ARMA estimation

Giovani Valdrighi, Vitória Guardieiro

05/11/2020

Modelling process

- Plot the data and look for patters.
- If necessary, use BoxCox transformation to stabilize the variance.
- If necessary, remove seasonality.
- If necessary, difference the data until is stationary. Use of Dickey Fuller test.
- Plot the ACF and PACF to identify the model order, p and q for $ARMA(p, q)$.
- Compare identified models, chose the one that minimize the AIC
- Analysis of the residuals of the model, with ACF and histogram.
- If the residuals look like white noise, make forecasts.

WMurders

Monthly female murder rate
(per 100,000 standard population)
Rate
4.5
4.0
3.5
3.0
2.5
1950
1960
1970
1980
1990
2000
Time

Variance stabilization and stationarity

- There is no need to stabilize data, we can look at the tendency with Dick-Fulley.

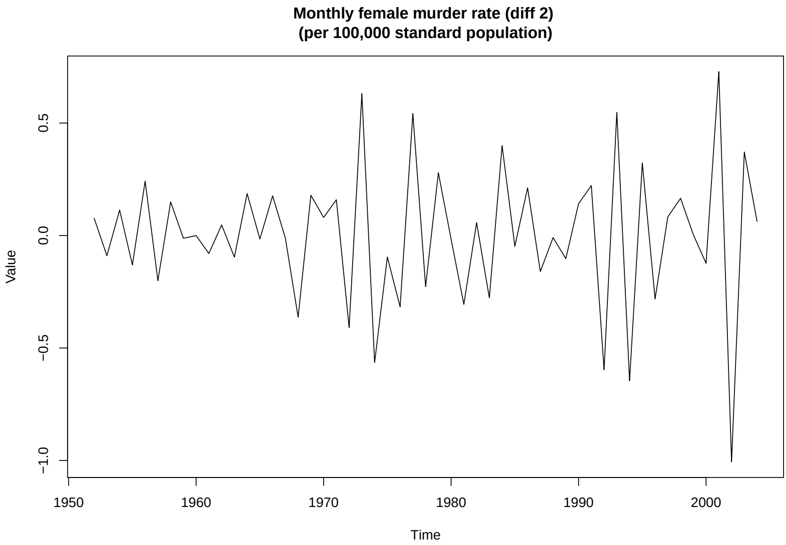
```
##
##  Augmented Dickey-Fuller Test
##
## data:  data
## Dickey-Fuller = -0.29243, Lag order = 3, p-value = 0.987
## alternative hypothesis: stationary
```

- Data is not stationary, we are going to test with one and two differences.

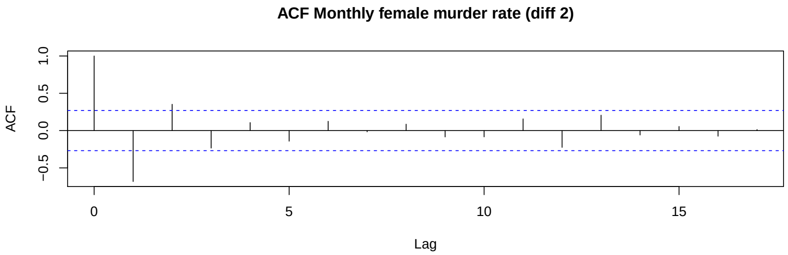
```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(data)
## Dickey-Fuller = -3.7688, Lag order = 3, p-value = 0.0272
## alternative hypothesis: stationary
```

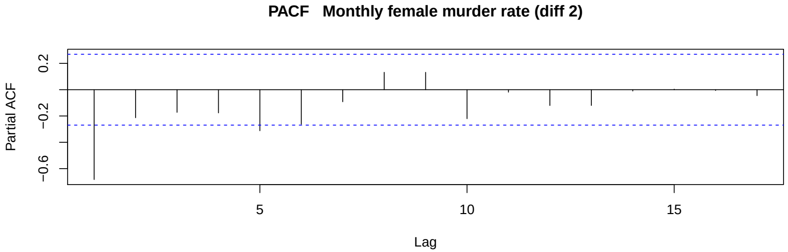
```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(data, differences = 2)
## Dickey-Fuller = -5.1646, Lag order = 3, p-value = 0.01
## alternative hypothesis: stationary
```

- Diff 1 P-value: 0.02726
- Diff 2 P-value: less than 0.01 <- choosen.


Monthly female murder rate (diff 2)
(per 100,000 standard population)
Value
Time
0.5
0.0
−0.5
−1.0
1950
1960
1970
1980
1990
2000

ACF and PACF

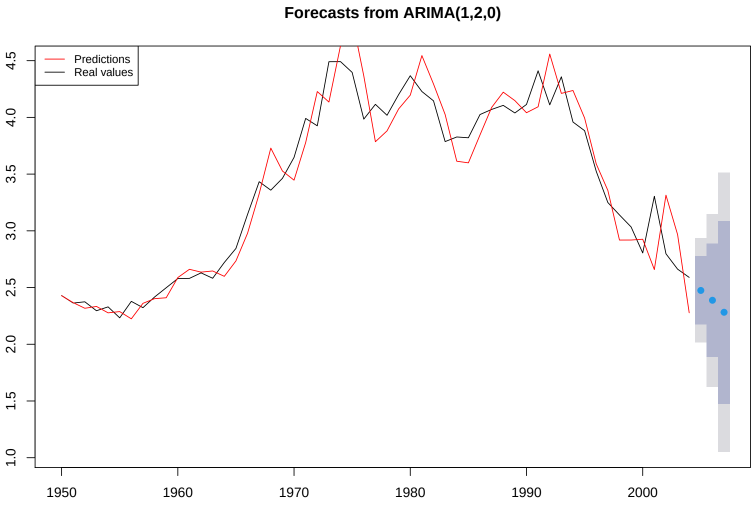




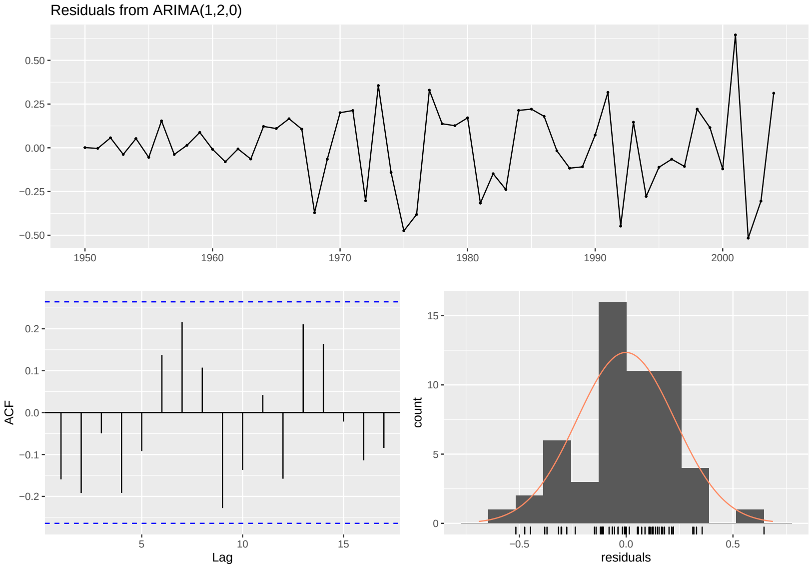
- Looks like AR(1) because the first spike of the PACF. The models that will be tested are AR(1), MA(1), ARMA(1, 1).

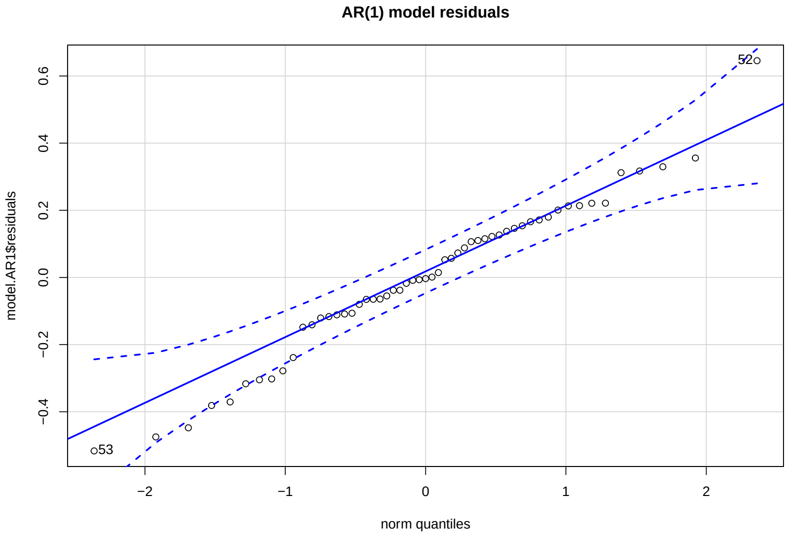
AR(1)

```
## Series: data
## ARIMA(1,2,0)
##
## Coefficients:
##           ar1
##       -0.6719
## s.e.   0.0981
##
## sigma^2 estimated as 0.05471:  log likelihood=2
## AIC=0   AICc=0.24   BIC=3.94
##
## Training set error measures:
##                        ME      RMSE       MAE         MP
## Training set -0.001376898 0.2274352 0.1777919 0.0014867
##                    ACF1
## Training set -0.1593845
```


Forecasts from ARIMA(1,2,0)
Predictions
Real values
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
1950
1960
1970
1980
1990
2000

Residuals analysis



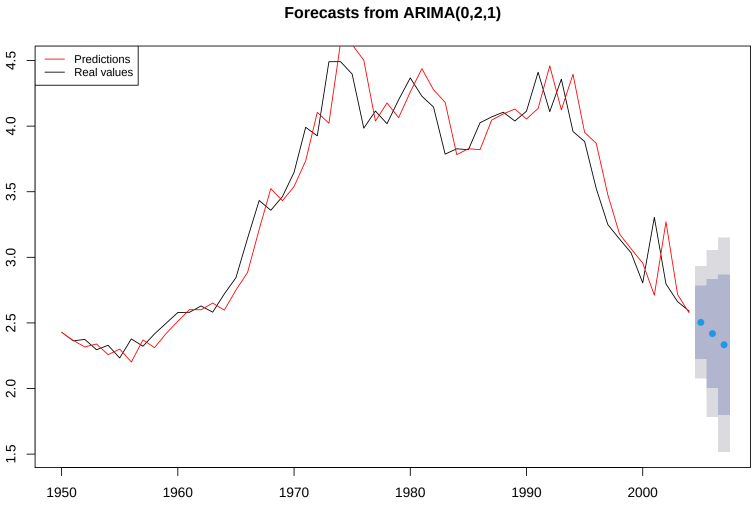

AR(1) model residuals
model.AR1$residuals
norm quantiles
0.6
0.4
0.2
0.0
−0.2
−0.4
−2
−1
0
1
2
52
53

```
## [1] 52 53
```

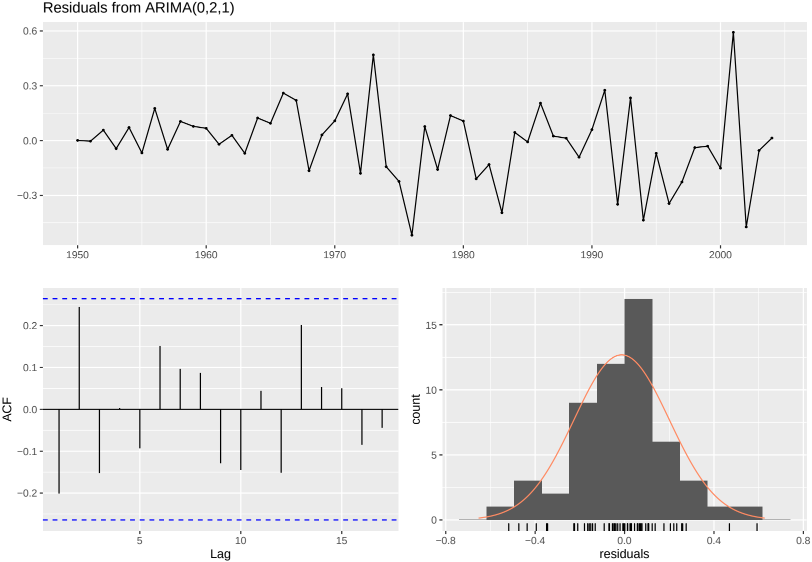
```
##
##  Jarque Bera Test
##
## data:  model.AR1$residuals
## X-squared = 0.24457, df = 2, p-value = 0.8849
```

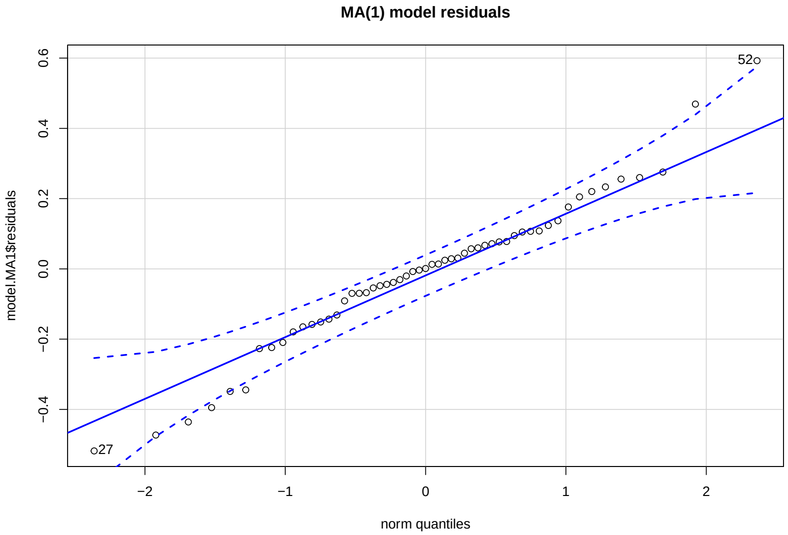
MA(1)

```
## Series: data
## ARIMA(0,2,1)
##
## Coefficients:
##           ma1
##       -0.8995
## s.e.   0.0669
##
## sigma^2 estimated as 0.04747:  log likelihood=5.24
## AIC=-6.48   AICc=-6.24   BIC=-2.54
##
## Training set error measures:
##                       ME      RMSE       MAE        MPE
## Training set -0.01306101 0.2118445 0.1559694 -0.3151353
##                    ACF1
## Training set -0.2011523
```


Forecasts from ARIMA(0,2,1)
Predictions
Real values
1.5
2.0
2.5
3.0
3.5
4.0
4.5
1950
1960
1970
1980
1990
2000

Residuals analysis



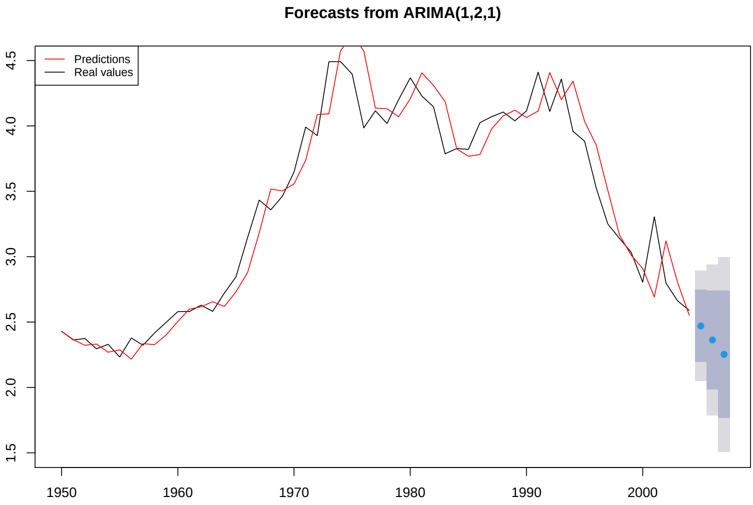

MA(1) model residuals
model.MA1$residuals
norm quantiles
0.6
0.4
0.2
0.0
−0.2
−0.4
−2
−1
0
1
2
52
27

```
## [1] 52 27
```

```
##
##  Jarque Bera Test
##
## data:  model.MA1$residuals
## X-squared = 1.3749, df = 2, p-value = 0.5028
```

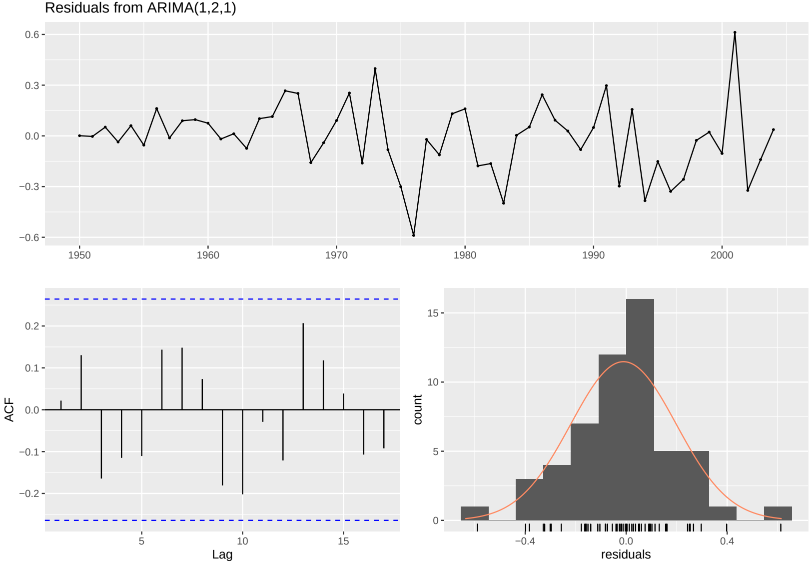
ARMA(1, 1)

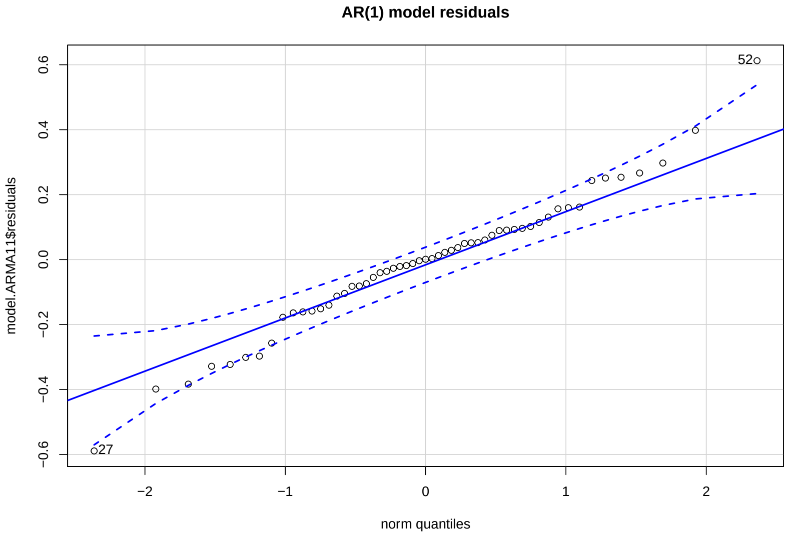
```
## Series: data
## ARIMA(1,2,1)
##
## Coefficients:
##           ar1      ma1
##       -0.2434  -0.8261
## s.e.   0.1553   0.1143
##
## sigma^2 estimated as 0.04632:  log likelihood=6.44
## AIC=-6.88   AICc=-6.39   BIC=-0.97
##
## Training set error measures:
##                       ME      RMSE       MAE        MPE
## Training set -0.01065956 0.2072523 0.1528734 -0.2149476
##                    ACF1
## Training set 0.02176343
```



Forecasts from ARIMA(1,2,1)
Predictions
Real values
1.5
2.0
2.5
3.0
3.5
4.0
4.5
1950
1960
1970
1980
1990
2000

Residuals analysis




AR(1) model residuals
model.ARMA11$residuals
norm quantiles
52
27
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−2
−1
0
1
2

```
## [1] 52 27
```

```
##
##  Jarque Bera Test
##
## data:  model.ARMA11$residuals
## X-squared = 2.3784, df = 2, p-value = 0.3045
```

Model selection

- The model with lowest AIC is the $ARMA(1,1)$. The final forecast is: