Additional Time Series Materials

Sean Hellingman ©

Regression for Applied Data Science (ADSC2020)

shellingman@tru.ca

Winter 2025

THOMPSON RIVERS UNIVERSITY

Topics

Introduction

- So far, we have only relied on the correlation structures of Y_t.
- We can also include *exogenous* explanatory variables $X_{j,t}$.
- There are machine learning algorithms specifically designed to forecast time-dependent data.

ARIMAX

ARIMAX

- ARIMA with eXplanatory/eXogenous variables.
- Traditional approach:

$$Y_t \sim N(\mu + \text{AR terms} + \text{MA terms} + \phi_1 X_{1,t} + ... + \phi_r X_{r,t}, \sigma^2)$$

- *forecast* R package approach:

$$Y_t = \phi_1 X_{1,t} + ... + \phi_r X_{r,t} + e_t$$
$$e_t \sim N(\mu + \text{AR terms} + \text{MA terms}, \sigma^2)$$

ARIMAX Comments

- The interpretations of $\hat{\phi}_t$ can be tricky.
  - Especially when the series is differenced.
- *If you are only interested in forecasting this isn't as much of an issue.*
- If you are interested in making inferences you can try dynamic regression.

ARIMAX in R

- Can use the `Arima(ts,order = c(p,d,q), seasonal = c(P,D,Q), xreg = matrix)` function from the *forecast package*.
- The `xreg = matrix` argument also works in the `auto.arima()` function.
- Your matrix needs to have the same number of rows as the time series.

Example 1

1. Import the *BagComplaints.csv* dataset into R and run the provided code to format the data.
2. Use the techniques that we have covered to estimate an appropriate SARIMA model.
3. Estimate an appropriate ARIMAX model.

Forecasting ARIMAX

- Once the model is estimated you can use the the `forecast(model,h,xreg=xreg1)` function.
- You need to include a matrix of your explanatory variables.
  - *It must have h rows.*
- This function will forecast your Y_{t+h} series.

Example 2

1. Forecast $h = 6$ for each of the models estimated in Example 1.
2. Which model is better?

Neural Network Autoregression

Neural Networks

- **Artificial neural networks** are forecasting methods that are based on simple mathematical models of the brain.
- They allow for non-linear relationships between explanatory variables and the response variable.
- Lagged values of the time series can be used as inputs to a neural network.

Neural Networks Illustration

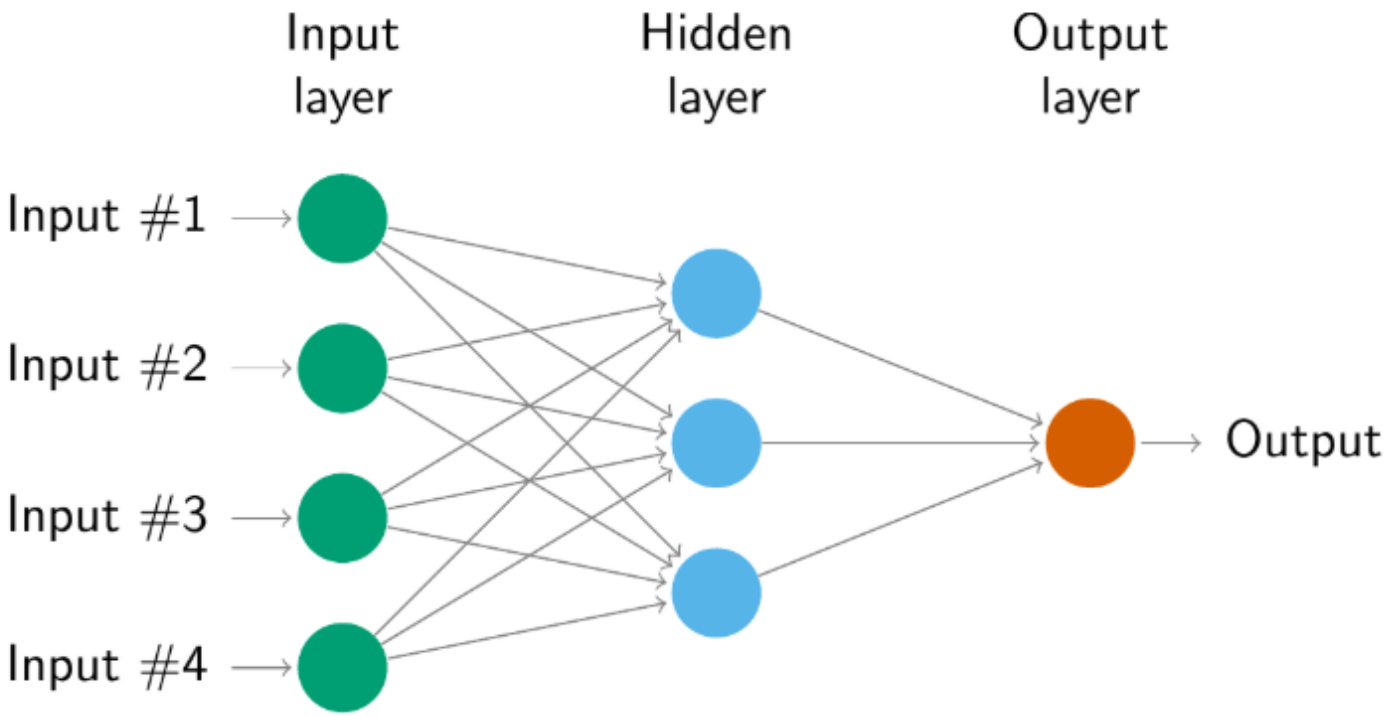

Figure: Source: (3)

Neural Network Autoregression

- Lagged values of the time series can be used as inputs to a neural network (**NNAR model**).
- We are only going to consider one hidden layer in this course.
- NNAR$(p, P, k)_s$ inputs:
  - $y_{t-1}, y_{t-2}, \dots, y_{t-p}$
  - $y_{t-s}, y_{t-2s}, \dots, y_{t-Ps}$
  - k neurons in the hidden layer.

NNAR in R

- In R: `nnetar(ts,p,P,size=`k`,xreg = matrix)`
- If not specified:
  - `p` selected based on best AIC for AR(p).
  - `P` selected based on best AIC for seasonal AR(P).
  - $k = (p + P + 1)/2$ rounded to the nearest integer.
- When it comes to forecasting, the network is applied iteratively.
  - *Forecasts are fed into the model to produce further forecasts.*

Example 3

1. Fit two NNAR models, one with and one without the X matrix.
2. Forecast $h = 6$ for each of the models.
3. Which model is better?

Forecast Combinations

Forecast Combinations

- Combining multiple forecasts leads to increased forecast accuracy.
- Simply averaging the forecasts can improve accuracy.
- Algorithms may also be used to select the weights.
  - *We will only look at averages, but you can check the ForecastComb for additional methods.*

Forecast Combination Weights

- Combining multiple forecasts leads to increased forecast accuracy.

$$\hat{Y}_{t+h} = \frac{w_1}{w_1 + w_2 + ... + w_k} \cdot \hat{Y}^{(1)}_{t+h} + ... + \frac{w_k}{w_1 + w_2 + ... + w_k} \cdot \hat{Y}^{(k)}_{t+h}$$

- *Weights must sum to 1.*

Example 4

1. Generate a combined forecast from the best forecasts in Example 2 and Example 3.

2. Did we improve our forecasts?

Example 4

1. Generate a combined forecast from the two best forecasts in Example 2 and Example 3.

2. Did we improve our forecast?

Example 5

1. Adjust the provided code to visualize all of your mean forecasts.
2. How do they look?

Exercise 1

- You are now fully equipped to forecast most time series.
- Use what you have learned to forecast some data.
- *A well done project in this area will look really good in your portfolio.*

References & Resources

1. Shumway, R. H., & Stoffer, D. S. (2011). *Time Series Analysis and Its Applications: With R Examples*. Springer Texts in Statistics. Link

2. Jonathan, D. C., & Kung-Sik, C. (2008). *Time series analysis with applications in R*.

3. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice.* OTexts. Link

- TSA
- forecast package
- Arima()
- auto.arima()
- Arima()
- nnetar()