6.11 ARCH and GARCH Models

Autoregressive Conditional Heteroskedasticity (ARCH)

Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

1. ARCH (p) Model

   $$\epsilon_t|\epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_1 \sim N(0, h_t),$$

   where,

   $$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \cdots + \alpha_p\epsilon_{t-p}^2.$$

   The unconditional variance of ϵ_t is:

   $$\sigma_\epsilon^2 = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \cdots - \alpha_p}$$

2. GARCH (p, q) Model

   $$\epsilon_t|\epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_1 \sim N(0, h_t),$$

where

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_p \epsilon_{t-p}^2 + \beta_1 h_{t-1} + \cdots + \beta_q h_{t-q}.$$

3. Application to OLS (Case of ARCH(1) Model):

$$y_t = x_t \beta + \epsilon_t, \qquad \epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_1 \sim N(0, \alpha_0 + \alpha_1 \epsilon_{t-1}^2).$$

The joint density of $\epsilon_1, \epsilon_2, \cdots, \epsilon_T$ is:

$$\begin{aligned} f(\epsilon_1, \cdots, \epsilon_T) &= f(\epsilon_1) \prod_{t=2}^{T} f(\epsilon_t | \epsilon_{t-1}, \cdots, \epsilon_1) \\ &= (2\pi)^{-1/2} \left(\frac{\alpha_0}{1-\alpha_1} \right)^{-1/2} \exp\left(-\frac{1}{2\alpha_0/(1-\alpha_1)} \epsilon_1^2 \right) \\ &\quad \times (2\pi)^{-(T-1)/2} \prod_{t=2}^{T} (\alpha_0 + \alpha_1 \epsilon_{t-1}^2)^{-1/2} \exp\left(-\frac{1}{2} \sum_{t=2}^{T} \frac{\epsilon_t^2}{\alpha_0 + \alpha_1 \epsilon_{t-1}^2} \right). \end{aligned}$$

The log-likelihood function is:

$$
\begin{aligned}
&\log L(\beta, \alpha_0, \alpha_1; y_1, \cdots, y_T) \\
&\quad = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\left(\frac{\alpha_0}{1-\alpha_1}\right) - \frac{1}{2\alpha_0/(1-\alpha_1)}(y_1 - x_1\beta)^2 \\
&\qquad -\frac{T-1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{T}\log\left(\alpha_0 + \alpha_1(y_{t-1} - x_{t-1}\beta)^2\right) \\
&\qquad -\frac{1}{2}\sum_{t=2}^{T}\frac{(y_t - x_t\beta)^2}{\alpha_0 + \alpha_1(y_{t-1} - x_{t-1}\beta)^2}.
\end{aligned}
$$

Obtain α_0, α_1 and β such that the log-likelihood function is maximized.

$\alpha_0 > 0$ and $\alpha_1 > 0$ have to be satisfied.

These two conditions are explicitly included, when the model is modified to: $\mathrm{E}(\epsilon_t^2|\epsilon_{t-1}, \epsilon_{t-2}, \cdot$

$\alpha_0^2 + \alpha_1^2\epsilon_{t-1}^2$.

Testing the ARCH(1) Effect:

(a) Estimate $y_t = x_t\beta + u_t$ by OLS, and compute $\hat{\beta}$ and $\hat{u}_t = y_t - x_t\hat{\beta}$.

(b) Estimate $\hat{u}_t^2 = \alpha_0 + \alpha_1\hat{u}_{t-1}^2$ by OLS. If $\hat{\alpha}_1$ is significant, there is the ARCH(1) effect in the error term.

This test corresponds to LM test.

Example: GARCH(1,1) Model

```
. arch sdex l.sdex l2.sdex, arch(1) garch(1)        → 誤差項の MA 項を GRARCH に

(setting optimization to BHHH)
Iteration 0:   log likelihood = -5089.3558
Iteration 1:   log likelihood = -5086.7468
.........
.........

Iteration 22:  log likelihood = -5064.9328  (backed up)
Iteration 23:  log likelihood = -5064.9328

ARCH family regression
```

Sample: 16 - 516 Number of obs = 501
Distribution: Gaussian Wald chi2(2) = 225.19
Log likelihood = -5064.933 Prob > chi2 = 0.0000

sdex	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
sdex						
sdex						
L1.	-.6357273	.0426939	-14.89	0.000	-.7194059	-.5520488
L2.	-.370862	.0466222	-7.95	0.000	-.4622398	-.2794842
_cons	-55.28043	261.2057	-0.21	0.832	-567.2341	456.6733
ARCH						
arch						
L1.	.041632	.0123474	3.37	0.001	.0174317	.0658324
garch						
L1.	.9526041	.0148639	64.09	0.000	.9234715	.9817367
_cons	312143.8	227564.3	1.37	0.170	-133873.9	758161.6

7 Vector Autoregressive (VAR) Model – Causality, Impulse Response Function and etc

We can consider VARMA (vector autoregressive moving average) model.

However, it is very difficult to estimate MA terms in the case of vector.

Usually, we consider VAR (Vector Autoregressive) process:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where

$$y_t : k \times 1, \qquad \mu : k \times 1, \qquad \epsilon_t : k \times 1, \qquad \phi_i : k \times k.$$

Rewriting the above equation,

$$\phi(L) y_t = \mu + \epsilon_t,$$

where $\phi(L) = I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$.

VAR(1) Model:

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \text{i.e.,} \qquad (I_k - \phi_1 L) y_t = \epsilon_t.$$

When y_t is stationary, we obtain:

$$\begin{aligned} y_t &= (I_k - \phi_1 L)^{-1} \epsilon_t \\ &= (I_k + \phi_1 L + \phi_1^2 L^2 + \phi_1^3 L^3 + \cdots) \epsilon_t \\ &= \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_1^2 \epsilon_{t-2} + \phi_1^3 \epsilon_{t-3} + \cdots \end{aligned}$$

VAR(1)=VMA(∞)

VAR(2) Model:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t, \qquad \text{i.e.,} \qquad (I_k - \phi_1 L - \phi_2 L^2) y_{t-1} = \epsilon_t.$$

When y_t is stationary, we obtain:

$$\begin{aligned} y_{t-1} &= (I_k - \phi_1 L - \phi_2 L^2)^{-1} \epsilon_t \\ &= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots \end{aligned}$$

VAR(2)=VMA(∞)

VAR(p) Model:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

i.e.,

$$(I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_t = \epsilon_t.$$

When y_t is stationary, we obtain:

$$\begin{aligned} y_t &= (I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)^{-1} \epsilon_t \\ &= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots \end{aligned}$$

VAR(p)=VMA(∞)

7.1 Autocovariance Matrix and Autocorrelation Matrix

Let y_t be a $k \times 1$ vector.

Autocovariance Function Matrix:

$$\Gamma(\tau) = \mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu)'), \qquad \tau = 0, 1, 2, \cdots,$$

where $\mathrm{E}(y_t) = \mu$. $\Gamma(\tau)$ is a $k \times k$ matrix.

$$\Gamma(\tau) = \Gamma(-\tau)'$$

Autocorrelation Function Matrix:

$$\rho(\tau) = D^{-1/2}\Gamma(\tau)D^{-1/2},$$

where the (i, j)th element of D is given by $\gamma_{ii}(0) = \mathrm{V}(y_{it})$ for $i = j$ and zero otherwise.

$$\rho(\tau) = \rho(-\tau)'$$

7.2 Granger Cuasality Test (グレンジャー因果性テスト)

Consider the bivariate case:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \phi_{11,1} & \phi_{12,1} \\ \phi_{21,1} & \phi_{22,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \phi_{11,p} & \phi_{12,p} \\ \phi_{21,p} & \phi_{22,p} \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}$$

Take an example of the first equation.

- Unrestricted Model (Sum of Squared Residuals, denoted by SSR_1):

$$\begin{aligned} y_{1,t} = \mu_1 &+ \phi_{11,1} y_{1,t-1} + \phi_{12,1} y_{2,t-1} \\ &+ \phi_{11,2} y_{1,t-2} + \phi_{12,2} y_{2,t-2} \\ &\quad \cdots \\ &+ \phi_{11,p} y_{1,t-p} + \phi_{12,p} y_{2,t-p} + \epsilon_{1,t} \end{aligned}$$

$$\longrightarrow\ \hat{\epsilon}_{1,t} \quad \longrightarrow\ \text{Unrestricted VAR}(p)\ \longrightarrow\ \text{SSR}_1 = \textstyle\sum_{t=p+1}^{T} \hat{\epsilon}_{1,t}^2$$

Test $H_0: \ \phi_{12,1} = \phi_{12,2} = \cdots = \phi_{12,p} = 0.$

When H_0 is correct, we say there is no causality from y_2 to y_1.

$\Longrightarrow$ **Granger Causality Test**

- Restricted Model (Sum of Squared Residuals, denoted by SSR_0):

Under $H_0: \ \phi_{12,1} = \phi_{12,2} = \cdots = \phi_{12,p} = 0$, we estimate the following regression:

$$
\begin{aligned}
y_{1,t} = \mu_1 &+ \phi_{11,1} y_{1,t-1} + 0 \times y_{2,t-1} \\
&+ \phi_{11,2} y_{1,t-2} + 0 \times y_{2,t-2} \\
&\quad \cdots \\
&+ \phi_{11,p} y_{1,t-p} + 0 \times y_{2,t-p} + \epsilon_{1,t} \\
= \mu_1 &+ \phi_{11,1} y_{1,t-1} \\
&+ \phi_{11,2} y_{1,t-2} \\
&\quad \cdots
\end{aligned}
$$

$$+ \phi_{11,p} y_{1,t-p} + \epsilon_1$$

$$\longrightarrow \tilde{\epsilon}_{1,t} \quad \longrightarrow \text{Restricted VAR}(p) \longrightarrow \text{SSR}_0 = \sum_{t=p+1}^{T} \tilde{\epsilon}_{1,t}^2$$

The number of parameters to be estimated:

- Unestricted Model: $2p + 1$

  $\longrightarrow$ VAR(p), p lagged coefficients, one constant term $\longrightarrow$ $2p + 1$

- Restricted Model: $p + 1$

  $\longrightarrow$ The number of restrictions = p (= G)

Therefore, asymptotically we have the following distribution:

$$F = \frac{(\text{SSR}_0 - \text{SSR}_1)/p}{\text{SSR}_1/(T - 2p - 1)} \sim F(p, T - 2p - 1),$$

or

$$pF \sim \chi^2(p).$$

In general, we consider testing the Granger causality from y_j to y_i.

VAR(p) model:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where $y_t : k \times 1$, $\mu : k \times 1$, $\phi_p : k \times k$, $\epsilon_t : k \times 1$.

ϕ_p is given by:

$$\phi_p = \begin{pmatrix} \phi_{11,p} & \cdots & \phi_{1j,p} & \cdots & \phi_{1k,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_{i1,p} & \cdots & \phi_{ij,p} & \cdots & \phi_{ik,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_{k1,p} & \cdots & \phi_{kj,p} & \cdots & \phi_{kk,p} \end{pmatrix}$$

Test the Granger causality from y_j to y_i.

$\longrightarrow$ Focus on the ith equation.

$$
\begin{aligned}
y_{i,t} = \mu_i &+ \phi_{i1,1}y_{1,t-1} + \cdots + \phi_{ij,1}y_{j,t-1} + \cdots + \phi_{ik,1}y_{k,t-1} \\
&+ \phi_{i1,2}y_{1,t-2} + \cdots + \phi_{ij,2}y_{j,t-2} + \cdots + \phi_{ik,2}y_{k,t-2} \\
&\quad \cdots \\
&+ \phi_{i1,p}y_{1,t-p} + \cdots + \phi_{ij,p}y_{j,t-p} + \cdots + \phi_{ik,p}y_{k,t-p} + \epsilon_{i,t}
\end{aligned}
$$

Supose that the null hypothesis is: $H_0: \ \phi_{ij,1} = \phi_{ij,2} = \cdots = \phi_{ij,p} = 0$, while the alternative hypothesis is: $H_1: \ \text{not } H_0$.

$$
\begin{aligned}
\text{SSR}_0 &= \text{Sum of Squared Residuals under } H_0 \\
\text{SSR}_1 &= \text{Sum of Squared Residuals under } H_1
\end{aligned}
$$

Under H_0, the asymptotic distribution is given by:

$$F = \frac{(\mathrm{SSR}_0 - \mathrm{SSR}_1)/p}{\mathrm{SSR}_1/(T-kp-1)} \sim F(p, T-kp-1),$$

or

$$pF \sim \chi^2(p).$$

Example: VAR(p): Data: 1994 年第一四半期～2014 年第一四半期

`gdp` = GDP (実質，10 億円，季調済，内閣府 HP から取得)

`def` = GDP デフレータ (季調済，内閣府 HP から取得)

`r` = 貸出約定平均金利 (%，新規，総合・国内銀行，日銀 HP から取得)

`m` = 通貨流通高（平均発行高，億円，季調済，日銀 HP から取得)

```
. gen time=_n                                    → time のデータ作成

. tsset time                                     → time が時間データとする
        time variable:  time, 1 to 81
                delta:  1 unit

. gen lgdp=log(gdp)                              → gdp の対数変換

. gen lm=log(m/(def/100))                        → マネーサプライ m を実質化で，対数変換

. varsoc d.lgdp d.r d.lm                         → 各変数階差を取って，ラグ次数を決める

 Selection-order criteria
 Sample:  6 - 81                              Number of obs      =        76
+---------------------------------------------------------------------------+
|lag |    LL      LR      df    p      FPE       AIC      HQIC      SBIC    |
|----+----------------------------------------------------------------------|
|  0 |  541.22                        1.4e-10  -14.1637  -14.1269  -14.0717  |
|  1 | 571.181  59.923*   9  0.000  8.2e-11* -14.7153* -14.5682* -14.3473* |
|  2 | 575.715  9.0675    9  0.431  9.2e-11  -14.5978  -14.3404  -13.9537  |
|  3 |  579.55  7.6704    9  0.568  1.1e-10  -14.4619  -14.0942  -13.5418  |
|  4 | 583.767  8.4328    9  0.491  1.2e-10   -14.336   -13.858  -13.1399  |
+---------------------------------------------------------------------------+
 Endogenous:  D.lgdp D.r D.lm
  Exogenous:  _cons
```

```
. var d.lgdp d.r d.lm, lags(1)
```

→ 各変数階差を取って，3 変数 VAR(1)

```
Vector autoregression

Sample:  3 - 81                                    No. of obs      =        79
Log likelihood =  592.2334                         AIC             = -14.68945
FPE            =  8.38e-11                         HQIC            = -14.54526
Det(Sigma_ml)  =  6.18e-11                         SBIC            = -14.32954

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
D_lgdp               4     .010717   0.0422   3.480972   0.3232
D_r                  4     .087186   0.2553    27.0782   0.0000
D_lm                 4     .009434   0.2903   32.30929   0.0000
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
D_lgdp       |
        lgdp |
         LD. |   .2031129   .1119361     1.81   0.070    -.0162778    .4225037
             |
           r |
         LD. |   .0045431   .0120151     0.38   0.705    -.0190061    .0280922
             |
          lm |
         LD. |   .0152162   .1086739     0.14   0.889    -.1977807     .228213
             |
       _cons |   .0019504   .0019124     1.02   0.308    -.0017978    .0056986
-------------+----------------------------------------------------------------
```

D_r						
lgdp						
LD.	.4341641	.9106374	0.48	0.634	-1.350652	2.218981
r						
LD.	.5085677	.0977469	5.20	0.000	.3169874	.7001481
lm						
LD.	.1845222	.8840978	0.21	0.835	-1.548278	1.917322
_cons	-.0202984	.0155578	-1.30	0.192	-.0507912	.0101943
D_lm						
lgdp						
LD.	-.1972406	.098541	-2.00	0.045	-.3903774	-.0041037
r						
LD.	-.029395	.0105773	-2.78	0.005	-.0501261	-.0086639
lm						
LD.	.4472679	.0956691	4.68	0.000	.2597599	.634776
_cons	.0071036	.0016835	4.22	0.000	.0038039	.0104033

```
. vargranger
```

→ Granger Cuasality Test (グレンジャー因果性テスト)

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
D_lgdp	D.r	.14297	1	0.705
D_lgdp	D.lm	.0196	1	0.889
D_lgdp	ALL	.15705	2	0.924
D_r	D.lgdp	.22731	1	0.634
D_r	D.lm	.04356	1	0.835
D_r	ALL	.3039	2	0.859
D_lm	D.lgdp	4.0064	1	0.045
D_lm	D.r	7.7232	1	0.005
D_lm	ALL	10.798	2	0.005

→ gdp から m への因果関係（p 値 0.045）
r から m への因果関係（p 値 0.005）

→ 学部の教科書レベルの貨幣需要関数
m=f(gdp,r)，すなわち，m は gdp と r の関数