

Arima order and seasonality

Evan Misshula

June 16, 2025

Introduction to ARIMA Models

- ARIMA stands for:
 - AR: Autoregressive
 - I: Integrated (differencing)
 - MA: Moving Average
- ARIMA models are used for univariate time series forecasting.
- General form: $\text{ARIMA}(p, d, q)$

Autoregressive (AR) Model

- An AR(p) model expresses current value y_t as a linear combination of past values:

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

where $\varepsilon_t \sim \text{IID}(0, \sigma^2)$

Moving Average (MA) Model

- An MA(q) model expresses y_t as a linear combination of past error terms:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

where ε_t is white noise

Autocorrelation Function (ACF)

- Measures linear correlation between y_t and y_{t-k}

$$\rho_k = \frac{\text{Cov}(y_t, y_{t-k})}{\sqrt{\text{Var}(y_t) \text{Var}(y_{t-k})}}$$

Partial Autocorrelation Function (PACF)

- Measures correlation between y_t and y_{t-k} , controlling for $y_{t-1}, \dots, y_{t-k+1}$
- Formally, PACF at lag k is the coefficient ϕ_{kk} in the AR(k) model

How to choose p , d , q ?

We use:

- **Differencing**: determines d , the number of times to difference the data to make it stationary
- **ACF plot**: helps choose q , the MA order
- **PACF plot**: helps choose p , the AR order

Rules of Thumb

- If ACF cuts off sharply after lag q : likely $MA(q)$
- If PACF cuts off sharply after lag p : likely $AR(p)$
- If both ACF and PACF decay gradually: try $ARMA(p,q)$

Let's run code

07-arima-order.py

Summary of S-Arima

- ARIMA models handle trend and autocorrelation
- Use differencing + ACF + PACF to identify parameters
- Fit models with 'statsmodels' in Python

References

- Box, Jenkins, Reinsel: **Time Series Analysis**
- Statsmodels documentation: ARIMA and plot_acf/pacf