Forecasting I

Bissan Ghaddar

Technical University of Denmark

42586 – Decisions under Uncertainty – 12 March 2024

Learning Objectives

At the end of this lecture (and related hands-on session), a student should be able to:

- Understand the three types of **exponential smoothing**
- Describe **ARIMA** models for times series
- Apply and evaluate different **forecasting models** on a given dataset

Outline

1. Recap
2. Holts-Winters Exponential Smoothing
3. AR and MA models
4. ARIMA models
5. Evaluating forecasting models

Seasonal Data with Trend: Holt Winter's Method

- The Holt-Winters seasonal method comprises the forecast equation and three smoothing equations
  - one for the level L_t,
  - one for trend B_t, and
  - one for the seasonal component denoted by S_t,
- with smoothing parameters α, β and γ.
- The parameter m denotes the period of the seasonality, i.e., the number of seasons in a year. For example, for quarterly data $m = 4$, and for monthly data $m = 12$.

Seasonal Series with Trend: Holt-Winter's Method

There are two variations to this method that differ in the nature of the seasonal component.

- The additive method is preferred when the seasonal variations are roughly constant through the series,
  - the seasonal component is expressed in absolute terms in the scale of the observed series,
  - in the level equation the series is seasonally adjusted by subtracting the seasonal component
  - within each year the seasonal component will add up to approximately zero
- the multiplicative method is preferred when the seasonal variations are changing proportional to the level of the series.
  - the seasonal component is expressed in relative terms (percentages)
  - the series is seasonally adjusted by dividing through by the seasonal component
  - within each year the seasonal component will add up to approximately m

We will focus on additive methods for this course.

Time Series: Holt-Winter's Method - Additive

The component form for the additive method is:

The series

$$L_t = \alpha(Y_t - S_{t-m}) + (1-\alpha)(L_{t-1} + B_{t-1})$$

The trend

$$B_t = \beta(L_t - L_{t-1}) + (1-\beta)B_{t-1}$$

Seasonal factors

$$S_t = \gamma(Y_t - L_{t-1} - B_{t-1}) + (1-\gamma)S_{t-m}$$

The forecast made in period t for any future period $t+\tau$ ($\tau \leq m$) is given by:

$$\hat{Y}_{t,t+\tau} = L_t + \tau B_t + S_{t-m+\tau_m}$$

where $\tau_m = \lfloor(\tau - 1) \mod m\rfloor + 1$.

- The level equation shows a weighted average between the seasonally adjusted observation $(Y_t - S_{t-m})$ and the non-seasonal forecast $(L_{t-1} + B_{t-1})$ for time t.
- The trend equation is identical to Holt's linear method.
- The seasonal equation shows a weighted average between the current seasonal index, $(Y_t - L_{t-1} - B_{t-1})$, and the seasonal index of the same season last year (i.e., m time periods ago).

Time Series: Holt-Winter's Method - Multiplicative

The component form for the multiplicative method is:

The series

$$L_t = \alpha \frac{Y_t}{S_{t-m}} + (1-\alpha)(L_{t-1} + B_{t-1})$$

The trend

$$B_t = \beta(L_t - L_{t-1}) + (1-\beta)B_{t-1}$$

Seasonal factors

$$S_t = \gamma \frac{Y_t}{L_{t-1} + B_{t-1}} + (1-\gamma)S_{t-m}$$

The forecast made in period t for any future period $t+\tau$ $(\tau \leq m)$ is given by:

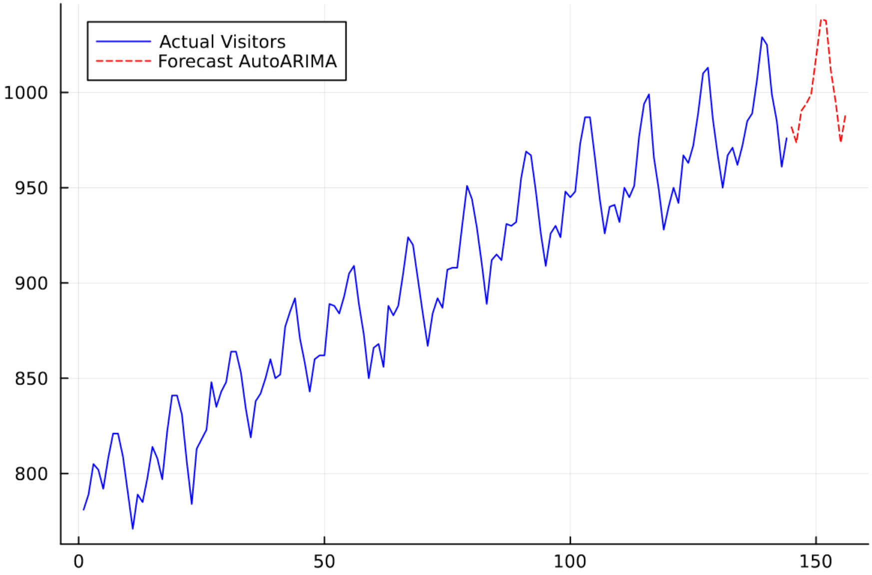
$$\hat{Y}_{t,t+\tau} = (L_t + \tau B_t)S_{t-m+\tau_m}$$

where $\tau_m = \lfloor (\tau - 1) \mod m \rfloor + 1$.

Time Series: Initialization of Holt-Winter's Method

- We must have initial estimates L_0, B_0 and the S_t for m previous periods.
- A minimum of $2m$ data points should be available for estimation.

Time Series: EatComo Example



Time Series: Exponential Smoothing

- Simple exponential smoothing: no trend
  - One smoothing equation.
  - Parameter: $0 \leq \alpha \leq 1$.
- Holt's method: linear trend
  - Holt's method extends exponential smoothing to capture trend.
  - Two smoothing equations-one for the level and one for trend.
  - Parameters: $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$.
- Holt and Winters extended Holt's method to capture seasonality
  - Holt and Winters extended Holt's method to capture seasonality.
  - Three smoothing equations-one for the level, one for trend, and one for seasonality.
  - Parameters: $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1 - \alpha$, and m =period of seasonality.

Time Series: Exponential Smoothing

- Stationary Series without Trend
  - Use Moving Average or Exponential Smoothing
- Stationary Series with Trend
  - Use Linear Regression Analysis
  - Use Holt's Method (Double Exponential Smoothing)
- Seasonal Series with Trend
  - Use Holt-Winter's Method (Triple Exponential Smoothing)

- ARIMA models provide another approach to time series forecasting. Exponential smoothing and ARIMA models are the two most widely-used approaches to time series forecasting, and provide complementary approaches to the problem.
- While exponential smoothing models were based on a description of trend and seasonality in the data, ARIMA models aim to describe the autocorrelations in the data.
- ARIMA models, also called Box-Jenkins models, are models that may include autoregressive terms, moving average terms, and differencing operations.

Time Series: AR

- One of the simplest ARIMA type models is a model in which we use a linear model to predict the value at the present time using the value at the previous time. This is called an AR(1) model, standing for autoregressive model of order 1.
- The order of the model indicates how many previous times we use to predict the present time.
- A start in evaluating whether an AR(1) might work is to plot values of the series against lag 1 values of the series.
- Let y_t denote the value of the series at any particular time t, so y_{t-1} denotes the value of the series one time before time t. That is, y_{t-1} is the lag 1 value of y_t.
- In a multiple regression model, we forecast the variable of interest using a linear combination of predictors.
- In an autoregression model, we forecast the variable of interest using a linear combination of past values of the variable.
- The term autoregression indicates that it is a regression of the variable against itself.

Time Series: AR

- For a series y_t, the level of its current observations depends on the level of its lagged observations.
- For example, if we observe a high GDP realization this quarter, we would expect that the GDP in the next few quarters are good as well. This way of thinking can be represented by an AR model.
- The AR(1) (autoregressive of order one) can be written as:

$$y_t = c + \phi_1 y_{t-1} + \epsilon_t$$

  where

$$\epsilon_t \sim iidN(0, \sigma_\epsilon^2)$$

- $AR(p)$ (autoregressive of order p) can be written as:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

Time Series: AR

- The AR(1) can be written as:

$$(1 - \phi_1 B)y_t = c + \epsilon_t$$

- $AR(p)$ can be written as:

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)y_t = c + \epsilon_t$$

We call lag polynomial:

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$$

Example:

$$y_t = 0.8y_{t-1} + 0.09y_{t-2} + \epsilon_t$$

Time Series: MA

- A moving average term in a time series model is a past error (multiplied by a coefficient).
- For example, if we observe a negative shock to the economy, say, a catastrophic earthquake, then we would expect that this negative effect affects the economy not only for the time it takes place, but also for the near future. This way of thinking can be represented by a MA model.
- The MA(1) (moving average of order one) can be written as:

$$y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1}$$

  where

$$\epsilon_t \sim iidN(0, \sigma_\epsilon^2)$$

- The qth order moving average model, denoted by MA(q) is

$$y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

- The MA(1) can be written as:
$$y_t = c + (1 + \theta_1 B)\epsilon_t$$
- $MA(q)$ can be written as:
$$y_t = c + (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)\epsilon_t$$

Time Series: ARMA

$$y_t = \overbrace{c} + \overbrace{\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}} + \overbrace{\theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}} + \epsilon_t.$$

- $AR(p)$ can be written as:

$$(1 - \phi_1 B - \phi_2 B^2 + \cdots - \phi_p B^p) y_t = c + \epsilon_t$$

- $MA(q)$ can be written as:

$$y_t = c + (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)\epsilon_t$$

- lag polynomials

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 + \cdots - \phi_p B^p)$$
$$\theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)$$

- ARMA Process are in general

$$\begin{aligned} AR: &\quad \phi(B) y_t = c + \epsilon_t \\ MA: &\quad y_t = c + \theta(B)\epsilon_t \\ ARMA: &\quad \phi(B) y_t = c + \theta(B)\epsilon_t \end{aligned}$$

Time Series: Estimation

- once the model order has been identified (i.e., the values of p and q), we need to estimate the parameters $c, \phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q$.
- Julia estimates the ARIMA model using maximum likelihood estimation (MLE).
- this technique finds the values of the parameters which maximize the probability of obtaining the data that have been observed.
- MLE is very similar to the least squares estimates that would be obtained by minimizing

$$\sum_{t=1}^{T} e_t^2$$

- different software give slightly different answers as they use different methods of estimation

Time Series: non-Seasonal ARIMA Models

- ARIMA combines differencing with autoregression and a moving average model.
- The full model can be written as

$$\underbrace{(1-\phi_1 B - \cdots - \phi_p B^p)}_{\text{AR}(p)} \quad \underbrace{(1-B)^d}_{d \text{ diffs}} y_t = c + \underbrace{(1+\theta_1 B + \cdots + \theta_q B^q)}_{\text{MA}(q)} e_t$$

- Selecting appropriate values for p, d and q can be done using the auto_arima in Julia.

- A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA models we have seen so far. It is written as follows:

$$\text{ARIMA} \quad \underbrace{(p, d, q)}_{\substack{\uparrow \\ \left(\begin{array}{l}\text{Non-seasonal part}\\ \text{of the model}\end{array}\right)}} \quad \underbrace{(P, D, Q)_m}_{\substack{\uparrow \\ \left(\begin{array}{l}\text{Seasonal part}\\ \text{of the model}\end{array}\right)}}$$

- where $m =$ number of periods per season.
- We use uppercase notation for the seasonal parts of the model, and lowercase notation for the non-seasonal parts of the model.

Time Series: Seasonal ARIMA Models

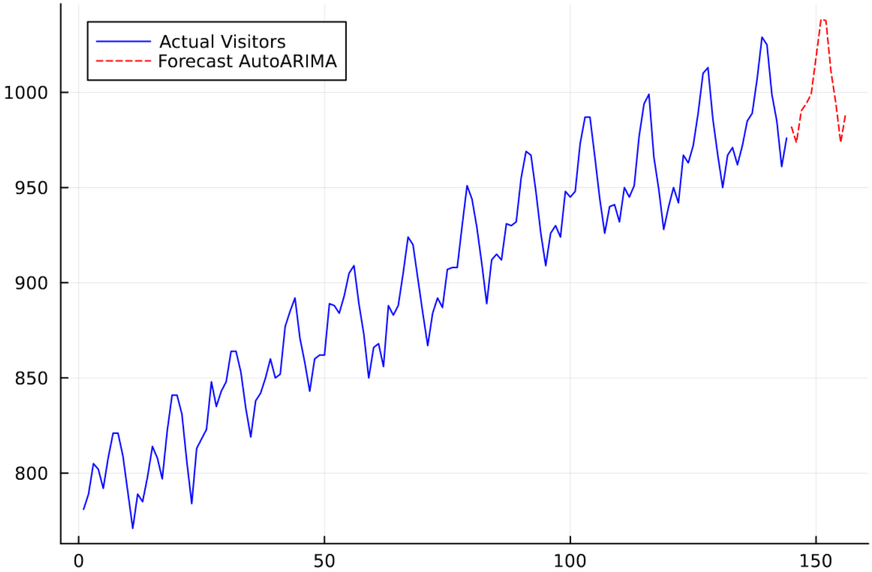
- The seasonal part of the model consists of terms that are very similar to the non-seasonal components of the model, but they involve backshifts of the seasonal period.
- For example, an ARIMA$(1,1,1)(1,1,1)_4$ model (without a constant) is for quarterly data (m=4) and can be written as

$$\underset{\text{Non-seasonal AR(1)}}{(1-\phi_1 B)}\ \underset{\text{Seasonal AR(1)}}{(1-\Phi_1 B^4)}\ \underset{\text{Non-seasonal difference}}{(1-B)}\ \underset{\text{Seasonal difference}}{(1-B^4)}y_t = \underset{\text{Non-seasonal MA(1)}}{(1+\theta_1 B)}\ \underset{\text{Seasonal MA(1)}}{(1+\Theta_1 B^4)}e_t.$$

Time Series: Evaluation

- Use maximum likelihood estimation
- Good models are obtained by minimizing either the AIC, AICc$_c$ or BIC.

Time Series: ARIMA EatComo Example



Time Series: ARIMA EatComo Example

```
                             Results
==============================================================
Model:                        SARIMA(1, 0, 1)x(0, 1, 1, 12) with zero mean
Number of observations:       144
Number of unknown parameters: 4
Log-likelihood:               -410.7826
AIC:                          829.5652
AICc:                         829.8530
BIC:                          841.4445
--------------------------------------------------------------
Parameter       Estimate       Std.Error       z stat       p-value
ar_L1             0.9960               -            -             -
ma_L1            -0.4069               -            -             -
s_ma_L12         -0.5505               -            -             -
sigma2_η         28.1018               -            -             -
```

Time Series: ARIMA EatComo Example

The ARIMA(1, 0, 1)x(0, 1, 1, 12) model is used.

- The model can be written as follows

$$(1 - \phi_1 B)(1 - B^{12})y_t = (1 + \theta_1 B)(1 + \Theta_1 B^{12})\epsilon_t,$$

  where $\phi_1 = 0.9960$, $\theta_1 = -0.4069$, and $\Theta_1 = -0.5505$.

- Then we expand the left-hand side to obtain

$$[1 - \phi_1 B - B^{12} + \phi_1 B^{13}]y_t = (1 + \theta_1 B + \Theta_1 B^{12} + \theta_1 \Theta_1 B^{13})\epsilon_t,$$

- and applying the backshift operator give

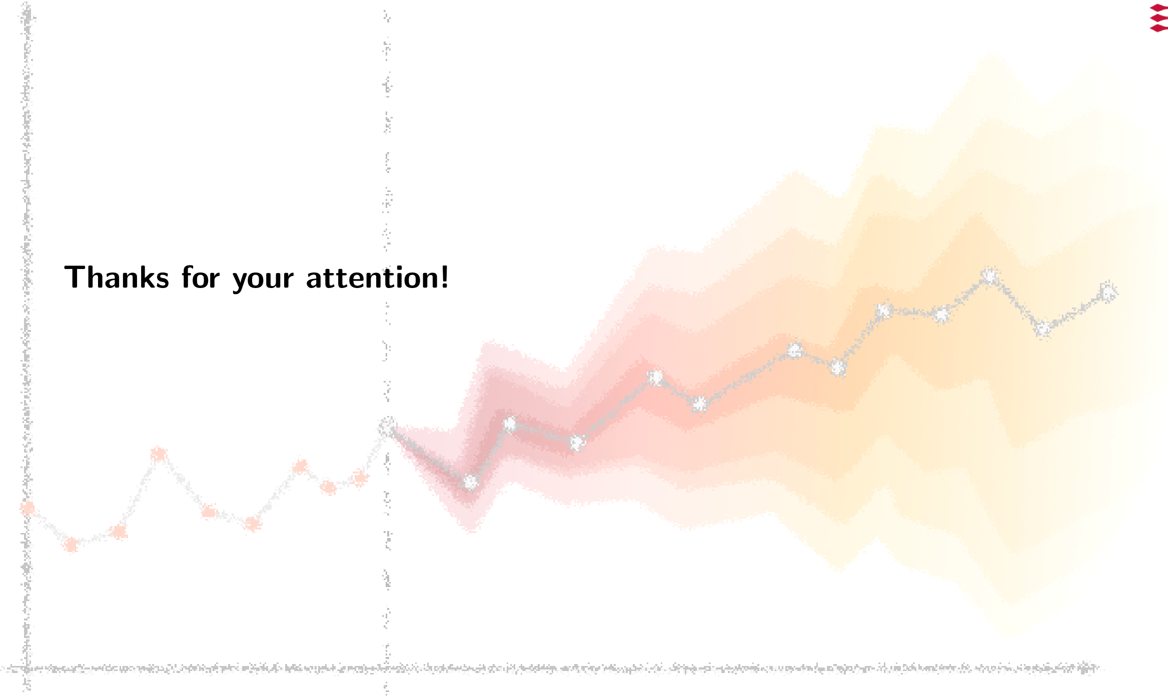
$$y_t - \phi_1 y_{t-1} - y_{t-12} + \phi_1 y_{t-13} = \epsilon_t + \theta_1 \epsilon_{t-1} + \Theta_1 \epsilon_{t-12} + \theta_1 \Theta_1 \epsilon_{t-13}.$$

- Finally, we move all terms other than y_t to the right-hand side:

$$y_t = \phi_1 y_{t-1} + y_{t-12} - \phi_1 y_{t-13} + \epsilon_t + \theta_1 \epsilon_{t-1} + \Theta_1 \epsilon_{t-12} + \theta_1 \Theta_1 \epsilon_{t-13}.$$

- Linear exponential smoothing models all special cases of ARIMA models.
- Non-linear exponential smoothing models have no equivalent ARIMA counterparts.
- Many ARIMA models have no exponential smoothing counterparts.

Wrapping up



Thanks for your attention!