

# — ARIMA Modeling

*Matt Brems*

# Learning Objectives

- Describe the purpose of the autoregressive and moving average components.
- Define hyperparameters  $p$ ,  $d$ , and  $q$ .
- Describe AIC.
- Fit an ARIMA model.



# **We have multiple approaches to work with time series data.**

**ARIMA  
Models**

**Exponential  
Smoothing  
Methods**

**Recurrent Neural  
Networks (RNNs)**

Note: This is not an exhaustive list of models, but the most common ones.



# Why ARIMA?

- Among the most common approaches to time series modeling.
- Highly flexible; it can model time series with varying characteristics.
  - It takes information from **both long-term trends and sudden shocks!**
- Can easily be extended into more advanced models.
- Tends to perform well with moderate amounts of data.
  - It can be hard to get lots of time series data!

## Downsides of ARIMA Models

- ARIMA models are best suited for **short-term forecasts**, but very quickly will start predicting the mean.
  - Some extensions to ARIMA models can handle this better.



## What is an **ARIMA** model?

# ARIMA

## What is an ARIMA model?

ARIMA

Autoregressive Integrated Moving Average



**What do you think the word “autoregressive” means?**



# What do you think the word “autoregressive” means?

- Autocorrelation is the correlation of one variable with itself.
- An autobiography is a book written by a person, about that same person.
- An autotransplant is a surgical procedure in which an organ is transplanted from a person to that same person.
- **Autoregressive means we regress a variable on itself.**
  - We'll regress newer values on older values.



## **$AR(p)$ : An autoregressive model of order $p$**

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} \\ &= \beta_0 + \sum_{k=1}^p \beta_k Y_{t-k} \end{aligned}$$

## **AR(p): An autoregressive model of order $p$**

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} \\ &= \beta_0 + \sum_{k=1}^p \beta_k Y_{t-k} \end{aligned}$$

**Purpose:** An autoregressive model explains **long-term trends** in our data.

**Hyperparameter:**  $p$ , the number of previous values of  $Y$  to put into our model.



We'll **GridSearch** to find this value!

## What is an ARIMA model?

ARIMA

Autoregressive Integrated Moving Average

# Moving Average Models

- A moving average model takes **previous error terms as inputs**.
- The goal is to predict future values based on recent forecasting errors.
  - This isn't identical to boosting, but is similar in that fitting is driven by errors.
- *Annoying: this isn't the same thing as moving average smoothing.*



## **MA(q): A moving average model of order q**

$$\begin{aligned} Y_t &= \mu + w_1 \varepsilon_{t-1} + w_2 \varepsilon_{t-2} + \cdots + w_p \varepsilon_{t-q} \\ &= \mu + \sum_{k=1}^q w_k \varepsilon_{t-k} \end{aligned}$$

## **MA(q): A moving average model of order q**

$$\begin{aligned} Y_t &= \mu + w_1 \varepsilon_{t-1} + w_2 \varepsilon_{t-2} + \cdots + w_p \varepsilon_{t-q} \\ &= \mu + \sum_{k=1}^q w_k \varepsilon_{t-k} \end{aligned}$$

**Purpose:** A moving average model explains **sudden shocks** in our data.

**Hyperparameter:**  $q$ , the number of previous errors  $\varepsilon$  to put into our model.



We'll **GridSearch** to find this value!

## How do we GridSearch to find the best values of $p$ and $q$ ?

- Because we're working in statsmodels, we will **manually GridSearch** values of  $p$  and  $q$  to see which gives us the **lowest AIC**.
- AIC, or Akaike Information Criterion, is a common way to evaluate time series models. (AIC is an attribute in **statsmodels**.)
- Remember that a model is a simplification of reality?
  - AIC attempts to measure how much information we lose when we simplify reality with a model.

$$AIC = 2 \times [\text{\# of model parameters}] - 2 \times \log(\text{likelihood})$$



## What is an ARIMA model?

ARIMA

Autoregressive Integrated Moving Average

At this point, it's helpful to see what an **ARIMA** model is.

$$Y_t^{(d)} = \beta_0 + \sum_{k=1}^p \beta_k Y_{t-k}^{(d)} + \sum_{i=1}^q w_i \varepsilon_{t-i} + \varepsilon_t$$

At this point, it's helpful to see what an **ARIMA** model is.

ARIMA

We literally just added together the AR(p) and MA(q) models.

$$Y_t^{(d)} = \beta_0 + \sum_{k=1}^p \beta_k Y_{t-k}^{(d)} + \sum_{i=1}^q w_i \varepsilon_{t-i} + \varepsilon_t$$

Autoregressive Integrated Moving Average

# What is that $Y_t^{(d)}$ ?

- Onto the notebook!

*...but first!*

# ARIMA Cheat Sheet

	AR	I	MA
<b>Stands for:</b>	Autoregressive	Integrated	Moving Average
<b>Summary:</b>	Regress future values on <b>past values</b> .	Differences our Y variable.	Regress future values on <b>past errors</b> .
<b>Looks Like:</b>	$\beta_0 + \sum_{k=1}^p \beta_k Y_{t-k}^{(d)}$	$Y_t^{(d)}$	$\beta_0 + \sum_{i=1}^q w_i \varepsilon_{t-i} + \varepsilon_t$
<b>Purpose:</b>	Long-term trends.	Ensure stationarity.	Sudden shocks.
<b>Hyperparameter:</b>	$p$	$d$	$q$
<b>Find good value of hyperparameter by:</b>	GridSearch	Augmented Dickey-Fuller Test	GridSearch