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

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

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

Pick up the ith equation:

$$y_{i,t} = \mu_i + \begin{pmatrix} \phi_{i1,1} & \phi_{i2,1} & \cdots & \phi_{ik,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \begin{pmatrix} \phi_{i1,2} & \phi_{i2,2} & \cdots & \phi_{ik,2} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix}$$
$$+ \cdots + \begin{pmatrix} \phi_{i1,p} & \phi_{i2,p} & \cdots & \phi_{ik,p} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \epsilon_{i,t}.$$

The null hypothesis is: $H_0: \phi_{ij,1} = \phi_{ij,2} = \cdots = \phi_{ij,p} = 0.$

The alternative hypothesis is: $H_1:$ not H_0.

SSR_0 = Sum of Squared Residuals under H_0, which is computed from:

$$
\begin{aligned}
y_{i,t} = \mu_i &+ (\,\phi_{i,1,1} \quad \cdots \quad \phi_{i,j-1,1} \quad \phi_{i,j+1,1} \quad \cdots \quad \phi_{i,k,1}\,) \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} \\
&+ (\,\phi_{i,1,2} \quad \cdots \quad \phi_{i,j-1,2} \quad \phi_{i,j+1,2} \quad \cdots \quad \phi_{i,k,2}\,) \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \cdots \\
&+ (\,\phi_{i,1,p} \quad \cdots \quad \phi_{i,j-1,p} \quad \phi_{i,j+1,p} \quad \cdots \quad \phi_{i,k,p}\,) \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \epsilon_{i,t}.
\end{aligned}
$$

$\Longrightarrow$ Restricted Model

SSR_1 = Sum of Squared Residuals under H_1, which is computed from:

$$y_{i,t} = \mu_i + \begin{pmatrix} \phi_{i1,1} & \cdots & \phi_{ik,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \begin{pmatrix} \phi_{i1,2} & \cdots & \phi_{ik,2} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \cdots$$

$$+ \begin{pmatrix} \phi_{i1,p} & \cdots & \phi_{ik,p} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{k,t-1} \end{pmatrix} + \epsilon_{i,t}.$$

$\Longrightarrow$ Unrestricted Model

Under H_0, the asymptotic distribution is given by:

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

or

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

Example:

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

. tsset time
        time variable:  time, 1 to 81
                delta:  1 unit

. gen lgdp=log(gdp)

. gen lm=log(m/(def/10))

. varsoc d.lgdp d.r d.lm

   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)

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 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

8.3 Impulse Response Function (インパルス応答関数):

$$\frac{\partial y_{i,t+m}}{\partial \epsilon_{j,t}}, \qquad m = 1, 2, \cdots,$$

where $i, j = 1, 2, \cdots, k$.

Example: AR(p) Process:

When y_t is stationary, we obtain:

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

The impulse response function is:

$$\frac{\partial y_{i,t+k}}{\partial \epsilon_{j,t}} = \theta_{ij,k}, \qquad k = 1, 2, \cdots,$$

where $\theta_{ij,k}$ denotes the (i, j)th element of θ_k.

$$\begin{aligned} y_t &= \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots \\ &= PP^{-1}\epsilon_t + \theta_1 PP^{-1}\epsilon_{t-1} + \theta_2 PP^{-1}\epsilon_{t-2} + \cdots \\ &= \Omega_0\eta_t + \Omega_1\eta_{t-1} + \Omega_2\eta_{t-2} + \cdots, \end{aligned}$$

where $\mathrm{V}(\eta_t) = I_k$, and $\Omega_i = \theta_i P$ for $i = 0, 1, 2, \cdots$ and $\Omega_0 = P$.

$$\frac{\partial y_{i,t+m}}{\partial \eta_{j,t}}, \qquad m = 1, 2, \cdots,$$

where $i, j = 1, 2, \cdots, k$.

$\Longrightarrow$ **Orthogonalized Impulse Response Function** (直交化インパルス応答関数)

Example:

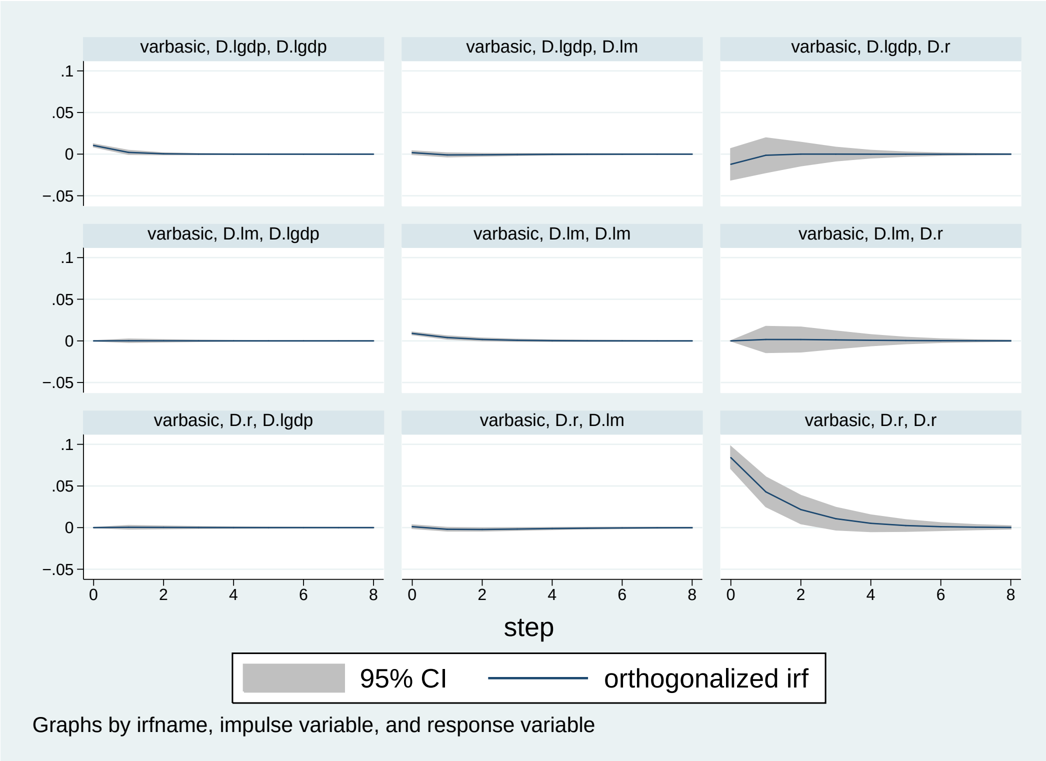
```
. varbasic d.lgdp d.r d.lm, lags(1)
Vector autoregression
```

```
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
------------------------------------------------------------------------------
```


varbasic, D.lgdp, D.lgdp
varbasic, D.lgdp, D.lm
varbasic, D.lgdp, D.r
varbasic, D.lm, D.lgdp
varbasic, D.lm, D.lm
varbasic, D.lm, D.r
varbasic, D.r, D.lgdp
varbasic, D.r, D.lm
varbasic, D.r, D.r
.1
.05
0
−.05
0
2
4
6
8
step
95% CI
orthogonalized irf
Graphs by irfname, impulse variable, and response variable

9 Unit Root (単位根) and Cointegration (共和分)

9.1 Unit Root (単位根) Test (Dickey-Fuller (DF) Test)

1. Why is a unit root problem important?

   (a) Economic variables increase over time in general.

   One of the assumptions of OLS is stationarity on y_t and x_t.

   This assumption implies that $\dfrac{1}{T}X'X$ converges to a fixed matrix as T is large.

   That is, asymptotic normality of OLS estimator goes not hold.

   (b) In nonstationary time series, the unit root is the most important.

   In the case of unit root, OLSE of the first-order autoregressive coefficient is consistent.

OLSE is $\sqrt{T}$-consistent in the case of stationary AR(1) process, but OLSE is T-consistent in the case of nonstationay AR(1) process.

(c) A lot of economic variables increase over time.

It is important to check an economic variable is trend stationary (i.e., $y_t = a_0 + a_1 t + \epsilon_t$) or difference stationary (i.e., $y_t = b_0 + y_{t-1} + \epsilon_t$).

Consider k-step ahead prediction for both cases.

$$\text{(Trend Stationarity)} \qquad y_{t+k|t} = a_0 + a_1(t+k)$$

$$\text{(Difference Stationarity)} \qquad y_{t+k|t} = b_0 k + y_t$$

2. **The Case of $|\phi_1| < 1$:**

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } N(0, \sigma_\epsilon^2), \qquad y_0 = 0, \qquad t = 1, \cdots, T$$

Then, OLSE of ϕ_1 is:

$$\hat{\phi}_1 = \frac{\displaystyle\sum_{t=1}^{T} y_{t-1} y_t}{\displaystyle\sum_{t=1}^{T} y_{t-1}^2}.$$

In the case of $|\phi_1| < 1$,

$$\hat{\phi}_1 = \phi_1 + \frac{\displaystyle\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\displaystyle\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow \phi_1 + \frac{\mathrm{E}(y_{t-1}\epsilon_t)}{\mathrm{E}(y_{t-1}^2)} = \phi_1.$$

Note as follows:

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow \mathrm{E}(y_{t-1}\epsilon_t) = 0.$$

By the central limit theorem,

$$\frac{\overline{y\epsilon} - \mathrm{E}(\overline{y\epsilon})}{\sqrt{\mathrm{V}(\overline{y\epsilon})}} \longrightarrow N(0,1)$$

where

$$\overline{y\epsilon} = \frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t.$$

$$\begin{aligned}
\mathrm{E}(\overline{y\epsilon}) &= 0, \\
\mathrm{V}(\overline{y\epsilon}) &= \mathrm{V}(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t) = \mathrm{E}\Big((\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t)^2\Big) \\
&= \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T}\sum_{s=1}^{T} y_{t-1}y_{s-1}\epsilon_t\epsilon_s\Big) = \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T} y_{t-1}^2\epsilon_t^2\Big) = \frac{1}{T}\sigma^2\gamma(0).
\end{aligned}$$

Therefore,

$$\frac{\overline{y\epsilon}}{\sqrt{\sigma^2\gamma(0)/T}} = \frac{1}{\sigma_\epsilon\sqrt{\gamma(0)}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0,1),$$

which is rewritten as:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0, \sigma_\epsilon^2\gamma(0)).$$

Using $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \mathrm{E}(y_{t-1}^2) = \gamma(0)$, we have the following asymptotic distribution:

$$\sqrt{T}(\hat{\phi}_1 - \phi_1) = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow N\left(0, \frac{\sigma_\epsilon^2}{\gamma(0)}\right) = N\left(0, 1-\phi_1^2\right).$$

Note that $\gamma(0) = \dfrac{\sigma_\epsilon^2}{1-\phi_1^2}$.

3. In the case of $\phi_1 = 1$, as expected, we have:

$$\sqrt{T}(\hat{\phi}_1 - 1) \longrightarrow 0.$$

That is, $\hat{\phi}_1$ has the distribution which converges in probability to $\phi_1 = 1$ (i.e., degenerated distribution).

Is this true?

4. **The Case of $\boldsymbol{\phi_1 = 1}$:** $\Longrightarrow$ Random Walk Process

$y_t = y_{t-1} + \epsilon_t$ with $y_0 = 0$ is written as:

$$y_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \cdots + \epsilon_1.$$

Therefore, we can obtain:

$$y_t \sim N(0, \sigma_\epsilon^2 t).$$

The variance of y_t depends on time t. $\Longrightarrow$ y_t is nonstationary.

5. Remember that $\hat{\phi}_1 = \phi_1 + \dfrac{\sum y_{t-1}\epsilon_t}{\sum y_{t-1}^2}$.

(a) First, consider the numerator $\sum y_{t-1}\epsilon_t$.

We have $y_t^2 = (y_{t-1} + \epsilon_t)^2 = y_{t-1}^2 + 2y_{t-1}\epsilon_t + \epsilon_t^2$.

Therefore, we obtain:

$$y_{t-1}\epsilon_t = \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2).$$

Taking into account $y_0 = 0$, we have:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2.$$

Divided by $\sigma_\epsilon^2 T$ on both sides, we have the following:

$$\frac{1}{\sigma_\epsilon^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2.$$

From $y_t \sim N(0, \sigma_\epsilon^2 t)$, we obtain the following result:

$$\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 \sim \chi^2(1).$$

Moreover, the second term is derived from:

$$\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2 \longrightarrow \sigma_\epsilon^2.$$

Therefore,

$$\frac{1}{\sigma_\epsilon^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma\sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2 \longrightarrow \frac{1}{2}(\chi^2(1)-1).$$

(b) Next, consider $\sum y_{t-1}^2$.

$$\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) = \sum_{t=1}^{T}\mathrm{E}(y_{t-1}^2) = \sum_{t=1}^{T}\sigma_\epsilon^2(t-1) = \sigma_\epsilon^2\frac{T(T-1)}{2}.$$

Thus, we obtain the following result:

$$\frac{1}{T^2}\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) \longrightarrow \text{a fixed value.}$$

Therefore,

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \text{ a distribution.}$$

6. Summarizing the results up to now, $T(\hat{\phi}_1 - \phi_1)$, not $\sqrt{T}(\hat{\phi}_1 - \phi_1)$, has limiting distribution in the case of $\phi_1 = 1$.

$$T(\hat{\phi}_1 - \phi_1) = \frac{(1/T)\sum y_{t-1}\epsilon_t}{(1/T^2)\sum y_{t-1}^2} \longrightarrow \text{ a distribution.}$$

The distributions of the t statistic: $\dfrac{\hat{\phi}_1 - 1}{s_\phi}$, where s_ϕ denotes the standard error of $\hat{\phi}_1$.

$\Longrightarrow$ Compare t distribution with (a) – (c).

$\Longrightarrow$ **Unit Root Test (単位根検定, or Dickey-Fuller (DF) Test)**