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: $y'_t = y_t - y_{t-m}$.

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:

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

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

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

Backshift operator notation

- B shifts the data back one period. $By_t = y_{t-1}$
- B^2 shifts the data back two periods: $B(By_t) = B^2 y_t = y_{t-2}$
- A difference can be written as $(1 - B)y_t$
- A dth-order difference can be written as $(1 - B)^d y_t$
- A seasonal difference followed by a first difference can be written as $(1 - B)(1 - B^m)y_t$

AR(1) model

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$$

- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

Autoregressive models

A multiple regression with **lagged values** of y_t as predictors.

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \\ &= c + (\phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p) y_t + \varepsilon_t \end{aligned}$$

Autoregressive models

A multiple regression with **lagged values** of y_t as predictors.

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \\ &= c + (\phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p) y_t + \varepsilon_t \end{aligned}$$

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) y_t &= c + \varepsilon_t \\ \phi(B) y_t &= c + \varepsilon_t \end{aligned}$$

- ε_t is white noise.
- $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$

Stationarity conditions

We normally restrict autoregressive models to stationary data, and then some constraints on the values of the parameters are required.

General condition for stationarity

Complex roots of $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$ lie outside the unit circle on the complex plane.

Stationarity conditions

We normally restrict autoregressive models to stationary data, and then some constraints on the values of the parameters are required.

General condition for stationarity

Complex roots of $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$ lie outside the unit circle on the complex plane.

- For $p = 1$: $-1 < \phi_1 < 1$.
- For $p = 2$:

  $-1 < \phi_2 < 1 \qquad \phi_2 + \phi_1 < 1 \qquad \phi_2 - \phi_1 < 1.$
- More complicated conditions hold for $p \geq 3$.
- fable takes care of this.

Moving Average (MA) models

A multiple regression with **past *errors*** as predictors.

$$\begin{aligned} y_t &= c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \\ &= c + (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)\varepsilon_t \\ &= c + \theta(B)\varepsilon_t \end{aligned}$$

Moving Average (MA) models

A multiple regression with **past *errors*** as predictors.

$$
\begin{aligned}
y_t &= c + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q} \\
&= c + (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)\varepsilon_t \\
&= c + \theta(B)\varepsilon_t
\end{aligned}
$$

- ε_t is white noise.
- $\theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)$

Invertibility

General condition for invertibility

Complex roots of $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q$ lie outside the unit circle on the complex plane.

Invertibility

General condition for invertibility

Complex roots of $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q$ lie outside the unit circle on the complex plane.

- For $q = 1$: $-1 < \theta_1 < 1$.
- For $q = 2$:

  $-1 < \theta_2 < 1 \qquad \theta_2 + \theta_1 > -1 \qquad \theta_1 - \theta_2 < 1.$
- More complicated conditions hold for $q \geq 3$.
- fable takes care of this.

ARIMA models

ARIMA(p, d, q) model: $\phi(B)(1 - B)^d y_t = c + \theta(B)\varepsilon_t$

AR: p = order of the autoregressive part

I: d = degree of first differencing involved

MA: q = order of the moving average part.

ARIMA models

ARIMA(p, d, q) model: $\phi(B)(1 - B)^d y_t = c + \theta(B)\varepsilon_t$

AR: p = order of the autoregressive part
I: d = degree of first differencing involved
MA: q = order of the moving average part.

- Conditions on AR coefficients ensure stationarity.
- Conditions on MA coefficients ensure invertibility.
- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)