Problem Set 9

Due Wednesday March 26, 4pm

Data Exercises

(1) In the file "realgdpgrowth.xlsx" the series *pdi* contains quarterly U.S. seasonally adjusted aggregate investment growth rates, from the BEA. Estimate an AR(4) model for this series. Which lags appear to be most important? Generate a point and 90% interval one-step-ahead forecast for 2024q4.

(2) In the same file the series *exports* is aggregate U.S. exports (growth rates). Estimate an AR(4) model for this series. Which lags appear to be important? Do you notice anything interesting about the coefficients on the relevant lags? Generate a point and 90% interval one-step-ahead forecast for 2024q4.

(3) In the same file the series *pdi_residential* is aggregate residential investment growth rates. Using an AR(4) model, generate point and interval forecasts for 2024q4 through 2025q3. Create a plot of the point and interval forecasts.

(4) The FRED labels for monthly retail sales of sporting goods stores is MRTSSM45111USN. Use it to get the data for 1992m1-2025m1.

  (a) Graph the time series.
  (b) What model should be used for the trend? Seasonal? Cycle?
  (c) Estimate the model for forecast horizons 1 through 12.
  (d) Generate point and 90% interval forecast for each horizon, and plot your forecasts.

Theoretical Questions

(5) Take a stationary AR(p) process

$$y_t = \alpha + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_t$$

Calculate the mean $E(y_t) = \mu$.

(6) Consider AR(3) process

$$y_t = \frac{11}{6}y_{t-1} - y_{t-2} + \frac{1}{6}y_{t-3} + \varepsilon_t.$$

Is y_t stationary? Is $z_t := y_t - y_{t-1}$ stationary?

(7) Consider a components model

$$y_t = T_t + C_t,$$

$$T_t = 1 + 4t,$$

$$C_t = \beta_1 C_{t-1} + \beta_2 C_{t-2} + \varepsilon_t. \qquad AR(2).$$

(a) Rewrite the components model as a regression on the trend and lags of y_t.

(b) Compare trends in your regression when $\beta_1 + \beta_2 < 1$ and when $\beta_1 + \beta_2 = 1$.

(a). $y_t - \beta_1 y_{t-1} - \beta_2 y_{t-2} = T_t + C_t - \beta_1(T_{t-1} + C_{t-1}) - \beta_2(T_{t-2} + C_{t-2})$

$= ((1+4t) + C_t) - \beta_1(4t-3 + C_{t-1}) - \beta_2(4t-7 + C_{t-2})$

$= (1+4t) - \beta_1(4t-3) - \beta_2(4t-7) + \underbrace{C_t - \beta_1 \cdot C_{t-1} - \beta_2 \cdot C_{t-2}}_{\varepsilon_t}$

$\Rightarrow y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + (1+4t) - \beta_1(4t-3) - \beta_2(4t-7) + \varepsilon_t$

$\Rightarrow y_t = (1 + 3\beta_1 + 7\beta_2) + (4t - 4\beta_1 t - 4\beta_2 t) + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t$

$\Rightarrow y_t = \underbrace{(1 + 3\beta_1 + 7\beta_2)}_{\text{Constants}} + \underbrace{4(1-\beta_1-\beta_2)t}_{\text{trend}} + \underbrace{\beta_1 y_{t-1} + \beta_2 y_{t-2}}_{\text{lags of } y_t} + \varepsilon_t$

Constants and trend

(b). When $\beta_1 + \beta_2 = 1$, $(1-\beta_1-\beta_2) = 0$, So the trend effect is constant over time

when $\beta_1 + \beta_2 < 1$, $(1-\beta_1-\beta_2) > 0$. It means that as $t \uparrow$, the term $4(1-\beta_1-\beta_2)$ will increase, That means the trend effect is positive and linear increase as time increase,

(5) Take a stationary AR(p) process

$$y_t = \alpha + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_t$$

Calculate the mean $E(y_t) = \mu$.

5. $E[y_t] = E[\alpha + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_t] = \alpha + \beta_1 \cdot E[y_{t-1}] + \cdots + \beta_p \cdot E[y_{t-p}] + \underbrace{E[\varepsilon_t]}_{0}$

Since we assume it's stationary, we have $E(y_t) = \mu \quad \forall t$.

Thus, $E[y_t] = \mu = \alpha + \beta_1 \mu + \cdots + \beta_p \mu$

$\Rightarrow (1 - \beta_1 - \beta_2 - \cdots - \beta_p) \cdot \mu = \alpha$

$$\mu = \frac{\alpha}{1 - \beta_1 - \beta_2 - \cdots - \beta_p}$$

(6) Consider AR(3) process

$$y_t = \frac{11}{6} y_{t-1} - y_{t-2} + \frac{1}{6} y_{t-3} + \varepsilon_t.$$

① Is y_t stationary? ② Is $z_t := y_t - y_{t-1}$ stationary?

6. For AR(3), $y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} \iff \underbrace{(1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3)}_{A} y_t = \varepsilon_t$.

We know that the process is stationary if the roots of polynomial A are larger than one in absolute value. A neccessary condition for the correct roots is $\beta_1 + \cdots + \beta_p < 1$. Thus, if $\beta_1 + \cdots + \beta_p \geq 1$ by contrapositive, we know the process can not be stationary.

①. For AR(3), we have $\beta_1 = \frac{11}{6}$, $\beta_2 = -1$, $\beta_3 = \frac{1}{6}$, $(p=3)$

$\beta_1 + \beta_2 + \beta_3 = 1 \not< 1$ Since $\beta_1 + \beta_2 + \beta_3 \geq 1$.

y_t is not stationary.

②

$$Z_{t-1} = y_{t-1} - y_{t-2}$$
$$Z_{t-2} = y_{t-2} - y_{t-3}$$

$$Z_t = y_t - y_{t-1} = \left(\frac{11}{6} y_{t-1} - y_{t-2} + \frac{1}{6} y_{t-3} + \varepsilon_t\right) - y_{t-1}$$

$$= \frac{5}{6} y_{t-1} - y_{t-2} + \frac{1}{6} y_{t-3} + \varepsilon_t$$

$$= \frac{5}{6} y_{t-1} - \frac{5}{6} y_{t-2} - \frac{1}{6} y_{t-2} + \frac{1}{6} y_{t-3} + \varepsilon_t$$

$$= \frac{5}{6} (y_{t-1} - y_{t-2}) - \left(\frac{1}{6} y_{t-2} - \frac{1}{6} y_{t-3}\right) + \varepsilon_t$$

$$= \frac{5}{6} Z_{t-1} - \frac{1}{6} Z_{t-2} + \varepsilon_t$$

(we use ε_t to denote all error terms in different model. It doesn't impact the invertibility)

$$\Rightarrow \quad Z_t - \frac{5}{6} Z_{t-1} + \frac{1}{6} Z_{t-2} = \varepsilon_t$$

$1 - \frac{5}{6} + \frac{1}{6} = \frac{2}{6} < 1$, it may be invertible

$$Z_t \cdot \left(1 - \frac{5}{6} L + \frac{1}{6} L^2\right) = \varepsilon_t.$$

we then solve $1 - \frac{5}{6} L + \frac{1}{6} L^2 = 0$.

$$6 - 5L + L^2 = 0$$

$\Rightarrow$ $root_1 = 2$ $\quad root_2 = 3$. $\quad$ both $|root_i| > 1$.

Thus. Z_t is stationary.