Mathematical Formulations of ARIMA

Jerome Chou

January 31, 2025

What is ARIMA?

- ARIMA (Autoregressive Integrated Moving Average) is a statistical model for time series forecasting.
- It consists of three components:
  - **Autoregressive (AR)**: Uses past values to predict future values.
  - **Integrated (I)**: Applies differencing to make the time series stationary.
  - **Moving Average (MA)**: Models the residual errors of past observations.

Mathematical Definition of ARIMA

- An ARIMA(p, d, q) model is defined as:

$$\Phi_p(B)(1 - B)^d y_t = \Theta_q(B)\epsilon_t$$

where:

- y_t is the observed time series.
- B is the backshift operator ($By_t = y_{t-1}$).
- d is the order of differencing.
- ϵ_t is white noise.

Autoregressive (AR) Process

- An AR(p) process is given by:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$$

where:

- ϕ_i are the autoregressive coefficients.
- c is a constant term.
- ϵ_t is white noise.

Moving Average (MA) Process

- An MA(q) process is given by:

$$y_t = \mu + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

where:

- θ_j are the moving average coefficients.
- μ is the mean of the series.
- ϵ_t is white noise.

Differencing for Stationarity

- To remove trends, differencing is applied:

$$y'_t = y_t - y_{t-1} = (1 - B)y_t$$

- For higher-order differencing (d times):

$$y_t^{(d)} = (1 - B)^d y_t$$

- Differencing transforms a non-stationary series into a stationary one.

Estimation of ARIMA Model

- The ARIMA(p, d, q) equation combines AR and MA components:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t$$

- Parameters (ϕ, θ, c) are estimated using Maximum Likelihood Estimation (MLE) or Least Squares.

Loss Function for ARIMA

- Model parameters are estimated by minimizing the error.
- **Mean Squared Error (MSE):**

$$J(\theta) = \frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{y}_t)^2$$

- **Log-Likelihood Function for MLE:**

$$L(\theta) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^{N} (y_t - \hat{y}_t)^2$$

Model Selection Criteria

- **Akaike Information Criterion (AIC):**

$$AIC = 2k - 2\log L$$

  where k is the number of parameters.
- **Bayesian Information Criterion (BIC):**

$$BIC = k\log N - 2\log L$$

  where N is the number of observations.
- A lower AIC or BIC value indicates a better model.

Conclusion

- ARIMA models are widely used for time series forecasting.
- The model consists of AR, MA, and differencing components.
- Parameters are estimated using MLE or Least Squares.
- Model selection is based on AIC/BIC to balance complexity and fit.