Homework 3

Nick Climaco

February 12, 2024

Table of contents

Chapter 5: The Forecaster's Toolbox

Exercise 1

Produce forecasts for the following series using whichever of NAIVE(y), SNAIVE(y) or RW(y ~ drift()) is more appropriate in each case:

- Australian Population (global_economy)
- Bricks (aus_production)
- NSW Lambs (aus_livestock)
- Household wealth (hh_budget).
- Australian takeaway food turnover (aus_retail).

```
df_aus_prod = pd.read_csv("../rdata/global_economy.csv", parse_dates=["Year"],
  index_col=['Year'])
df_aus_prod
```

Year	Unnamed: 0	Country	Code	...	Imports	Exports	Population
1960-01-01	1	Afghanistan	AFG	...	7.024793	4.132233	8996351.0
1961-01-01	2	Afghanistan	AFG	...	8.097166	4.453443	9166764.0
1962-01-01	3	Afghanistan	AFG	...	9.349593	4.878051	9345868.0
1963-01-01	4	Afghanistan	AFG	...	16.863910	9.171601	9533954.0
1964-01-01	5	Afghanistan	AFG	...	18.055555	8.888893	9731361.0
...	...	...	...	...	...	...	...
2013-01-01	15146	Zimbabwe	ZWE	...	36.668735	21.987759	15054506.0
2014-01-01	15147	Zimbabwe	ZWE	...	33.741470	20.930146	15411675.0
2015-01-01	15148	Zimbabwe	ZWE	...	37.588635	19.160176	15777451.0
2016-01-01	15149	Zimbabwe	ZWE	...	31.275493	19.943532	16150362.0
2017-01-01	15150	Zimbabwe	ZWE	...	30.370273	19.658023	16529904.0

Exercise 2

Use the Facebook stock price (data set gafa_stock) to do the following:

- Produce a time plot of the series.
- Produce forecasts using the drift method and plot them.
- Show that the forecasts are identical to extending the line drawn between the first and last observations.
- Try using some of the other benchmark functions to forecast the same data set. Which do you think is best? Why?

Exercise 3

Apply a seasonal naïve method to the quarterly Australian beer production data from 1992. Check if the residuals look like white noise, and plot the forecasts. The following code will help.

```
# Extract data of interest
recent_production <- aus_production |>
  filter(year(Quarter) >= 1992)

# Define and estimate a model
fit <- recent_production |> model(SNAIVE(Beer))
#Look at the residuals

fit |> gg_tsresiduals()

# Look a some forecasts
fit |> forecast() |> autoplot(recent_production)
```

What do you conclude?

Exercise 4

Repeat the previous exercise using the Australian Exports series from global_economy and the Bricks series from aus_production. Use whichever of NAIVE() or SNAIVE() is more appropriate in each case

Exercise 7

For your retail time series (from Exercise 7 in Section 2.10):

a. Create a training dataset consisting of observations before 2011 using

```
myseries_train <- myseries |>
  filter(year(Month) < 2011)
```

b. Check that your data have been split appropriately by producing the following plot.

```
autoplot(myseries, Turnover) +
  autolayer(myseries_train, Turnover, colour = "red")
```

c. Fit a seasonal naïve model using SNAIVE() applied to your training data (myseries_train).

```
fit <- myseries_train |>
  model(SNAIVE())
```

d. Check the residuals.

```
fit |> gg_tsresiduals()
```

Do the residuals appear to be uncorrelated and normally distributed?

e.Produce forecasts for the test data

```
fc <- fit |>
  forecast(new_data = anti_join(myseries, myseries_train))
fc |> autoplot(myseries)
```

f. Compare the accuracy of your forecasts against the actual values.

```
fit |> accuracy()
fc |> accuracy(myseries)
```

g.How sensitive are the accuracy measures to the amount of training data used?