13주차 3차시 : Py 실습(시차분포모형)

1. Koyck 추정방법
2. Almon 추정방법

1. Koyck 추정방법

- 시차분포모형은 다음과 같음

$$y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_k X_{t-k} + u_t$$

- Koyck은 다음의 시차계수 구조를 가정함

$$\beta_k = \beta_0 \lambda^k$$

- 시차계수 구조를 시차분포모형에 대입하면 정리하면 ⓔ식이 됨

$$y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda y_{t-1} + \nu_t \text{ (단, } \nu_t = u_t - \lambda u_{t-1}) \quad ⓔ$$

- ⓔ식에서 y_{t-1}대신에 X_{t-1}을 사용하여 OLS로 추정
- OLS로 추정하면 $\hat{\alpha}, \hat{\beta}_0, \hat{\lambda}$을 구할 수 있고, 따라서 $\hat{\beta}_k = \hat{\beta}_0 \hat{\lambda}^k$를 이용하여 시차분포모형의 $\hat{\beta}_k$를 구할 수 있음
- 또는 ⓔ식을 최우추정법(MLE)으로 추정

1.Koyck 추정

b2-ch7-1-mle.Py

```
import pandas as pd
from statsmodels.tsa.ardl import ARDL
df = pd.read_table("http://kanggc.iptime.org/book/data/ar.txt")
df
# quarterly time series from 1995 Q1 to 2001 Q1:
start_date = "01/01/1995"
end_date = "01/01/2001"
#pd.date_range(start_date, end_date, freq='QS').quarter
df.index = pd.date_range(start_date, end_date, freq='QS')
df
print("Data Frame of df is : ",f'\n{df}\n')
#df.index = pd.date_range("04/01/1995", end_date, freq='QS')
c = df['consume']
y = df['gdp']
c_t = pd.DataFrame(c[1:25])
c_t
c_t = c_t.rename(columns={'consume': 'consume_t'})
c_t
c_t1 = pd.DataFrame(c[0:24])
c_t1
c_t1.index = pd.date_range("04/01/1995", "01/01/2001", freq='QS')
c_t1 = c_t1.rename(columns={'consume': 'consume_t1'})
c_t1
y_t = pd.DataFrame(y[1:25])
y_t
y_t = y_t.rename(columns={'gdp': 'gdp_t'})
y_t
```

b2-ch7-1-mle.Py

(앞에서 계속)

```
y_t1 = pd.DataFrame(y[0:24])
y_t1
y_t1.index = pd.date_range("04/01/1995", "01/01/2001", freq='QS')
y_t1 = y_t1.rename(columns={'gdp': 'gdp_t1'})
y_t1
df_t = pd.concat([c_t, c_t1, y_t, y_t1], axis=1)
df_t
#df_t.reset_index(inplace=True)
print("Data Frame of df_t is : ",f'\n{df_t}\n')
model = ARDL(df_t['consume_t'],1, df_t[["gdp_t"]], 0)
results = model.fit()
print(results.summary())
b = results.params
b
print("MLE Coefficients are :", f'b\n{b}\n')
beta0 = b[2]
lamb = b[1]
beta1 = beta0 * lamb
beta2 = beta0 * lamb**2
beta3 = beta0 * lamb**3
beta4 = beta0 * lamb**4
beta0, lamb
beta1, beta2, beta3, beta4
print("Estimated beta0 is : ", round(beta0, 5))
print("Estimated lambda is : ", round(lamb, 5))
print("Calculated beta1 is : ", round(beta1, 5))
print("Calculated beta2 is : ", round(beta2, 5))
print("Calculated beta3 is : ", round(beta3, 5))
print("Calculated beta4 is : ", round(beta4, 5))
```

```
                              ARDL Model Results
==============================================================================
Dep. Variable:              consume_t   No. Observations:                   24
Model:                     ARDL(1, 0)   Log Likelihood                -207.034
Method:               Conditional MLE   S.D. of innovations           1963.640
Date:                Sat, 01 Jun 2024   AIC                            422.069
Time:                        16:27:14   BIC                            426.611
Sample:                    07-01-1995   HQIC                           423.211
                         - 01-01-2001
================================================================================
                   coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------------------------
const         2.204e+04   7512.074      2.934      0.008    6368.486    3.77e+04
consume_t.L1     0.1390      0.192      0.724      0.477      -0.261       0.539
gdp_t.L0         0.3249      0.089      3.655      0.002       0.139       0.510
================================================================================
```

```
MLE Coefficients are : b
const           22038.396988
consume_t.L1        0.139025
gdp_t.L0            0.324857
```

```
Estimated beta0 is :  0.32486
Estimated lambda is :  0.13903
Calculated beta1 is :  0.04516
Calculated beta2 is :  0.00628
Calculated beta3 is :  0.00087
Calculated beta4 is :  0.00012
```

2. Almon 추정방법

- Almon은 시차계수(β_i)가 다음 식과 같이 시차의 길이인 i의 적절한 차수의 다항식의 근사치로 계산할 수 있다고 가정
  $\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \cdots + \alpha_k i^r$

- 예를 들어, r=2인 경우 다음 식과 같이 2차 다항식이 됨
  $\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2$
- 2차 다항식을 가정하는 경우 β는 α들과 다음의 관계에 있으므로 $\alpha_0, \alpha_1, \alpha_2$을 알면 β의 값들을 알 수 있음

$$\beta_0 = \alpha_0$$
$$\beta_1 = \alpha_0 + \alpha_1 + \alpha_2$$
$$\beta_2 = \alpha_0 + 2\alpha_1 + 4\alpha_2$$
$$\vdots$$
$$\beta_k = \alpha_0 + k\alpha_1 + k^2\alpha_2$$

- 이를 시차분포모형에 대입하면 다음의 ⓕ식을 얻게 됨
  $y_t = \alpha + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + u_t$ ⓕ

- ⓕ식을 OLS로 추정하여 $\hat{\alpha}, \hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2$을 구하고 β와 α의 관계를 이용하여 시차분포모형의 $\hat{\beta}_0, \hat{\beta}_1, \ldots\ldots, \hat{\beta}_k$를 계산

2.Almon 추정

b2-ch7-2.py

```
import pandas as pd
import statsmodels.formula.api as smf
data = pd.read_csv("http://kanggc.iptime.org/book/data/ar-1.txt",index_col='quarter',sep='\s+',nrows=25)
y = data['gdp']
c = data['consume']
#c1 = c.shift(1)
y1 = y.shift(1)
y2 = y.shift(2)
y3 = y.shift(3)
y4 = y.shift(4)
df = pd.DataFrame({"y":y,"c":c,"y1":y1,"y2":y2,"y3":y3,"y4":y4})
df = df.dropna()
z1 = df['y'] + df['y1'] + df['y2'] + df['y3'] + df['y4']
z2 = df['y1'] + 2*df['y2'] + 3*df['y3'] + 4*df['y4']
z3 = df['y1'] + 4*df['y2'] + 9*df['y3'] + 16*df['y4']
```

b2-ch7-2.py

```
(앞에서 계속)
almon = smf.ols(formula ='c ~ z1 + z2 + z3', data = df)
res_almon = almon.fit()
print(res_almon.summary())
print(res_almon.summary().tables[1])
b = res_almon.params
print("OLS Estimated Coefficients Almon is :",f'\n{round(b, 5)}\n')
beta0 = b[1]
beta0
beta1 = b[1] + b[2] + b[3]
beta1
beta2 = b[1] + 2*b[2] + 4*b[3]
beta2
beta3 = b[1] + 3*b[2] + 9*b[3]
beta3
beta4 = b[1] + 4*b[2] + 16*b[3]
beta4
print("Calculated Value of beta0 is :", round(beta0, 5))
print("Calculated Value of beta1 is :", round(beta1, 5))
print("Calculated Value of beta2 is :", round(beta2, 5))
print("Calculated Value of beta3 is :", round(beta3, 5))
print("Calculated Value of beta4 is :", round(beta4, 5))
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      c   R-squared:                       0.757
Model:                            OLS   Adj. R-squared:                  0.714
Method:                 Least Squares   F-statistic:                     17.62
Date:                Sat, 01 Jun 2024   Prob (F-statistic):           1.84e-05
Time:                        16:46:28   Log-Likelihood:                -187.54
No. Observations:                  21   AIC:                             383.1
Df Residuals:                      17   BIC:                             387.3
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept   3.192e+04   7704.999      4.143      0.001    1.57e+04    4.82e+04
z1             0.4386      0.126      3.491      0.003       0.174       0.704
z2            -0.3777      0.227     -1.662      0.115      -0.857       0.102
z3             0.0631      0.057      1.104      0.285      -0.058       0.184
==============================================================================
```

```
OLS Estimated Coefficients Almon is :
Intercept    31920.60289
z1               0.43862
z2              -0.37773
z3               0.06313
```

```
Calculated Value of beta0 is : 0.43862
Calculated Value of beta1 is : 0.12403
Calculated Value of beta2 is : -0.06431
Calculated Value of beta3 is : -0.12639
Calculated Value of beta4 is : -0.06222
```