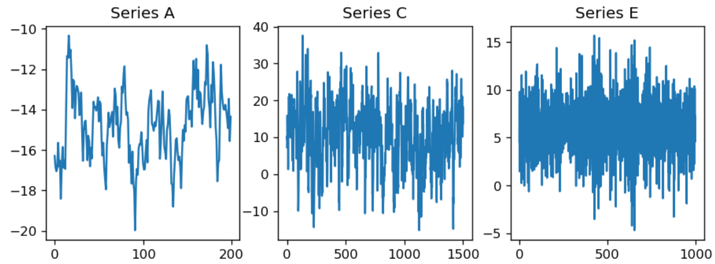
Code [will be] available @ https://github.com/szakrytnoy/financial-econometrics

1 Introduction 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.

2 Methodology

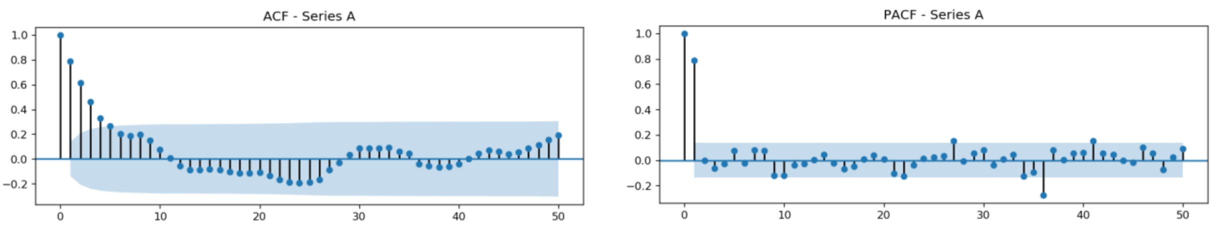
One of the common ways to determine optimal combinations of parameters (p, q) is to plot autocorrelation function (ACF) and partial autocorrelation function (PACF). The latter represents correlation coefficient between the original series and itself with lag = p and excludes the effects of lagged series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

3 Series A



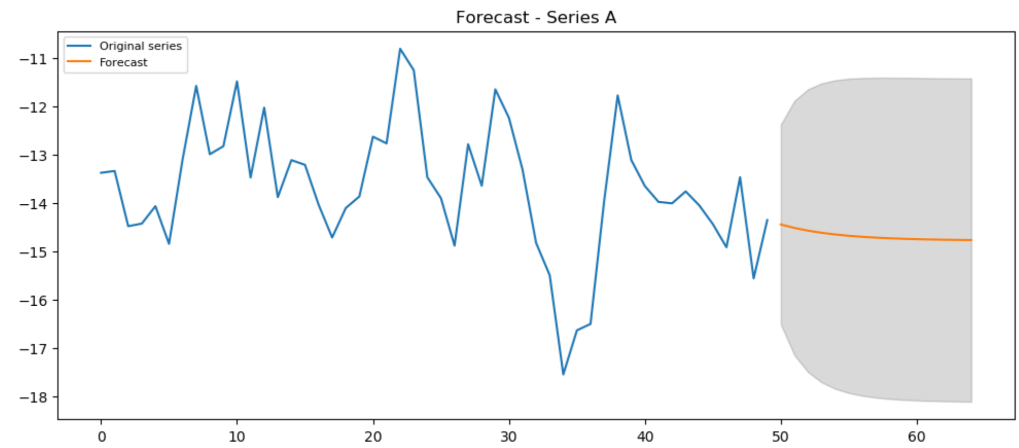
AR(2) process is clearly visible from the charts with 2 spikes in PACF and decaying ACF. However, when running an automated script with values of p and q from 0 to 3, the one that generates minimal AIC is ARMA(1, 0). Trying to override this choice with ARMA(2, 0) results in an insignificant coefficient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

```
                              ARMA Model Results
==============================================================================
Dep. Variable:                      A   No. Observations:                  200
Model:                     ARMA(1, 0)   Log Likelihood                -295.227
Method:                       css-mle   S.D. of innovations              1.056
Date:                Sun, 06 Oct 2019   AIC                            596.454
Time:                        17:10:29   BIC                            606.349
Sample:                             0   HQIC                           600.458

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const        -14.7821      0.342    -43.258      0.000     -15.452     -14.112
ar.L1.A        0.7853      0.043     18.167      0.000       0.701       0.870
```

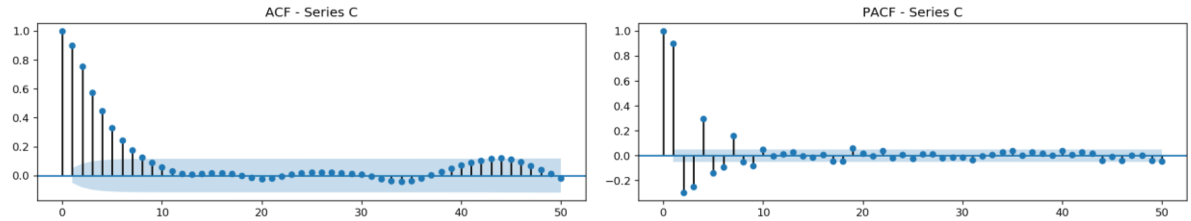
```
                              ARMA Model Results
==============================================================================
Dep. Variable:                      A   No. Observations:                  200
Model:                     ARMA(2, 0)   Log Likelihood                -295.227
Method:                       css-mle   S.D. of innovations              1.056
Date:                Sun, 06 Oct 2019   AIC                            598.454
Time:                        17:11:26   BIC                            611.647
Sample:                             0   HQIC                           603.793

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const        -14.7820      0.342    -43.258      0.000     -15.452     -14.112
ar.L1.A        0.7854      0.071     11.124      0.000       0.647       0.924
ar.L2.A       -0.0002      0.071     -0.002      0.998      -0.139       0.139
```

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.



4 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):

ARMA Model Results

Dep. Variable:	C	No. Observations:	1500
Model:	ARMA(2, 2)	Log Likelihood	-3888.369
Method:	css-mle	S.D. of innovations	3.230
Date:	Sun, 06 Oct 2019	AIC	7788.738
Time:	17:32:15	BIC	7820.617
Sample:	0	HQIC	7800.614

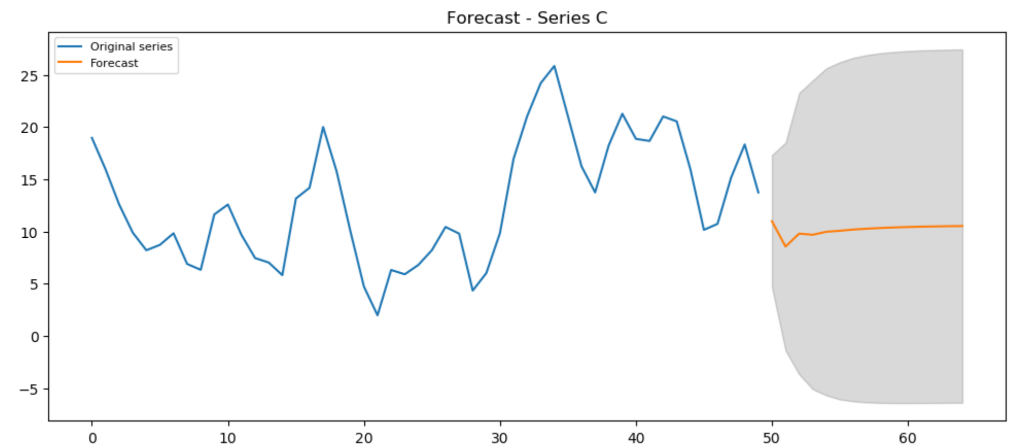
	coef	std err	z	P>|z|	[0.025	0.975]
const	10.5792	0.696	15.199	0.000	9.215	11.943
ar.L1.C	0.4431	0.041	10.742	0.000	0.362	0.524
ar.L2.C	0.2708	0.040	6.792	0.000	0.193	0.349
ma.L1.C	0.7721	0.033	23.071	0.000	0.707	0.838
ma.L2.C	0.6235	0.021	29.133	0.000	0.582	0.665

ARMA Model Results

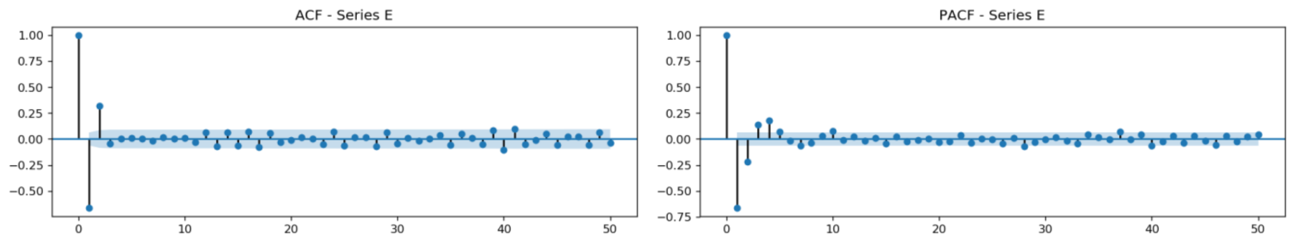
Dep. Variable:	C	No. Observations:	1500
Model:	ARMA(5, 0)	Log Likelihood	-3918.940
Method:	css-mle	S.D. of innovations	3.296
Date:	Sun, 06 Oct 2019	AIC	7851.879
Time:	17:34:25	BIC	7889.072
Sample:	0	HQIC	7865.735

	coef	std err	z	P>|z|	[0.025	0.975]
const	10.5764	0.656	16.131	0.000	9.291	11.862
ar.L1.C	1.2055	0.026	47.123	0.000	1.155	1.256
ar.L2.C	-0.0819	0.039	-2.122	0.034	-0.158	-0.006
ar.L3.C	-0.5687	0.036	-15.902	0.000	-0.639	-0.499
ar.L4.C	0.4523	0.039	11.697	0.000	0.377	0.528
ar.L5.C	-0.1367	0.026	-5.338	0.000	-0.187	-0.087

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)**



5 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:

```
                              ARMA Model Results
==============================================================================
Dep. Variable:                      E   No. Observations:                 1000
Model:                     ARMA(4, 5)   Log Likelihood               -2211.947
Method:                       css-mle   S.D. of innovations              2.202
Date:                Sun, 06 Oct 2019   AIC                           4445.893
Time:                        18:19:18   BIC                           4499.878
Sample:                             0   HQIC                          4466.411

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          5.9930      0.015    391.903      0.000       5.963       6.023
ar.L1.E       -0.7036      0.061    -11.474      0.000      -0.824      -0.583
ar.L2.E        0.5458      0.039     13.859      0.000       0.469       0.623
ar.L3.E        1.0216      0.040     25.739      0.000       0.944       1.099
ar.L4.E        0.1026      0.061      1.680      0.093      -0.017       0.222
ma.L1.E       -0.1028      0.054     -1.886      0.060      -0.210       0.004
ma.L2.E       -0.5665      0.025    -22.287      0.000      -0.616      -0.517
ma.L3.E       -0.2619      0.055     -4.801      0.000      -0.369      -0.155
ma.L4.E        0.4178      0.039     10.748      0.000       0.342       0.494
ma.L5.E       -0.4866      0.039    -12.370      0.000      -0.564      -0.410
```

```
                              ARMA Model Results
==============================================================================
Dep. Variable:                      E   No. Observations:                 1000
Model:                     ARMA(3, 5)   Log Likelihood               -2214.264
Method:                       css-mle   S.D. of innovations              2.209
Date:                Sun, 06 Oct 2019   AIC                           4448.528
Time:                        18:20:35   BIC                           4497.605
Sample:                             0   HQIC                          4467.181

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          5.9794      0.051    117.227      0.000       5.879       6.079
ar.L1.E       -1.8692      0.004   -519.471      0.000      -1.876      -1.862
ar.L2.E       -1.8594      0.004   -485.738      0.000      -1.867      -1.852
ar.L3.E       -0.9901      0.000  -3065.576      0.000      -0.991      -0.989
ma.L1.E        1.0817      0.026     41.563      0.000       1.031       1.133
ma.L2.E        0.9199      0.044     20.928      0.000       0.834       1.006
ma.L3.E        0.4967      0.051      9.710      0.000       0.396       0.597
ma.L4.E        0.1703      0.045      3.782      0.000       0.082       0.259
ma.L5.E        0.5095      0.027     18.585      0.000       0.456       0.563
```

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