Time Series Modeling: ADF, ACF, and PACF

Jerome Chou

February 24, 2025

Introduction

- Time series modeling is essential for forecasting and understanding data trends.
- ARIMA (AutoRegressive Integrated Moving Average) is a popular method in time series analysis.
- Key concepts:
    - ADF (Augmented Dickey-Fuller Test) for stationarity
    - ACF (Autocorrelation Function) for lag dependency
    - PACF (Partial Autocorrelation Function) for AR order selection

Augmented Dickey-Fuller (ADF) Test

Definition:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \epsilon_t \tag{1}$$

where:

- y_t is the time series,
- $\Delta y_t = y_t - y_{t-1}$ (first difference),
- t is the time index,
- ϵ_t is white noise.

Hypothesis Testing:

- $H_0 : \gamma = 0$ (unit root exists, non-stationary)
- $H_1 : \gamma < 0$ (stationary process)

ADF Test: Intuition

- Checks if a time series has a unit root (i.e., non-stationary).
- If H_0 is rejected, the series is stationary.
- ADF includes lagged differences to handle autocorrelation.
- Used before applying ARIMA, as ARIMA requires stationarity.

Autocorrelation Function (ACF)

Definition:

$$\rho_k = \frac{\mathsf{Cov}(y_t, y_{t-k})}{\mathsf{Var}(y_t)} \tag{2}$$

Alternative Form:

$$\rho_k = \frac{\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2} \tag{3}$$

where:

- ρ_k is the autocorrelation at lag k.
- $\bar{y}$ is the mean of the time series.

ACF: Intuition

▶ Measures how strongly past values influence the present.
▶ Helps determine the order of MA (Moving Average) models in ARIMA.
▶ If ACF decays gradually, the series may have an AR component.
▶ If ACF cuts off after a few lags, the series likely follows an MA process.

Partial Autocorrelation Function (PACF)

Definition (Recursive Formula):

$$\phi_{kk} = \rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j} \qquad (4)$$

Alternative Representation:

$$\phi_{kk} = \frac{\text{Cov}(y_t, y_{t-k}|y_{t-1}, \ldots, y_{t-k+1})}{\text{Var}(y_t)} \qquad (5)$$

Interpretation:

▶ PACF measures the correlation between y_t and y_{t-k}, removing the effect of intermediate lags.

▶ Used to identify the AR (AutoRegressive) order in ARIMA.

PACF: Intuition

- ▶ PACF measures direct influence at each lag, ignoring indirect influences.
- ▶ If PACF cuts off after lag p, an AR(p) model is appropriate.
- ▶ If PACF shows slow decay, a higher-order AR model may be needed.

Summary

- ADF test is used to check stationarity before applying ARIMA.
- ACF helps identify MA components.
- PACF helps identify AR components.
- Together, these tools guide ARIMA model selection.