ETC3550/ETC5550
Applied forecasting

Week 9: ARIMA models

Random walk model

If differenced series is white noise with zero mean:

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

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 model

$$y_t = y_{t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t \sim NID(0, \sigma^2).$$

Random walk model

$$y_t = y_{t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t \sim NID(0, \sigma^2).$$

$$\begin{aligned} y_{T+h} &= y_{T+h-1} + \varepsilon_{T+h} \\ &= y_{T+h-2} + \varepsilon_{T+h-1} + \varepsilon_{T+h} \\ &= \ldots \\ &= y_T + \varepsilon_{T+1} + \cdots + \varepsilon_{T+h} \end{aligned}$$

Random walk model

$$y_t = y_{t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t \sim NID(0, \sigma^2).$$

$$\begin{aligned}
y_{T+h} &= y_{T+h-1} + \varepsilon_{T+h} \\
&= y_{T+h-2} + \varepsilon_{T+h-1} + \varepsilon_{T+h} \\
&= \dots \\
&= y_T + \varepsilon_{T+1} + \cdots + \varepsilon_{T+h}
\end{aligned}$$

$$\begin{aligned}
\text{So} \qquad & \mathrm{E}(y_{T+h}|y_1, \dots, y_T) = y_T \\
\text{and} \qquad & \mathrm{Var}(y_{T+h}|y_1, \dots, y_T) = h\sigma^2
\end{aligned}$$

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: $B(By_t) = B^2y_t = y_{t-2}$
- 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

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

- When $\phi_1 = 0$, y_t is **equivalent to WN** (with mean c)
- 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 parameters are needed.

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 parameters are needed.

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 $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$ 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)

R model

Intercept form

$$(1 - \phi_1 B - \cdots - \phi_p B^p) y'_t = c + (1 + \theta_1 B + \cdots + \theta_q B^q)\varepsilon_t$$

Mean form

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(y'_t - \mu) = (1 + \theta_1 B + \cdots + \theta_q B^q)\varepsilon_t$$

- $y'_t = (1 - B)^d y_t$
- μ is the mean of y'_t.
- $c = \mu(1 - \phi_1 - \cdots - \phi_p)$.
- fable uses intercept form

Understanding ARIMA models

- If $c = 0$ and $d = 0$, the long-term forecasts will go to zero.
- If $c = 0$ and $d = 1$, the long-term forecasts will go to a non-zero constant.
- If $c = 0$ and $d = 2$, the long-term forecasts will follow a straight line.
- If $c \neq 0$ and $d = 0$, the long-term forecasts will go to the mean of the data.
- If $c \neq 0$ and $d = 1$, the long-term forecasts will follow a straight line.
- If $c \neq 0$ and $d = 2$, the long-term forecasts will follow a quadratic trend.

Understanding ARIMA models

Forecast variance and d

- The higher the value of d, the more rapidly the prediction intervals increase in size.
- For $d = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.

Cyclic behaviour

- For cyclic forecasts, $p \geq 2$ and some restrictions on coefficients are required.
- If $p = 2$, we need $\phi_1^2 + 4\phi_2 < 0$. Then average cycle of length

$$(2\pi)/\left[\text{arc cos}(-\phi_1(1-\phi_2)/(4\phi_2))\right].$$

Exercise

- Find an ARIMA model for the `pelt::Lynx` data