ARIMA Modeling

Matt Brems

GENERAL ASSEMBLY

Learning Objectives

- Describe the purpose of the autoregressive and moving average components.
- Define hyperparameters *p*, *d*, and *q*.
- Describe AIC.
- Find the right value of *p* and *q* using AIC.
- Find the right value of *d* using the augmented Dickey-Fuller test.
- Complete a manual GridSearch.
- Fit an ARIMA model.

We have multiple approaches to work with time series data.

✓

Linear Models

trend
seasonality

right now

ARIMA Models

not in DSI

Exponential Smoothing Methods

Forecasting: Principles & Practice

wk 11

Recurrent Neural Networks (RNNs)

Note: This is not an exhaustive list of models, but lists 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.

Seasonal ARIMA

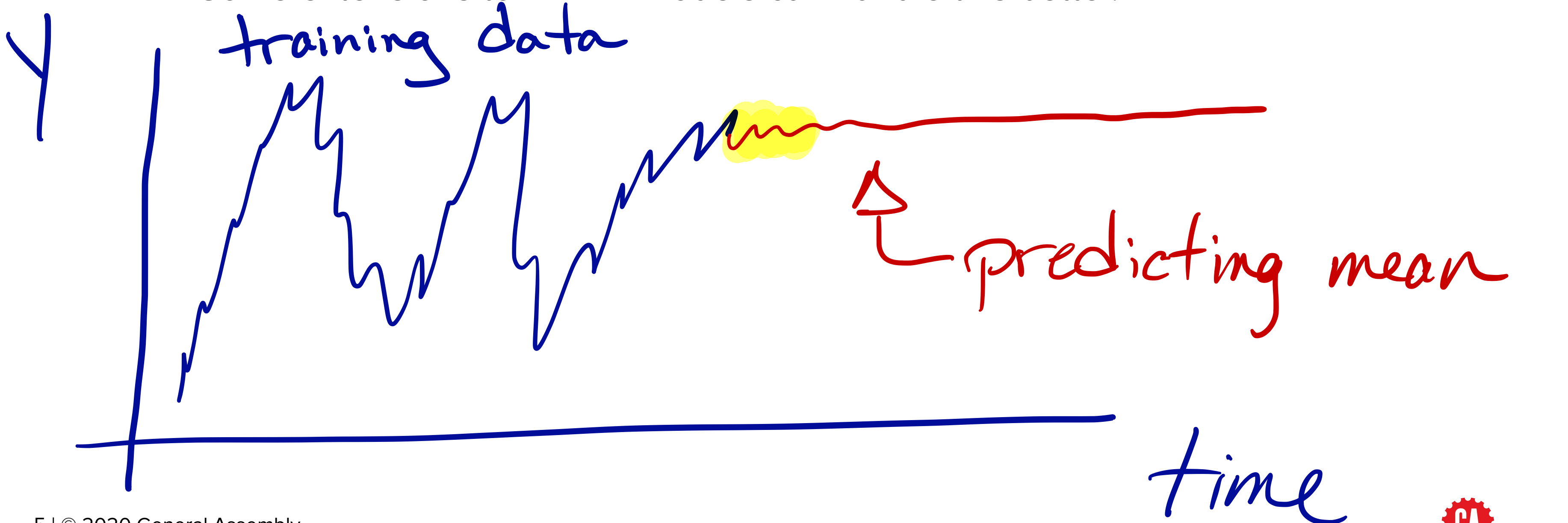
Vector ARIMA → multiple y variables

ARIMA w/ eXogenous predictors → X variables

- 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

1 3 2

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*

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

AR(1): $Y_t = \beta_0 + \beta_1 Y_{t-1}$

AR(2): AR(1) $+ \beta_2 Y_{t-2} \Rightarrow Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2}$

 |

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

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

$$\cancel{Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 Y_{t-12}}$$

not AR(3)

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*

$$Y_t = \mu + w_1\varepsilon_{t-1} + w_2\varepsilon_{t-2} + \cdots + w_{\cancel{3}\,q}\varepsilon_{t-q}$$

$$= \mu + \sum_{k=1}^{q} w_q \varepsilon_{t-q}$$

MA(1): $Y_t = \mu + w_1 \varepsilon_{t-1}$

\varepsilon → ε

\epsilon → ϵ

 |

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

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

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

Hyperparameter: *q*, the number of previous errors ε 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**. less information lost

- 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.

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

```
for p in range(5):
    for q in range(5):
        function(___)
        AIC
```

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