MONASH University

MONASH BUSINESS SCHOOL

ETC3550/ETC5550 Applied forecasting

Ch9. ARIMA models

OTexts.org/fpp3/

ARIMA models

AR: autoregressive (lagged observations as inputs)
I: integrated (differencing to make series stationary)
MA: moving average (lagged errors as inputs)

An ARIMA model is rarely interpretable in terms of visible data structures like trend and seasonality. But it can capture a huge range of time series patterns.

Stationarity

Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \dots, y_{t+s})$ does not depend on t.

Stationarity

Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \dots, y_{t+s})$ does not depend on t.

Transformations help to **stabilize the variance**.

For ARIMA modelling, we also need to **stabilize the mean**.

Differencing

- Differencing helps to **stabilize the mean**.
- First differencing: *change* between consecutive observations: $y'_t = y_t - y_{t-1}$.
- Seasonal differencing: *change* between years:

Automatic differencing

Using unit root tests for first differencing

1. Augmented Dickey Fuller test: null hypothesis is that the data are non-stationary and non-seasonal.
2. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test: null hypothesis is that the data are stationary and non-seasonal.

Seasonal strength

STL decomposition: $y_t = T_t + S_t + R_t$

Seasonal strength $F_s = \max\left(0, 1 - \frac{\text{Var}(R_t)}{\text{Var}(S_t + R_t)}\right)$

If $F_s > 0.64$, do one seasonal difference.

Random walk model

If differenced series is white noise with zero mean:

$$y_t - y_{t-1} = \varepsilon_t \qquad \text{or} \qquad y_t = y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- Model behind the **naïve method**.
- Forecast are equal to the last observation (future movements up or down are equally likely).

Random walk with drift model

If differenced series is white noise with non-zero mean:

$$y_t - y_{t-1} = c + \varepsilon_t \qquad \text{or} \qquad y_t = c + y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- c is the **average change** between consecutive observations.
- Model behind the **drift method**.