Cointegration

Series de Tiempo

UTDT

January 2020

Cointegration

Importance of checking for cointegration:

- If variables do not cointegrate there is not a meaningful relationship between those variables, then Change the Theory.
- If variables do cointegrate we need to write the model correctly.

STRUCTURE OF THE LECTURE

- Definition and Properties.
- Valid representations if the variables cointegrate
- Testing for Cointegration.

Cointegration

- A Cointegration relationship captures the long-run relationship between the variables of interest.
- The aim behind cointegration is the detection and analysis of long run relationships amongst economic time series variables.
- Cointegration analysis provides a way of retaining both short-run and long-run information.
- Cointegration is sometimes thought to be a pre-requisite for the validity of some economic theory.

Definition

Consider two $I(1)$ processes $\{Y_t\}$ and $\{X_t\}$.Their linear combination maybe $I(1)$ or $I(0)$. X and Y are said to cointegrate if there exists (a_1, a_2) such that $a_1 Y + a_2 X$ ˜ $I(0)$.

Example

Consider the following example where cointegration of Prices and Dividends is a necessary condition for markets efficiency in the Fama sense. Let us assume that stock prices might be written as

$$P_t = \sum_{i=1}^{\infty} (1/(1+r))^i E(D_{t+i}|I_t) + \varepsilon_t$$

where we assume that ε_t is an I(0) process

Example

Let also assume D_t follows a random walk which is a special case of an I(1) variable.

$$D_t = D_{t-1} + \nu_t.$$

Then, we may express stock prices as

$$P_t = (1/r)D_t + \varepsilon_t$$

Example

Given that D_t are integrated of order one, stock prices also are integrated of order one. If the theory is valid, $(1, -(1/r))$ is going to be a cointegrating vector since

$$(1, -(1/r)) \begin{bmatrix} P_t \\ D_t \end{bmatrix} = Z_t = \varepsilon_t$$

Notice that we assumed that ε_t was I(0), therefore if the theory holds dividends and prices should be cointegrated.

- There are cases where there exist a theoretical long run cointegration relationship dictated by the model. In the previous example is not enough that Prices and Dividends cointegrate, the cointegrating vector has to be $(1, -(1/r))$ for the theory to hold.

Example

Consider the following model:

$$x_t + \beta y_t = u_t$$

$$x_t + \alpha y_t = e_t$$

$$u_t = u_{t-1} + \varepsilon_{1t}$$

$$e_t = \rho e_{t-1} + \varepsilon_{2t} \quad \text{with } |\rho| < 1$$

$(\varepsilon_{1t}, \varepsilon_{2t})'$ is distributed identically and independently as a bivariate normal with $E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0$, $var(\varepsilon_{1t}) = \sigma_{11}$, $var(\varepsilon_{2t}) = \sigma_{22}$ $cov(\varepsilon_{1t}\varepsilon_{2t}) = \sigma_{12}$

Solving for x_t and y_t form the above system with $\alpha \neq \beta$ gives

$$
\begin{aligned}
x_t &= \alpha(\alpha-\beta)^{-1}u_t - \beta(\alpha-\beta)^{-1}e_t, \\
y_t &= -(\alpha-\beta)^{-1}u_t + (\alpha-\beta)^{-1}e_t.
\end{aligned}
$$

Then:

1. Both x_t and y_t are integrated of order one i.e., $x_t \tilde{\ } I(1)$, $y_t \tilde{\ } I(1)$, since u_t is integrated of order one.
2. $x_t + \alpha y_t$ is $I(0)$ because e_t is stationary.
3. The cointegration vector is $(1, \alpha)$ and $x + \alpha y$ is the equilibrium relationship.

Proposition *In the bivariate case if the equilibrium condition exists, is unique.*

Proof.

Suppose that there exist two distinct co-integrating parameters α and γ such that $x + \alpha y$ and $x + \gamma y$ are both ˜$I(0)$. This implies that $(\alpha - \gamma)y_t$ is also $I(0)$ because a linear combination of two I(0) variable is also I(0). But we know that for $\alpha \neq \gamma$, $(\alpha - \gamma)y_t$ ˜I(1) therefore we have a contradiction unless $\alpha = \gamma$. □

Representation Theorems

- Consider the model in the first example, with $|\rho| < 1$.
- Whenever x_t and y_t are cointegrated, we can show, for the simple two variable model, that the system of two equations has the following representations:
  - **Error-correction representation**
  - **Vector autoregressive in levels representation**
  - **Moving-average representation.**

Error correction mechanism

- There is a strong empirical motivation for these models since they perform very well.
- This representation has the advantage that it keeps long run and short run information.
- Allows to rewrite RE models which variables are I(1) in terms of I(0) variables.

ECM Representation

I repeat the above model for convenience:

$$x_t + \beta y_t = u_t$$

$$x_t + \alpha y_t = e_t$$

$$u_t = u_{t-1} + \varepsilon_{1t}$$

$$e_t = \rho e_{t-1} + \varepsilon_{2t} \quad \text{with } |\rho| < 1$$

Lagging the equations and substracting the lagged values:

$$\Delta x_t + \beta \Delta y_t = \Delta u_t$$

$$\Delta x_t + \alpha \Delta y_t = \Delta e_t$$

and rewriting the equations we have

$$\begin{bmatrix} 1 & \beta \\ 1 & \alpha \end{bmatrix} \begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} \Delta u_t \\ \Delta e_t \end{bmatrix}$$

- Inverting the matrix and noting that

$$\begin{aligned} \Delta u_t &= \varepsilon_{1t} \\ \Delta e_t &= -(1-\rho)e_{t-1} + \varepsilon_{2t} \end{aligned}$$

we have

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \alpha & -\beta \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ -(1-\rho)e_{t-1} + \varepsilon_{2t} \end{bmatrix}$$

- The ECM representation would be:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \beta(1-\rho)(e_{t-1}) \\ -(1-\rho)(e_{t-1}) \end{bmatrix} + \begin{bmatrix} \zeta_{1t} \\ \zeta_{2t} \end{bmatrix}$$

where

$$\begin{bmatrix} \zeta_{1t} \\ \zeta_{2t} \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \alpha & -\beta \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

VAR representation

Now, notice that:

$$
\begin{aligned}
\Delta u_t &= \varepsilon_{1t} \\
\Delta e_t &= -(1-\rho)e_{t-1} + \varepsilon_{2t} \\
&= -(1-\rho)(x_{t-1} + \alpha y_{t-1}) + \varepsilon_{2t}
\end{aligned}
$$

Then using the ECM representation

$$
\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \beta(1-\rho)(e_{t-1}) \\ -(1-\rho)(e_{t-1}) \end{bmatrix} + \begin{bmatrix} \zeta_{1t} \\ \zeta_{2t} \end{bmatrix}
$$

This can be rewritten to obtain:

$$
\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \beta(1-\rho)(x_{t-1} + \alpha y_{t-1}) \\ -(1-\rho)(x_{t-1} + \alpha y_{t-1}) \end{bmatrix} + \begin{bmatrix} \zeta_{1t} \\ \zeta_{2t} \end{bmatrix}
$$

$$
\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \beta(1-\rho) & \beta(1-\rho)\alpha \\ -(1-\rho) & -(1-\rho)\alpha \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta_{1t} \\ \zeta_{2t} \end{bmatrix}
$$

Notice that this VAR representation **is not** a VAR in the first differences. The VAR can be written as

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \frac{1}{\alpha-\beta} \begin{bmatrix} \beta(1-\rho)+(\alpha-\beta) & \beta(1-\rho)\alpha \\ -(1-\rho) & -(1-\rho)\alpha+(\alpha-\beta) \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta_{1t} \\ \zeta_{2t}$$

Moving Average Representation

Using the fact that

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \alpha & -\beta \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \Delta u_t \\ \Delta e_t \end{bmatrix},$$

and noting that,

$$\begin{bmatrix} \Delta u_t \\ \Delta e_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ (1 - L)(1 - \rho L)^{-1} \varepsilon_{2t} \end{bmatrix},$$

The MA representation is given by:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} \alpha & -\beta \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ (1 - L)(1 - \rho L)^{-1} \varepsilon_{2t} \end{bmatrix}.$$

- The MA polynomial has a root equal to 1 and therefore is not invertible
- There isn't a VAR in differences representation.

Cointergration and Trends

- Two variables that are I(1) may cointegrate to an trend stationary representation.

Examples

Consider the following procesess for X_t and Y_t :

$$\begin{aligned} X_t &= \mu_x + X_{t-1} + \nu_t, \\ Y_t &= \mu_y + Y_{t-1} + \varepsilon_t. \end{aligned}$$

X and Y are said to cointegrate if there exists (a_1, a_2) such that $a_1 X + a_2 Y$ ˜ $I(0)$. From the above example is eassy to see that the variables may cointegrate to a trend stationary process since:

$$\begin{aligned} X_t &= \mu_x t + X_0 + \nu_t + \nu_{t-1} + \ldots\ldots\ldots + \nu_1, \\ Y_t &= \mu_y t + Y_0 + \varepsilon_t + \varepsilon_{t-1} + \ldots\ldots\ldots + \varepsilon_{1.} \end{aligned}$$

then we know that $a_1(\nu_t + \nu_{t-1} + \ldots\ldots\ldots + \nu_1) + a_2(\varepsilon_t + \varepsilon_{t-1} + \ldots\ldots\ldots + \varepsilon_1)$ ˜ $I(0)$, but $(a_1\mu_x + a_2\mu_y)t$ and/or $(a_1 X_0 + a_2 Y_0)$ might be different from zero.

Cointegration in the n variable case

Definition

Consider a n×1 vector of stochastic variables $X_t = (X_{1t}, X_{2t}, .., X_{n_t})$. We say that the elements of the vector are cointegrated of order (d, b), which we denote X_t ~CI(d, b) if

1. Each of the components of X_t are I(d).
2. There exists (at least) a vector such that $Z_t = \alpha' X_t$ is I$(d - b)$ for $d \geq b > 0$.

Then, α is called the cointegrating vector

Comments

1. If $d = b = 0$, then $\alpha' X_t = 0$ defines a long-run equilibrium relationship.
2. Notice that α is not unique.
3. If $n > 2$, there may be $r \leq n - 1$ linearly independent $(n \times 1)$ vectors $(\alpha_1, ..., \alpha_r)$ such that $A' X_t \sim I(0)$
4. The vectors $(\alpha_1, ..., \alpha_r)$ are not unique

Granger Representation Theorem

Theorem

Consider an n - vector time series X_t *which satisfies:*

$$\Phi(L)X_t = c + u_t$$

where $\Phi(L) = I_n - \sum_{i=1}^{P} \Phi_i L^i$ *, and* u_t *is a white noise with positive definite covariance matrix. It is assumed that* det$[\Phi(z)]= 0$ *which implies* $|z| \geq 1$*. Suppose that there exist exactly r cointegrating relationships among the elements of* X_t *. Then:*

1. *there exists an* $(n \times r)$ *matrix A, of rank* $r < n$ *such that* $A'X_t$ ~*I(0).*
2. ΔX_t *has an MA representation given by* $\Delta X_t = \mu + \Psi(L)u_t$ *with*

$$A'\Psi(1) = 0,$$

where

$$\Psi(L) = I_n + \sum_{i=1}^{\infty} \Psi_i L^i$$

Theorem

(continues)

[3.]$\Phi(1) = BA'$ *There exist a VAR representation in levels and the determinant of the polynomial in* Φ *has a unit root.There exists an ECM representation*

Error Correction Representation

Transform the original VAR

$$\Phi(L)X_t = c + u_t,$$

using the following relationship

$$I_n - \sum_{i=1}^{p} \Phi_i L^i = I_n - (\sum_{i=1}^{P} \Phi_i) L - (I - L) \sum_{i=1}^{p-1} \Gamma_i L^i$$

where

$$\Gamma_i = -\sum_{i=j}^{p-1} \Phi_{i+1} \qquad \textit{for } j = 1, ..., p-1$$

Then

$$(I_n - \sum_{i=1}^{p} \Phi_i L^i) X_t = \left(I_n - \left(\sum_{i=1}^{p} \Phi_i L + (I - L) \sum_{i=1}^{p-1} \Gamma_i L^i \right) \right) X_t = c + u_t.$$

Rearranging terms

$$\begin{aligned} X_t &= c + (\sum_{i=1}^{p} \Phi_i) X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + u_t \\ &= c + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + (I_n - \Phi(1)) X_{t-1} + u_t \end{aligned}$$

$$\Delta X_t = c + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} - \Phi(1) X_{t-1} + u_t$$

Thus:

- (a) If rank $[\Phi(1)] = 0$ then $\Phi(1) = 0$ and $X_t \tilde{} I(1)$
- (b) If rank $[\Phi(1)] = n$ then $det(\Phi(1)) \neq 0$ ($\Phi(L)$ does not have a unit root). This implies that $X_t \tilde{} I(0)$
- (c) If rank $[\Phi(1)] = r, \; 0 < r < n$ then $\Phi(1) = BA'$, where B is an nxr matrix .

Restrictions on the parameters of the ECM representation

$$\Delta X_t = c + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} - BZ_{t-1} + u_t, \; Z_t = A' X_t$$

- $A'X_t$ gives the "error" from the long run equilibrium relationship, and B gives the "correction" to X_t
- Taking expected values in both sides we get;

$$[I - \sum_{i=1}^{p-1} \Gamma_i L^i] E(\Delta X_t) = c - BE(Z_{t-1})$$

- If we assume that there is no autonomous growth component

$$[I - \sum_{i=1}^{p-1} \Gamma_i L^i] E(\Delta X_t) = c - BE(Z_{t-1}) = 0$$

or

$$c = BE(Z_{t-1})$$

- Thus,

$$\begin{aligned} \Delta X_t &= BE(Z_{t-1}) + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + -BZ_{t-1} + u_t \\ &= -B(Z_{t-1} - E(Z_{t-1})) + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + u_t \end{aligned}$$

- The intercept enters the system only via the error-correction term and there is no autonomous growth component.

Tests for Cointegration: Univariate

Two stages Aproaches:First stage

- Consider the following model:

$$y_t = \beta x_t + \varepsilon_t$$

- And suppose that the actual model is the following dynamic model

$$y_t = \gamma_o x_t + \gamma_1 x_{t-1} + \alpha y_{t-1} + \varepsilon_t$$

which can be re-written as

$$y_t = \lambda_0 x_t + \lambda_1 \Delta x_t + \lambda_2 \Delta y_t + \varepsilon_t$$

where

$$\lambda_0 = \frac{\gamma_0 + \gamma_1}{1-\alpha}, \lambda_1 = \frac{-\gamma_1}{1-\alpha}, \lambda_2 = \frac{-\alpha}{1-\alpha}.$$

- Then the Static model, long-run parameter $\beta \Leftrightarrow$ dynamic model without the short run terms. The omitted dynamic terms are captured by the residuals and since the variables in levels converge at speed T, and those in differences at $.\sqrt{T}$, that won't bias the estimates of the static model.

Two stage approaches Second stage

As a second stage, we can use alternative testing strategies

1. **Make an (ADF) test for unit roots for the residuals**: *The Engle - Granger Approach.*
2. **The null may be tested using the Sargan-Bhargava or CDW test**

The Engle - Granger Approach

- If you do not reject the Hypothesis that the residuals have a unit root, then Y and X are not cointegrated.
- Regress

$$\Delta\hat{\varepsilon}_t = \phi\hat{\varepsilon}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta\hat{\varepsilon}_{t-i} + \mu + \delta t + \zeta_t$$

- Standard Dickey - Fuller critical values: over reject the null of no cointegration. See MacKinon

The cointegrating Durbin Watson

- This test is very simple and consist in comparing the DW statistic with tabulated values.
- Consider equations :

$$x_t + \alpha y_t = e_t$$

$$e_t = \rho e_{t-1} + \varepsilon_{2t} \quad \text{with } |\rho| < 1$$

- A regression of x_t on y_t will yield serially correlated residuals. We can use the DW statistic to get information about ρ since this statistic is approximately $2(1-\rho)$.

$$DW \cong 2(1-\rho)$$

Three stages: The Engle - Granger - Yoo approach

- They assume a unique cointegrating vector and weak exogeneity of the short run parameters
- Third step: provides a correction for the first stage estimate of β
- They correct the long run relationship by the small sample bias $\frac{\gamma}{1-\alpha}$

The Johansen approach

- Define X_t, a vector of n potentially endogenous variables.
- It is possible to specify the following DGP and the model X_t as an unrestricted VAR involving p lags of X_t :

$$\Phi(L)X_t = c + u_t$$

- Using the ECM

$$\Delta X_t = c + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} - \Phi(1)X_{t-1} + u_t$$

where $\Gamma_i = -\sum_{i=j}^{p-1} \Phi_{i+1}$ *for* $j = 1, ..., p-1$, and $\Phi(1) = I_n - \sum_{i=1}^{P} \Phi_i$.

Consequently testing for cointegration amounts to a consideration of the rank of $\Phi(1)$, that is, finding the number of r linearly independent columns in $\Phi(1)$.

1. If $\Phi(1)$ is full rank the variables in X_t have to be I(0)
2. If $\Phi(1)$ is has zero rank there is no cointegrating vector
3. If $\Phi(1)$ is reduced rank the number of cointegrating vectors is the RANK($\Phi(1)$)

Canonical correlations

Population Canonical Correlations

Let the $(n_1 \times 1)$ vector y_t and the $(n_2 \times 1)$ vector x_t denote stationary random variables. In general

$$\begin{bmatrix} E(y_t y_t') & E(y_t x_t') \\ E(x_t y_t') & E(x_t x_t') \end{bmatrix} = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{bmatrix}$$

We can often gain some insight into the nature of these correlations by defining two new $(n \times 1)$ random vectors φ_t and ξ_t , where n is the smaller of n_1 and n_2 .

$$\begin{aligned} \varphi_t &= K' y_t \\ \xi_t &= A' x_t \end{aligned}$$

The matrices K' and A' are chosen such that

$$\begin{bmatrix} E(y_t y_t') & E(y_t x_t') \\ E(x_t y_t') & E(x_t x_t') \end{bmatrix} = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{bmatrix}$$

$$E(\varphi_t \varphi_t') = K' \Sigma_{YY} K = I$$

$$E(\xi_t \xi_t') = A' \Sigma_{XX} A = I$$

and

$$E(\varphi_t \xi_t') = R = \begin{bmatrix} r_1 & & 0 \\ & r_2 & \\ & & \\ 0 & & r_n \end{bmatrix}$$

where the elements of φ_t and ξ_t are ordered in such a way that $1 \geq r_1 \geq r_2 ...... \geq r_n \geq 0$

Definition

The population parameter r_i is known as the i^{th} population canonical correlation between y_t and x_t .

Lemma

The canonical correlations can be calculated by calculating the eigen values of

$$\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY},$$

$\lambda_1 > \lambda_2 > ...\lambda_n$ *, and the canonical correlations in turn are the square roots of these eigenvalues.*

The Johansen Method of Reduced Rank Regression

A way to assess the rank of If $\Phi(1)$ is to use the auxiliary regressions which allow to find an equivalent method to maximizing the likelihood of the ECM model. This equivalent expression allows to test how many cointegrating relationships do exist.

1. Rewrite the ECM Equation

$$\Delta X_t + BA' X_{t-1} \quad = c + \Gamma_1 \Delta X_{t-1} + ... + \Gamma_{p-1} \Delta X_{t-(p-1)} + u_t$$

2. It is possible to correct for the short run dynamics by regressing ΔX_t and X_{t-1} separately on the right hand side of the previous equation

$$\begin{aligned}
\Delta X_t &= c_1 + P_1 \Delta X_{t-1} + ... + P_{p-1} \Delta X_{t-(p-1)} + R_{0t} \\
X_{t-1} &= c_2 + T_1 \Delta X_{t-1} + ... + T_{p-1} \Delta X_{t-(p-1)} + R_{kt}
\end{aligned}$$

3. The latter can be used to form the residual (product moment) matrices
$$\widehat{S}_{ij} = T^{-1} \sum \widehat{R}_{it} \widehat{R}'_{jt} \quad i, j = 0, k$$
4. The maximum likelihood estimate of A is obtained as eigenvectors corresponding to the largest eigenvalues of:
$$|\lambda \widehat{S}_{kk} - \widehat{S}_{0k} \widehat{S}_{00}^{-1} \widehat{S}_{0k}| = 0$$
5. This gives n eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > .....\hat{\lambda}_n$ and their corresponding eigen vectors, $\hat{\varphi} = (\hat{\varphi}_1 > \hat{\varphi}_2 > ... > \hat{\varphi}_n)$

Testing for reduced rank

- The use of the auxiliary regressions allows to write the likelihood as a function of the eigen values. The likelihood that corresponds to the ECM without any restriction is

$$L^* = \begin{array}{c} -(Tn/2)log(2\pi) - (Tn/2) - (T/2)log\left|\widehat{S}_{00}\right| \\ -(T/2)\sum_{i=1}^{n}\log(1-\hat{\lambda}_i) \end{array}.$$

- To find the number of cointegrating vectors is equivalent to find the number of linearly independent columns in $\Phi(1)$ or the number of $n-r$ columns with small associated eigen values.
- Thus to test the null hypothesis that there are at most r cointegration vectors amounts to test

$$H_0)\lambda_i = 0 \quad i = r+1, ..., n.$$

where only the first r eigenvalues are non-zero.

- Likelihood that corresponds to the ECM without any restriction:

$$L^* = \begin{array}{c} -(Tn/2)log(2\pi) - (Tn/2) - (T/2)log\left|\widehat{S}_{00}\right| \\ -(T/2)\sum_{i=1}^{n}\log(1-\hat{\lambda}_i) \end{array}.$$

- Likelihood that corresponds to the ECM under the null that there are only r cointegrating vectors is:

$$L^*(H_0) = \begin{array}{c} -(Tn/2)log(2\pi) - (Tn/2) - (T/2)log\left|\widehat{S}_{00}\right| \\ -(T/2)\sum_{i=1}^{r}\log(1-\hat{\lambda}_i) \end{array}.$$

Trace statistic

- Thus to test the null hypothesis that there are at most r cointegration vectors amounts to test

$$H_0)\lambda_i = 0 \quad i = r+1, ..., n.$$

where only the first r eigenvalues are non-zero. (This is tested against the alternative of n cointegrating vectors).

- A likelihood ratio test, *using a non standard distribution*, can be constructed, using what is known as the **Trace statistic.**

$$\lambda_{trace} = -T\sum_{i=r+1}^{n} \log(1-\hat{\lambda}_i) \qquad r = 0, 1, ..., n-2, n-1$$

Max statistic

Another test of the significance of the largest λ_r is the so called maximal-eigenvalue or $\boldsymbol{\lambda}-$ **max statistic** :

$$\lambda_{\max} = -T\log(1-\hat{\lambda}_{r+1}) \qquad r = 0, 1, ..., n-2, n-1.$$

This tests the existence of r cointegrating vectors against the alternative that $r+1$ exist and is derived in exactly same way

Testing restrictions on the cointegrating vector

- Many times we are interested to test restrictions on the cointegrating vector.
- Therefore the asymptotic theory is standard since the regressions only involve variables which are I(0)and the test would be distributed *chi-square*. The LR test will be:

$$LR = -T \sum_{i=1}^{r} \log(1 - \hat{\lambda}_i) + T \sum_{i=1}^{r} \log(1 - \tilde{\lambda}_i)$$