時間序列分析

Ch2 時間序列導論

2.1 時間序列資料

- 資料型態
  - 時間序列 (time-series): $\{y_t : t \in T\}$
  - 橫斷面 (cross-section): $\{y_i : i \in N\}$
  - 追蹤型 (panel): $\{y_{i,t} : t \in T, i \in N\}$
  - T 代表指標集合
  - $y_t \in S$ 其中S代表狀態空間
- 觀察GDP與匯率資料, 我們可以發現以下規律
  - GDP有固定趨勢; 匯率並沒有固定趨勢
  - 資料存在某種序列相關
- 序列相關的衡量方式 $corr(y_t, y_{t-k})$

2.2 時間序列資料的特性

- 落後期: y_{t-k}
- 一階差分

$$\Delta y_t = y_t - y_{t-1} \tag{1}$$

- 重要近似式

$$\Delta \log(y_t) = \frac{\Delta y_t}{y_{t-1}} \tag{2}$$

證明如下

- 落後運算元 (L)
    - $Ly_t = y_{t-1}$
    - $L^k y_t = y_{t-k}$
    - $Lc = c$
    - $(L^k + L^j)y_t = L^k y_t + L^j y_t$
    - $L^k L^j y_t = L^k y_{t-j} = y_{t-k-j}$
    - $L^0 y_t = y_t$
    - $L^{-k} y_t = y_{t+k}$
    - $\forall|\phi| < 1$

$$(1 + \phi L + \phi^2 L^2 + \cdots)y_t = (\frac{1}{1-\phi L})y_t \tag{5}$$

- 有限期落後運算元多項式 (polynomial in the log operator)

$$\phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p = \sum_{j=0}^{p} \phi_j L^p \tag{6}$$

- 無窮期落後運算元多項式
    - $\phi_\infty(L)$定義如下:

$$\phi_\infty(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots = \sum_{j=0}^{\infty} \phi_j L^j \tag{7}$$

- 百分點(pct)和基點 (bp;basic point)
    - ex: $0.01 = 1$(percent)=100 bp
- 指數衡量 (index)
    - 兩種大小不一的指數如何比較?

- 動差

$$
\begin{aligned}
\mu_t &= E(y_t) \\
\sigma^2 &= Var(y_t)
\end{aligned}
\tag{8}
$$

- k階自我相關係數

$$
\rho(t,k) = (\frac{cov(y_t, y_{t-k})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-k})}})y_t \tag{9}
$$

  - Remark: 我們通常關注1階自我相關 $\rho(t,1)$

2.3 定態 stationary

- 弱定態 Weak Stationary
  - 定義: $\{y_t : t \in (-\infty, \infty)\}$滿足下列條件:
    - $E(y_t) = E(y_{t-k}) = \mu$
    - $Var(y_t) < \infty$
    - $Cov(y_t, y_{t-k}) = E(y_t - \mu)(y_{t-k} - \mu) = \gamma(k)$

    $\Rightarrow$ 我們稱$\{y_t\}$為**弱定態**或簡稱**定態**
  - Remark:
    - 隨時間改變，其結構是穩定的
  - 具有穩定結構才是可預測的
    - 可用歷史預測未來
  - 給定 $\{y_t\}$ 定態，可得以下性質:
    - $\gamma(0) = Var(y_t) = Var(y_{t-k})$
    - $\rho(j) = \frac{cov(y_t, y_{t-j})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-j})}} = \frac{\gamma(j)}{\gamma(0)}$
    - $\gamma(j) = \gamma(-j), \quad \rho(j) = \rho(-j)$
- 定義: 嚴格定態 Strong Stationary

$$
\forall\ k\ and\ (t_1, t_2, \cdots, t_n), \quad (y_1, y_2, \cdots, y_{in}) \overset{\mathrm{d}}{=} (y_{t_1+k}, y_{t_2+k}, \cdots, y_{t_n+k}) \tag{10}
$$

$\rightarrow$ 我們稱$\{y_t\}$為**強定態**，但實際上難以驗證。

  - Prop: 若$\{y_t\}$為強定態且$E(y_t^2) < \infty$，則$\{y_t\}$為弱定態。

- 白噪音 (White Noise)
  - 定義: 若$\{\varepsilon\}$有下列性質:
    - $E(\varepsilon_t) = 0 \quad \forall t$
    - $E(\varepsilon_t^2) = \sigma^2 \quad \forall t$
    - $E(\varepsilon_t \varepsilon_{t-k}) = 0 \quad \forall t, k$

  $\Rightarrow$ 則稱$\{\varepsilon_t\}$為白噪音, $\quad \varepsilon_t \sim WN(0, \sigma^2)$

2.4 樣本動差

- 利用實際資料估計動差
  - 定義:
    1. $\gamma\hat{(k)} = \frac{1}{T}\sum(y_t - \bar{y})(y_{t-k} - \bar{y})$
    2. $\rho\hat{(k)} = \frac{\gamma\hat{(k)}}{\gamma\hat{(0)}} = \frac{\frac{1}{T}\sum(y_t-\bar{y})(y_{t-k}-\bar{y})}{\frac{1}{T}\sum(y_t-\bar{y})(y_t-\bar{y})}$

    $*$當$\hat{\rho}$ (或$\rho\hat{(1)}$) 越高，持續性越大。
- Ljung-Box Q-stat

$$\begin{gathered} H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0 \\ where \quad Q(k) = T(T+2)\sum_{j=1}^{k}\frac{(\rho\hat{(j)})^2}{T-j} \sim \chi^2(k) \end{gathered} \tag{11}$$

-

2.5 固定趨勢

- 我們可以用以下式子來表達固定趨勢

$$y_t = \beta_0 + \beta_1 Time_t + \varepsilon_t \tag{12}$$

- 假設 $\quad Time_t = t$

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{13}$$

  其中 $\quad \beta_1$=時間趨勢(固定趨勢)

  若考慮非線性

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t \tag{14}$$

2.6 季節性

- 1990 Q1⟶1991 Q1的大幅成長

  =季節性趨勢+成長

$$D1 = \begin{cases} 1 & ,if \quad Q_1 \\ 0 & , \quad o,w \end{cases} \qquad D2 = \begin{cases} 1 & ,if \quad Q_2 \\ 0 & , \quad o,w \end{cases} \qquad D3 = \begin{cases} 1 & ,if \quad Q_3 \\ 0 & , \quad o,w \end{cases} \tag{15}$$

2.7 收集真實資料

2.8 Python 使用簡介