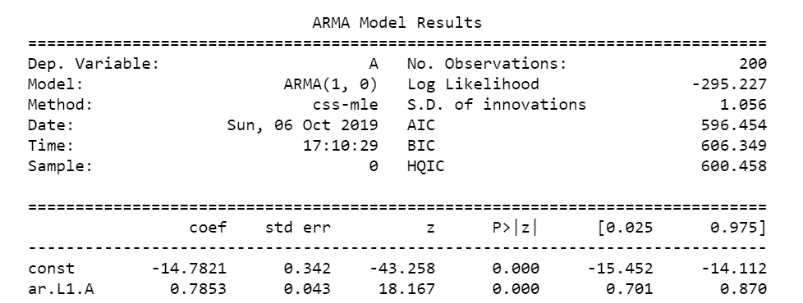
* [AR] p spikes in PACF, geometrically decaying ACF
* [MA] q spikes in ACF, geometrically decaying PACF

Once the graphical analysis is done, we will evaluate one or more models with different parameters (p, q) and apply AIC criterion (the smaller – the better) to identify the optimal combination. The model will then be checked for autocorrelation of residuals, which by definition of autoregressive model should not be there.

# C:\Users\sergeyz\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\6CE7DA34.tmpSeries A

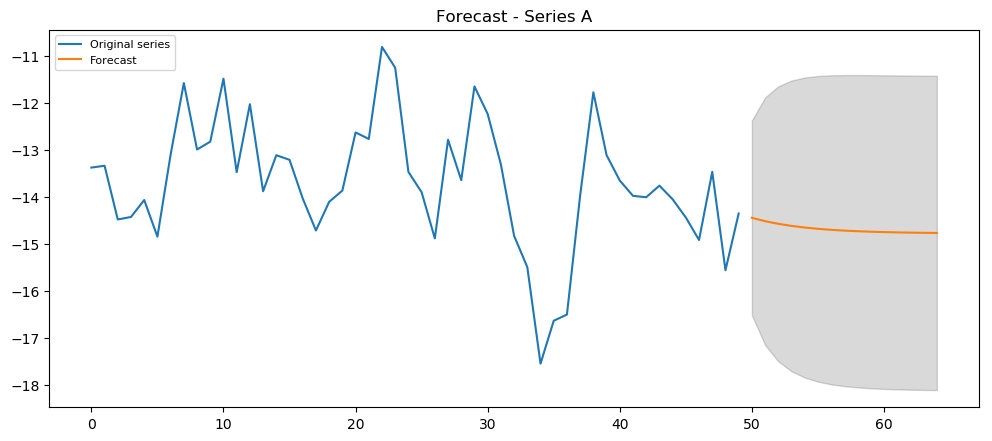


AR(2) process is clearly visible from the charts with 2 spikes in PACF and decaying ACF. However, when running an automated script with values of p and q from 0 to 3, the one that generates minimal AIC is ARMA(1, 0). Trying to override this choice with ARMA(2, 0) results in an insignificant coefficient at the second lag.

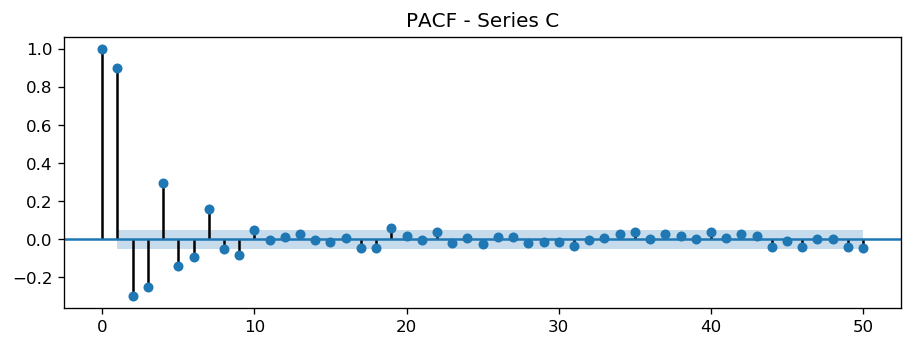
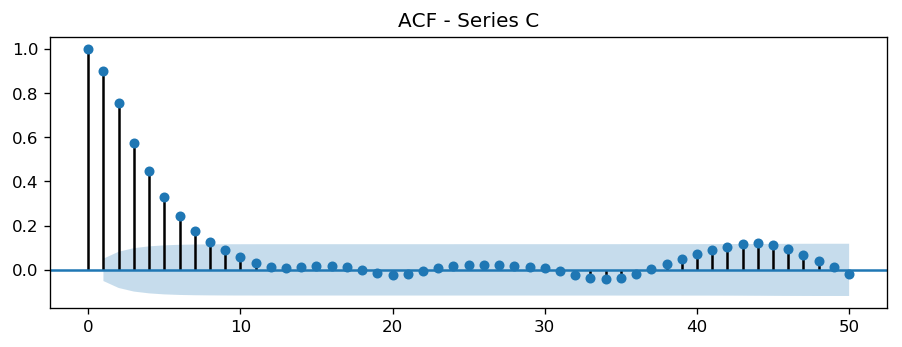


Thus, decision has been made to stick to ARMA(1, 0). The Ljung-Box test shows no autocorrelation of residuals (pval = 0.58).

Last 50 observations and the forecasted values from **ARMA(1, 0)** with confidence intervals are presented below.

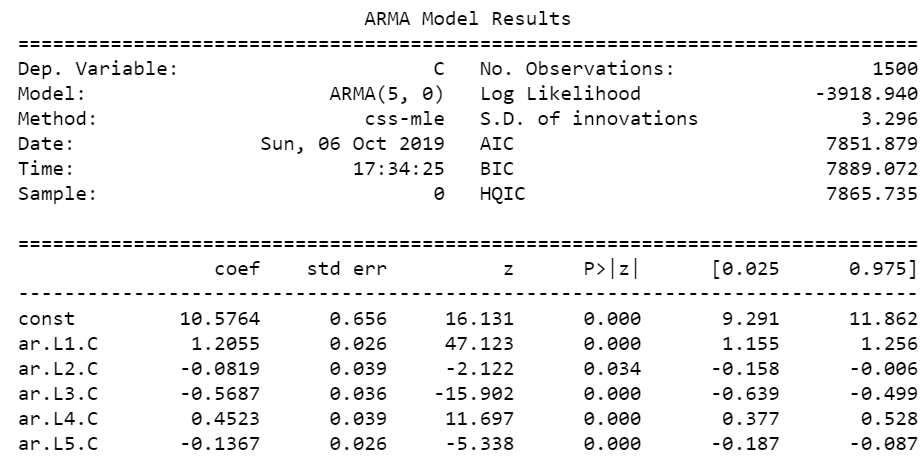
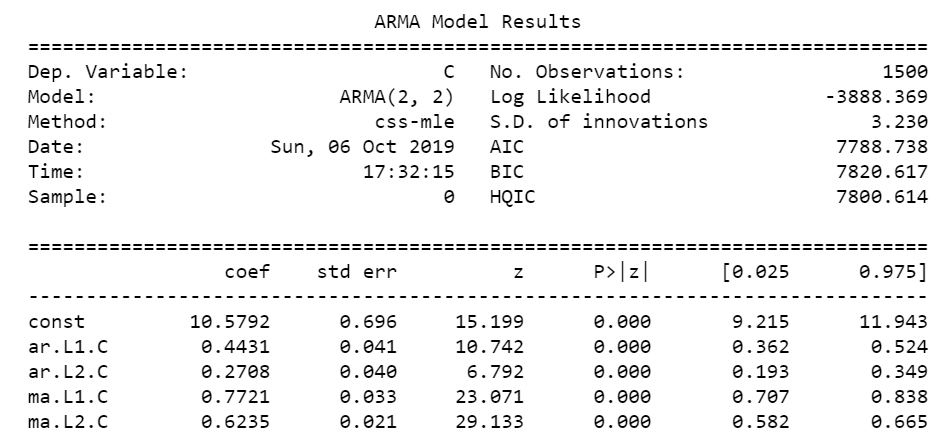


# Series C

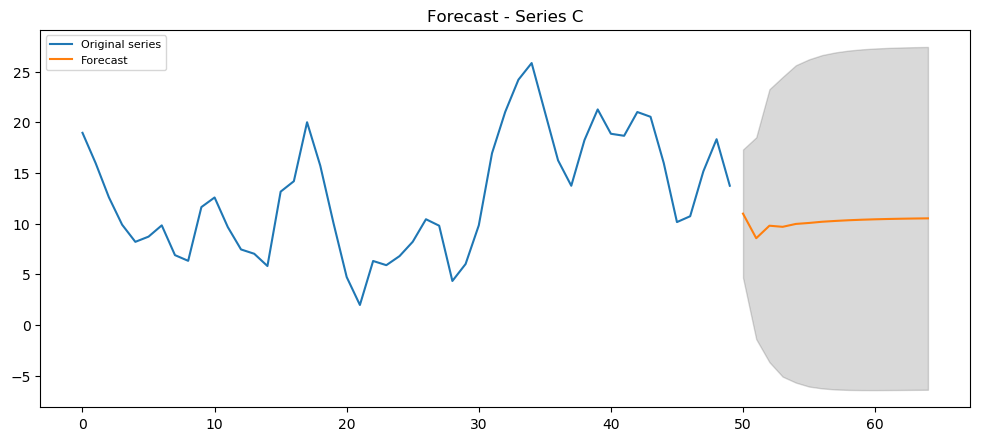


Similar to previous example, we see strong presence of PACF spikes that indicate AR process with geometrically decaying ACF. With 10 lags of significant PACF values at 5% confidence level, we will have to increase the maximal value of p when fitting semi-automated ARMA. Maximal values of p = 10 and q = 2 were chosen.

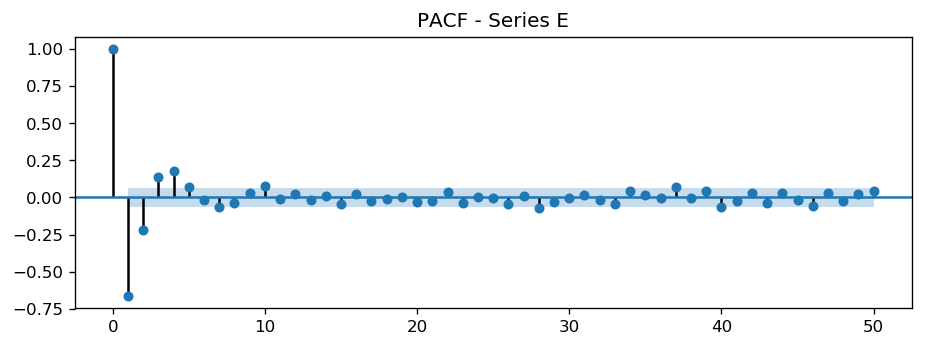
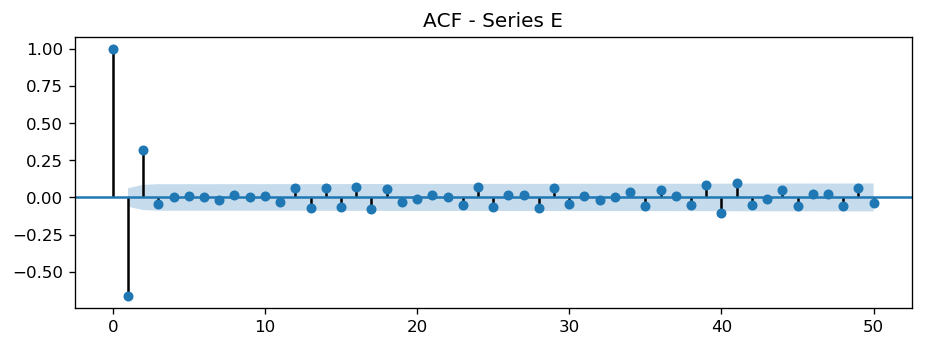
Unexpectedly, AIC criterion suggests going for ARMA(2, 2). Comparison against manual fit of ARMA(5, 0):



In spite of all significant lags from 1 to 5, there is strong autocorrelation observed in the residuals, pval = 0.00. Therefore, once again we stick to the automated selection, where pval = 0.97 for the same test statistics. Forecast from the selected model **ARMA(2, 2)**

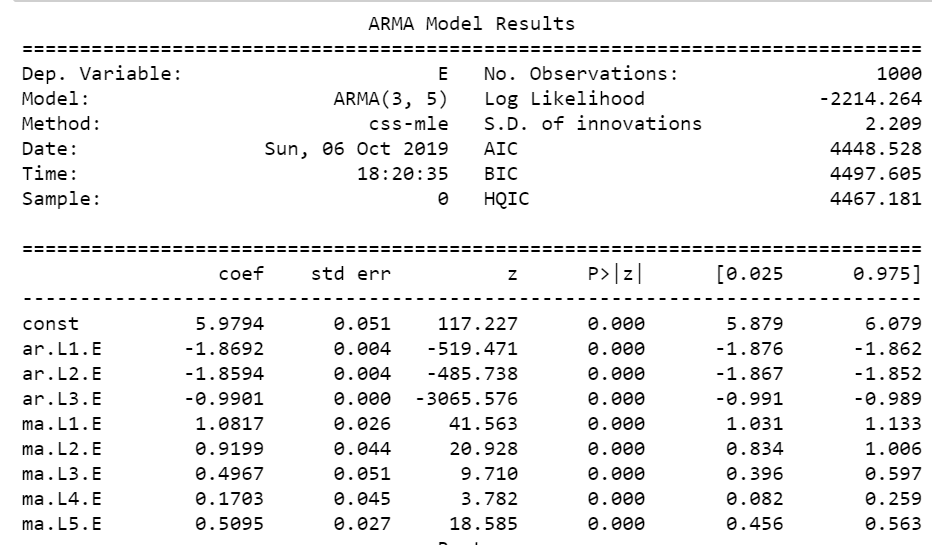
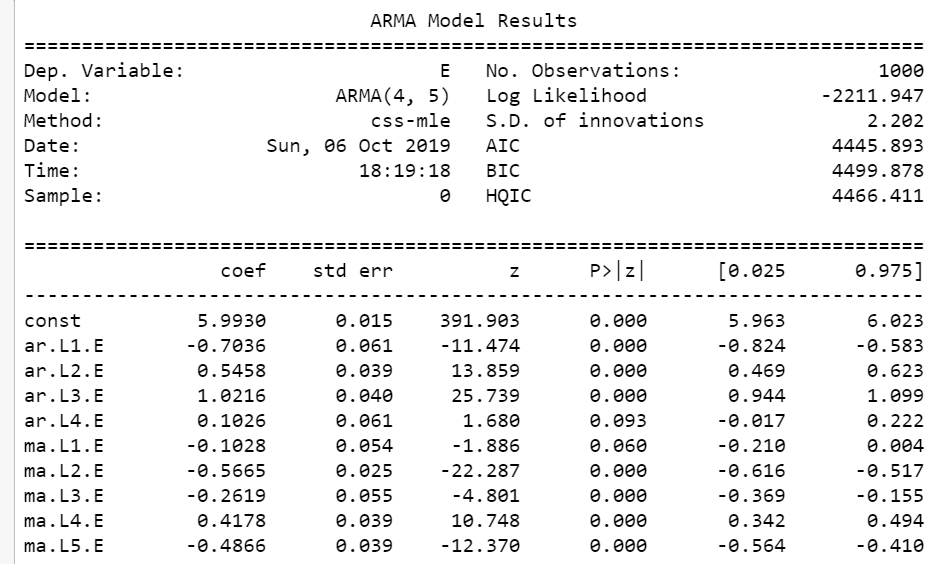


# Series E



For the first time within this assignment, we are observing a situation with spikes on both ACF and PACF plots. The number of spikes indicate possible efficiency with a model ARMA(2, 3) or ARMA(3, 3). Maximal values of 5 are selected for both parameters for automated search.

Minimal AIC criterion renders suggestion for model ARMA(4, 5). However, the latter results in insignificant component of fourth AR lag. I then remove the insignificant lag and leave the model at ARMA(3, 5) and all significant coefficients:



There is no autocorrelation present in the residuals, pval = 0.96. The final model is **ARMA(3, 5).**

