Introduction to Forecasting Models

Exercise 4: ARMA Models

Adam Cihlář[1]

13. March 2022

Abstract: The goal of this exercise is to estimate models for predicting the time series of interest rate of 10 year Norwegian state bond. We employ several ARMA models that we construct following the Box-Jenkins methodology. The models' performance is compared based on prediction metrics and properties of their prediction errors. All proposed models perform very similarly out-of-sample. Because of simplicity, we suggest using AR 1 process for modeling the time series.

Keywords: ARMA Models, Time Series Prediction, State Bonds

Introduction

Purpose of this exercise is to construct various ARMA models to predict Long-Term Government Bond Yields (10-year) for Norway. Firstly, we visualize the data and inspect the stationarity, then we split them to training and testing sets to save a portion for the out-of-sample predictions and evaluation of the models. Secondly, we define and estimate four models using the Box-Jenkins methodology and compare them based on how they can fit the data. Thirdly, the one-step out-of-sample predictions are performed and final comparison of the models is conducted based on visual analysis and properties of their prediction errors and of course based on the ability to predict future values - forecast metrics. The whole analysis is implemented in statistical software R.

[1]Masaryk University, Faculty of Economics and Administration, Field: Mathematical and Statistical Methods in Economics, 468087@mail.muni.cz

1 Data

In our modelling we will use only the time series of the interest rate for its prediction. The observed time series is Long-Term (10-year) Government Bond Yields for Norway (IRLTLT01NOM156N) as reported by Federal Reserve Bank of St. Louis.

Figure 1: Long-Term (10-year) Government Bond Yields for Norway (IRLTLT01NOM156N)

The data are available from 1. 1. 1985 to 1. 12. 2021 in monthly frequency. The models will be estimated on sample from 1. 1. 1985 to 1. 12. 2015 which is approximately 84 % of the full data set, that leaves us with one sixth of data for the testing purposes - out-of-sample predictions.

However, firstly, we will inspect the autocorrelation function to check the stationarity of the time series. As the figure 2 shows, there is very strong positive autocorrelation in the data reaching 0.991 for the first lag, so we are going to work with differences to make the series stationary.

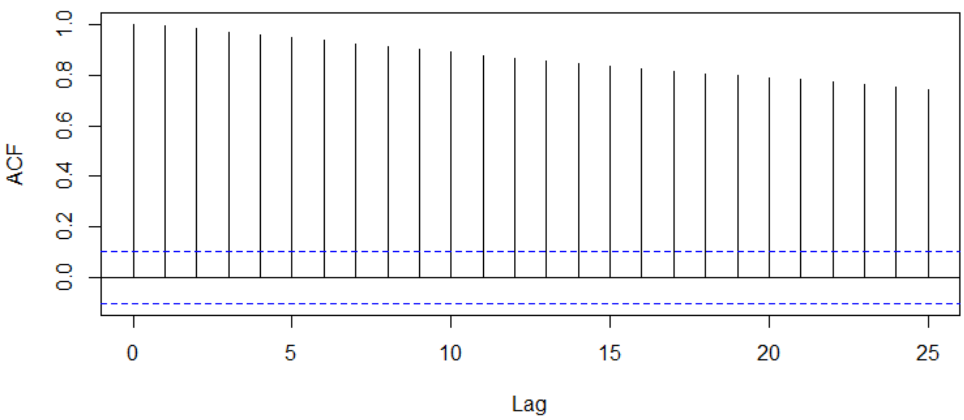

Figure 2: Autocorrelation of original time series

After detrending we get the following time series (figure 3). With regards to its autocorrelation and partial autocorrelation functions we consider it to be stationary and suitable to be modeled by ARMA processes.

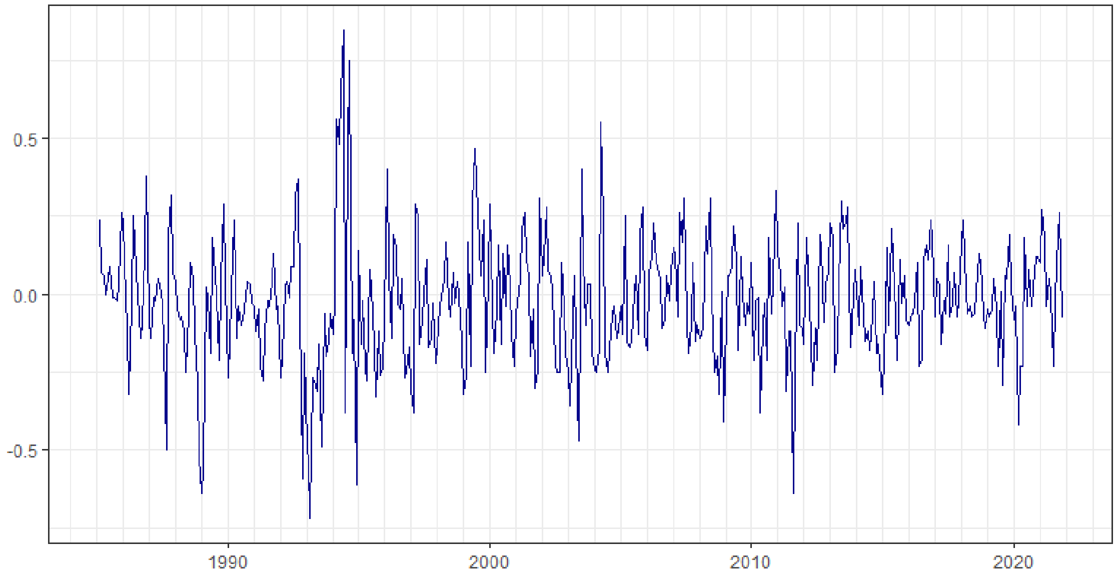

Figure 3: Month-to-month differences

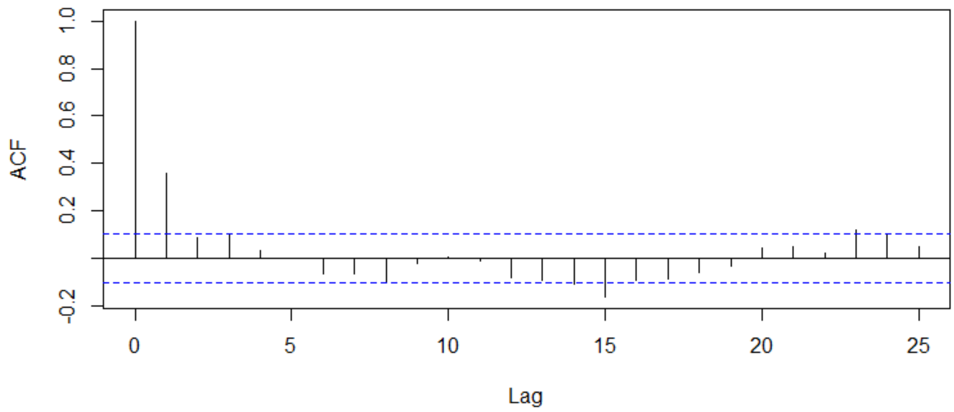

Figure 4: Autocorrelation of stationary series

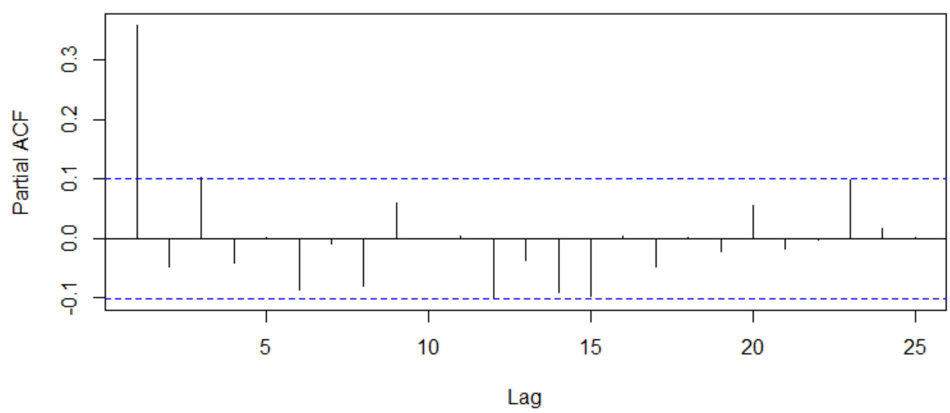

Figure 5: Partial autocorrelation of stationary series

2 Model selection and estimates

According to Box-Jenkins methodology, we first estimate several AR models with orders based on the autocorrelation functions:

- AR 1
- AR $\{1, 3\}$
- AR $\{1, 3, 12\}$

	Dependent variable:		
	Interest Rate (IR)		
	(1)	(2)	(3)
IR_{t-1}	0.360^{***}	0.354^{***}	0.353^{***}
	(0.048)	(0.048)	(0.048)
IR_{t-3}		0.075^{***}	0.073
		(0.048)	(0.048)
IR_{t-12}			-0.073^{***}
			(0.048)
intercept	-0.029^{*}	-0.029	-0.030
	(0.016)	(0.018)	(0.016)
Observations	371	371	371
Log Likelihood	75.48	76.66	77.81
σ^2	0.039	0.039	0.038
Akaike Inf. Crit.	-144.96	-145.33	-145.62
Bayes Inf. Crit.	-133.21	-129.66	-126.04
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 1: AR models

The more lagged variables we consider, the lower Akaike Information Criterion we get but the differences are almost negligible. However, the Bayes Information Criterion clearly prefers the first and most simple model - AR 1, so we will next analyse residuals of this model to determine the MA component.

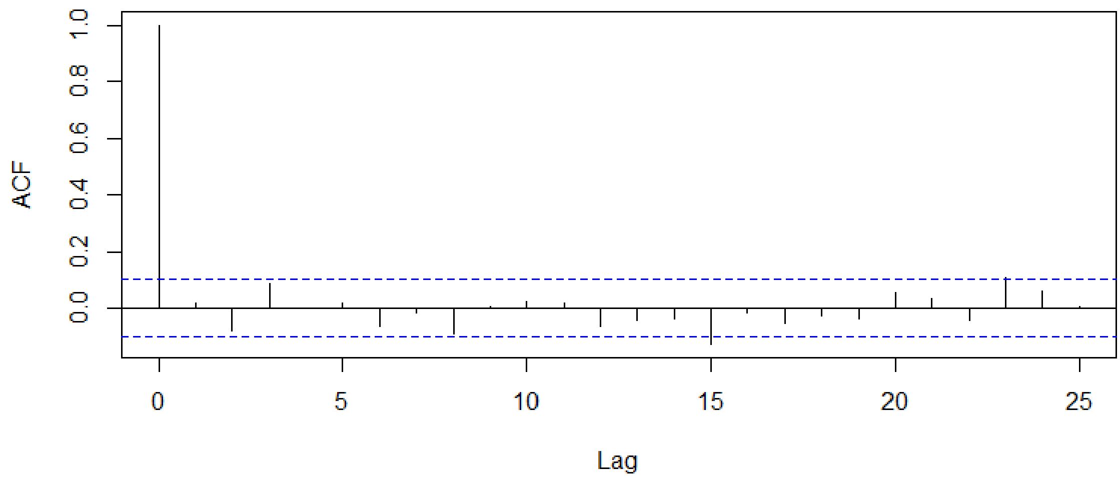

Figure 6: AR 1 residuals - autocorrelation

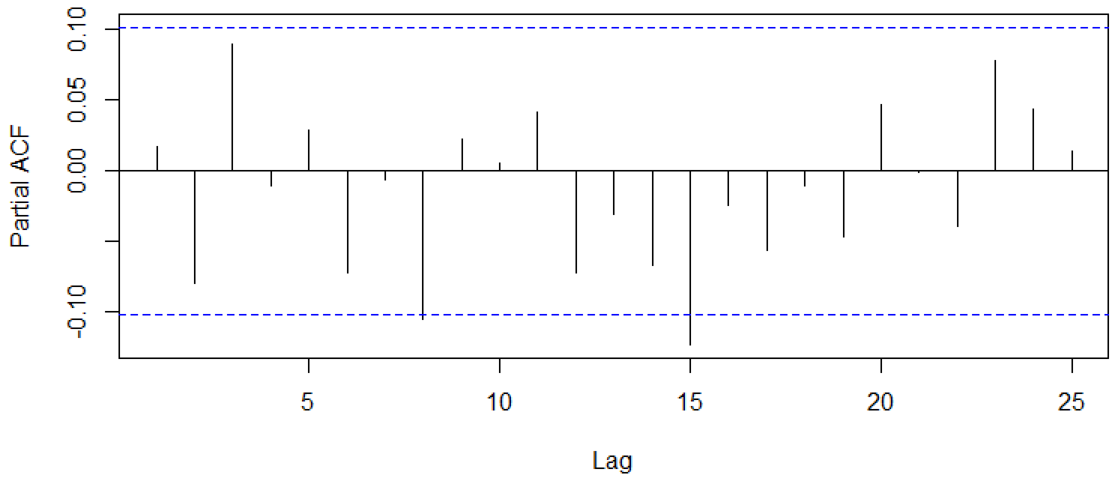

Figure 7: AR 1 residuals - partial autocorrelation

Both autocorrelation functions suggest that there might be an unexplained component for 8 and 15 lags. We define and estimate following four models for the final analysis and out-of-sample predictions:

- AR 1
- AR 1 MA {8}
- AR 1 MA {15}
- AR 1 MA {8, 15}

	Dependent variable:			
	Interest Rate (IR)			
	(1)	(2)	(3)	(4)
IR_{t-1}	0.360^{***}	0.357^{***}	0.349^{***}	0.347^{***}
	(0.048)	(0.049)	(0.049)	(0.049)
ε_{t-8}		-0.091^{***}		-0.074
		(0.052)		(0.053)
ε_{t-15}			-0.127^{***}	-0.119^{***}
			(0.051)	(0.052)
intercept	-0.029^{*}	-0.029	-0.030	-0.030
	(0.016)	(0.014)	(0.014)	(0.013)
Observations	371	371	371	371
Log Likelihood	75.483	76.992	78.577	79.545
σ^2	0.039	0.039	0.038	0.038
Akaike Inf. Crit.	-144.965	-145.984	-149.154	-149.091
Bayes Inf. Crit.	-133.216	-130.320	-133.489	-129.510
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2: ARMA models

The table 2 shows the estimates of all four models. Two best models according to fit metrics are AR 1 and AR 1 MA $\{15\}$ but as we are mainly interested in the prediction performance, we will use all four models to compute out-of-sample predictions.

3 Predictions

In this section we will predict the values of the Norwegian 10-Year bond interest rate in time $t+1$ - one-step predictions - for period from 1. 1. 2016 to 1. 12. 2021 using the four estimated models. Let's have a look on various prediction metrics and plot the predicted and true values.

	AR1	AR1 MA{8}	AR1 MA{15}	AR1 MA{8, 15}
MAE	0.104	0.102	0.107	0.106
RMSE	0.131	0.131	0.134	0.135
MAPE	8.5%	8.4%	8.8%	8.7%

Table 3: Prediction Metrics

All the prediction metrics (table 3) are very similar and we would need to employ a statistical test (e.g. Diebold-Mariano test) to assess if there is any statistically significant difference but that is out of scope of this exercise. Neither from the plot of predicted vs true values 8 are we able to clearly select the top performing model.

Figure 8: Predictions

4 Specification tests

Ultimately, we will have a closer look on the prediction errors and test if the specifications of the models are proper. This will be evaluated by prediction errors distribution and test of their normality.

	AR1	AR1 MA{8}	AR1 MA{15}	AR1 MA{8, 15}
Statistic	2.067	1.422	1.244	0.946
p-value	0.355	0.491	0.536	0.622

Table 4: Jarque-Bera test of normality of prediction errors

For any of the models we do not reject the null hypothesis that the residuals are normally distributed.

If we plot the prediction errors' distributions we can see that the dark red long-dashed line marking the mean of the errors lies below in all four cases. That means that all models are systematically underpredicting the true values. This may be caused by the fact that the models were trained on period when the interest rate was mostly declining (negative mean difference) while the predictions are performed in a period when the observed value stays almost on the same level (see figure 8).

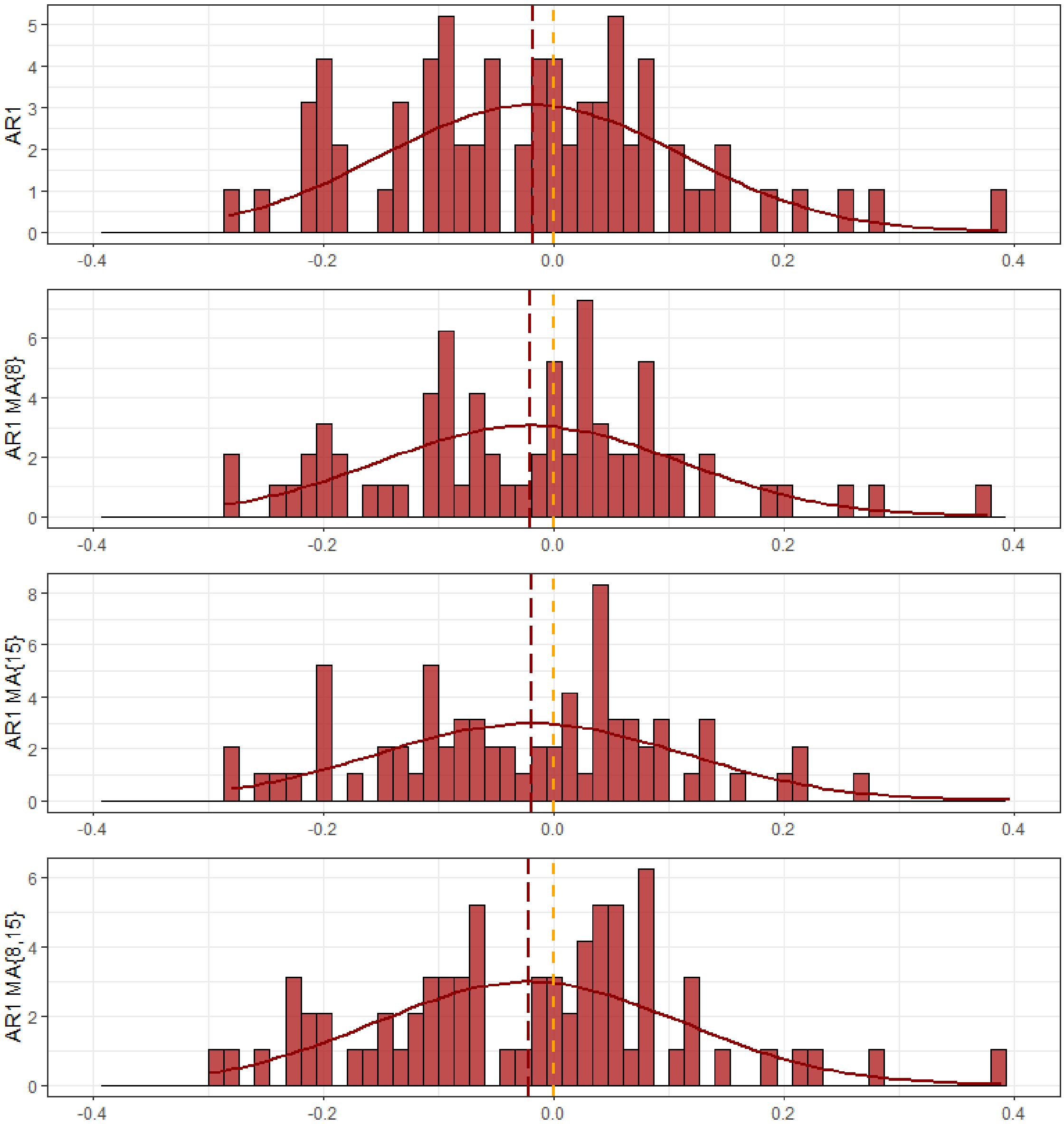

Figure 9: Prediction errors distributions

Conclusion

In this exercise we constructed four various AR/ARMA models to predict Long-Term Government Bond Yields (10-year) for Norway. The only data used for the prediction was the time series itself. We employed Box-Jenkins methodology to estimate the ARMA models. The data was available from 1. 1. 1985 to 1. 12. 2021 in monthly frequency.

Firstly, we visualized the time series, transformed it to stationary data using differences and split the it to save a portion for the out-of-sample predictions and evaluation of the models. Secondly, we specified and estimated the models based on autocorrelation and partial autocorrelation functions and compared them based on how they can fit the data, based on properties of their prediction errors and of course based on their ability to predict future values - forecast metrics.

All models are systematically underpredicting the out-of-sample values and all performed very similarly so we would choose the AR 1 model as it is the simplest one.