Applied Econometrics for Macro and Finance

Time Series Regression

Shiu-Sheng Chen

Department of Economics
National Taiwan University

Spring 2023

Part I

Time Series Regression Model

Time Series Regression Model

- Consider the time series regression model

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + \varepsilon_t = x_t{}'\beta + \varepsilon_t$$

where $x_t = (1 \; X_{1t} \cdots X_{kt})'$ is a $(k+1) \times 1$ vector of explanatory variables, $\beta = (\beta_0 \; \beta_1 \cdots \beta_k)'$ is a $(k+1) \times 1$ vector of coefficients, and ε_t is a random error term.

- The standard assumptions of the time series regression model are
  (a) $\{Y_t, x_t\}$ is jointly stationary and ergodic.
  (b) $E(\varepsilon_t | x_t) = 0$.
  (c) $Q = E(x_t x_t{}')$ is of full rank $k+1$ (nonsingular)
  (d) Let $g_t = x_t \varepsilon_t$, and $\{g_t\}$ is an MDS, $S = E(g_t g_t{}') < \infty$
  (e) $Var(\varepsilon_t | x_t) = E(\varepsilon_t^2 | x_t) = \sigma^2$ (conditional homoskedasticity)

Aside: MDS

- Martingale

$$E(Y_t | Y_{t-1}, Y_{t-2}, \ldots, Y_1) = Y_{t-1}$$

- Martingale Difference Sequence (MDS)

$$E(Y_t | Y_{t-1}, Y_{t-2}, \ldots, Y_1) = 0$$

- MDS-CLT:

  Let $\{g_t\}$ be a vector martingale difference sequence that is ergodic stationary,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} g_t \xrightarrow{d} N(0, S)$$

  where $S = E(g_t g_t{}') < \infty$.

Least Squares Estimation

- Ordinary least squares (OLS) estimation is based on minimizing the sum of squared residuals

$$SSR(\beta) \sum_{t=1}^{T} (Y_t - x_t{}'\beta)^2 = \sum_{t=1}^{T} \varepsilon_t^2$$

- The LS estimator is

$$\hat{\beta} = \left(\sum_{t=1}^{T} x_t x_t{}' \right)^{-1} \left(\sum_{t=1}^{T} x_t Y_t \right)$$

In matrix form

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

and

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{\varepsilon}_t^2$$

Least Squares Estimation

	常數項	新台幣匯率 (X_{1t})	日圓匯率 (X_{2t})	英鎊匯率 (X_{3t})
1990M01	1	26.08	145.09	1.65
1990M02	1	26.12	145.53	1.70
1990M03	1	26.35	153.08	1.63
1990M04	1	26.36	158.47	1.64
1990M05	1	26.97	153.52	1.68
1990M06	1	27.39	153.77	1.71
1990M07	1	27.16	149.28	1.81

Least Squares Estimation

$$x_t = (1, X_{1t}, \ldots, X_{kt})' = \begin{bmatrix} 1 \\ X_{1t} \\ \vdots \\ X_{kt} \end{bmatrix} = \begin{bmatrix} 1 \\ 26.08 \\ 145.09 \\ 1.65 \end{bmatrix}, \quad t = \text{1990M01}$$

$$x_t' = \begin{bmatrix} 1 & 26.08 & 145.09 & 1.65 \end{bmatrix}, \quad t = \text{1990M01}$$

$$\boldsymbol{X} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_T' \end{bmatrix} = \begin{bmatrix} 1 & 26.08 & 145.09 & 1.65 \\ 1 & 26.12 & 145.53 & 1.70 \\ 1 & 26.35 & 153.08 & 1.63 \\ 1 & 26.36 & 158.47 & 1.64 \\ 1 & 26.97 & 153.52 & 1.68 \\ 1 & 27.39 & 153.77 & 1.71 \\ 1 & 27.16 & 149.28 & 1.81 \end{bmatrix}$$

Properties of LS Estimator

- Consistency

$$
\begin{aligned}
\hat{\beta} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\beta + \epsilon) \\
&= \beta + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\epsilon \\
&= \beta + \left(\sum_t x_t x_t'\right)^{-1}\left(\sum_t x_t \varepsilon_t\right) \\
&= \beta + \left(\frac{1}{T}\sum_t x_t x_t'\right)^{-1}\left(\frac{1}{T}\sum_t x_t \varepsilon_t\right)
\end{aligned}
$$

$$
\hat{\beta} - \beta = \underbrace{\left(\frac{1}{T}\sum x_t x_t'\right)^{-1}}_{\xrightarrow{p} E[x_t x_t']^{-1}} \underbrace{\left(\frac{1}{T}\sum x_t \varepsilon_t\right)}_{\xrightarrow{p} E(x_t\varepsilon_t)=0} \xrightarrow{p} 0
$$

Properties of LS Estimator

- Asymptotic Distribution

$$\sqrt{T}(\hat{\beta} - \beta) = \sqrt{T}\left(\frac{1}{T}\sum x_t x_t'\right)^{-1}\left(\frac{1}{T}\sum x_t \varepsilon_t\right)$$

$$= \underbrace{\left(\frac{1}{T}\sum x_t x_t'\right)^{-1}}_{\xrightarrow{p} E[x_t x_t']^{-1} = Q^{-1}} \underbrace{\left(\frac{1}{\sqrt{T}}\sum x_t \varepsilon_t\right)}_{\xrightarrow{d} N(0, S)} \xrightarrow{d} N(0, V)$$

where $V = Q^{-1}SQ^{-1}$

$$S = E(g_t g_t') = E(x_t \varepsilon_t \varepsilon_t' x_t') = E(x_t x_t' \varepsilon_t^2)$$

- Note that under the assumption of homskedasticity,

$$S = E(x_t x_t' \varepsilon_t^2) = E\left[E(x_t x_t' \varepsilon_t^2 | x_t)\right] = E[x_t x_t' E(\varepsilon_t^2 | x_t)]$$

$$= E(x_t x_t' \sigma^2) = Q\sigma^2$$

Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation

- In empirical applications using macroeconomic and financial time series, it is often the case that the error terms ε_t has conditional heteroskedasticity ($Var(\varepsilon_t|x_t)$ depends on x_t) as well as autocorrelation ($E(\varepsilon_t \varepsilon_{t-j}) \neq 0$).

Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation

- The assumption (e) fails but assumption (d) holds (ε_t is serially uncorrelated).
- The conventional Eicker-White heteroskedasticity consistent (HC) standard error can be used.

$$\hat{V}^{HC} = \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_t{}' \right)^{-1} \hat{S}_{HC} \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_t{}' \right)^{-1}$$

where

$$\hat{S}_{HC} = \frac{1}{T} \sum_{t=1}^{T} \hat{\varepsilon}_t^2 x_t x_t{}'$$

Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation

- Both the assumption (d) and (e) fail.
- That is, ε_t is serially correlated (and hence $x_t\varepsilon_t$ is also serially correlated), then the assumption that $x_t\varepsilon_t$ is an MDS fails.
- Hence, we need to consider the long-run variance of $x_t\varepsilon_t$:

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, V)$$

where $V = Q^{-1}\Lambda Q^{-1}$ and $\Lambda = \sum_{j=-\infty}^{\infty} E(\varepsilon_t\varepsilon_{t-j}x_t x'_{t-j})$ is the long-run covariance matrix of $x_t\varepsilon_t$.

Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation

- If ε_t is serially correlated, the most popular heteroskedasticity and autocorrelation consistent (HAC) standard error, due to Newey and West (1987), has the form

$$\hat{V}^{HAC} = \left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t{}'\right)^{-1} \hat{\Lambda}_{HAC} \left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t{}'\right)^{-1}$$

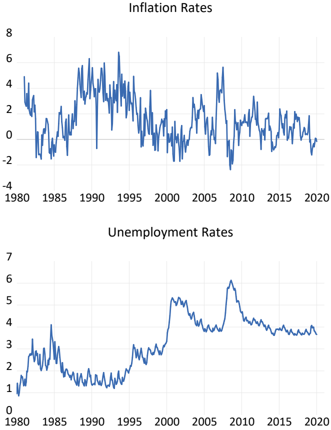
$$\hat{\Lambda}_{HAC} = \frac{1}{T}\left[\sum_{t=1}^{T} \hat{\varepsilon}_t^2 x_t x_t{}' + \sum_{\nu=1}^{q} w_\nu \sum_{t=\nu+1}^{T} (x_t \hat{\varepsilon}_t \hat{\varepsilon}_{t-\nu} x_{t-\nu}{}' + x_{t-\nu} \hat{\varepsilon}_{t-\nu} \hat{\varepsilon}_t x_t{}')\right]$$

$$w_\nu = \left(1 - \frac{\nu}{q+1}\right)$$

q is called the truncation lag, to ensure that $\hat{\Lambda}_{HAC}$ is positive-definite, and $b_T = q + 1$ is called the bandwidth.

Example: Phillips Curve

Figure: 台灣物價膨脹率 (年增率) 與失業率: 1980:M1–2021:M1



Example: Phillips Curve

- The time series regression model:

$$infl_t = \beta_0 + \beta_1 une_t + \varepsilon_t$$

- In EViews, the default is

$$q = \text{floor}(4(T/100)^{2/9})$$

- Stock and Watson (2020) suggest

$$q = \text{floor}(0.75T^{1/3})$$

- Andrews (1991) suggest

$$q = \text{floor}(1.4T^{1/3})$$

- Note that in EViews, the bandwidth setting is $b_T = q + 1$

EViews Code

```
wfcreate(wf=C5_1)  m 1980:1 2021:1
read(b4,s=C5_1) TSbookData.xls CPI une

genr infl = 100*log(CPI/CPI(-12))
equation eq1.ls infl c une
eq2.ls(cov=hac, covbwint) infl c une
```

Example: Phillips Curve

Dependent Variable: INFL
Method: Least Squares
Date: 03/08/23 Time: 08:55
Sample (adjusted): 1981M01 2020M01
Included observations: 469 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 6.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.916955	0.396769	9.872118	0.0000
UNE	-0.742245	0.102417	-7.247312	0.0000

R-squared
Adjusted R-s
S.E. of regre
Sum squared
Log likelihood
F-statistic

Dependent Variable: INFL
Method: Least Squares
Date: 03/08/23 Time: 08:32
Sample (adjusted): 1981M01 2020M01
Included observations: 469 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.916955	0.195404	20.04547	0.0000
UNE	-0.742245	0.056039	-13.24513	0.0000

Part II

Typical Time Series Regression Models

Dynamic Regression

- Often the time series regression model contains lagged variables as regressors to capture dynamic effects.
- We consider the following dynamic time series regression models:
  - AR Model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \varepsilon_t$$

  - ARDL Model

$$\begin{aligned} Y_t = \beta_0 &+ \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} \\ &+ \delta_1 Z_{t-1} + \delta_2 Z_{t-2} + \cdots + \delta_q Z_{t-q} + \varepsilon_t \end{aligned}$$

  - DL Model

$$Y_t = \beta_0 + \delta Z_t + \delta_1 Z_{t-1} + \delta_2 Z_{t-2} + \cdots + \delta_q Z_{t-q} + \varepsilon_t$$

Static DL: $\delta_1 = \delta_2 = \cdots = \delta_q = 0$

AR Models

- AR models are very popular in empirical macro and finance, in particular, the AR(1) model:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

- Although $\hat{\beta}_1$ is consistent, it is biased downward in small sample.
  - To obtain unbiasedness, we need the assumption that

$$E(\varepsilon_t | \boldsymbol{X}) = 0$$

  - But in AR(1), $x_t = (1 \;\; Y_{t-1})'$, which causes the assumption to fail.

Downward Bias of AR(1) Model