ETF3231/5231 Week 4 - Basic time series modelling, forecasting and evaluation

Sao Yang Hew

Friday - 10 March 2023 (12:30pm - 2:00pm)

2023

Table of Contents

Simple forecasting methods

- Here are some simple forecasting methods you would have learnt in the lectures (or Chapter 5):
  - Mean
  - Naive, or random walk
  - Seasonal naive
  - Drift, or random walk with drift
  - Seasonal naive with drift
- Why do we use these methods?
  - These methods are useful benchmarks to assess if complex models we formulate are appropriate
  - Sometimes, these forecasts are the best we can do at least mathematically (stock returns)
- Note that forecasting requires **business and contextual understanding**, these methods or methods we learn after may help - but are usually done alongside an understanding of the underlying item you are trying to forecast

Important functions

- Model (creates `mable`):
  - `fit <- table |> model(SNAIVE(variable))`
  - `fit <- table |> model(RW(variable))`
  - `fit <- table |> model(RW(variable ~ drift()))`
  - `fit <- table |> model(SNAIVE(variable ~ drift()))`
- Forecast (creates `fable`) and plot:
  - `fit |> forecast(h = '1 year') |> autoplot()`

Fitted values, residual diagnostics

Once we fit a model, we want to have a look at the residual components of the model (predicted value - actual value).

- Why do we assess model residuals?
  - To assess how well we have done in our modelling step.
  - If our residuals don't show any predictable patterns, it means that we picked up the necessary patterns in our model already.
- Essential properties of residuals (these allow production of good forecasts):
  - Zero mean
  - Uncorrelated across time
- Desirable (but not essential) properties of residuals (these make it easier to produce distributional forecasts to represent uncertainty):
  - Normally distributed
  - Homoscedastic

Evaluating point forecast accuracy

- Good forecasts can be defined by 2 metrics: point forecast accuracy (are our mean forecasts close to actual values?), and distributional forecast accuracy (how well did we characterise our uncertainty? are our prediction intervals narrow enough?)
- The main forecasting metrics we need to know are:
  - Point forecast metrics
    - MAE
    - RMSE
    - MAPE
    - MASE
  - Time-series cross validation
  - Distributional forecast metrics (which we don't focus in this tutorial)
    - Quantile score
    - Winkler score
    - CRPS
    - Skill score

Important functions

- Assessing residuals and accuracy (from a mable):
  - `fit |> gg_tsresiduals()`
- Train-test split forecast accuracy:
  1. `train <- table |> filter(...)`
  2. `fit <- train |> model(...)`
  3. `forecast <- fit |> forecast(...)`
  4. `forecast |> accuracy(table) |> arrange(...)`
- Time series cross-validation:
  1. ```
     train <- table |>
       stretch_tsibble(.init = init, .step = step) |>
       relocate(index, key, .id)
     ```
  2. ```
     train |> model(...)  |> forecast(...)  |>
     accuracy(table)
     ```