Capítulo 24. COINTEGRACIÓN Y MODELOS VAR

Capítulo 24. COINTEGRACIÓN Y MODELOS VAR

Este capítulo presenta herramientas que posibilitan el estudio del corto y el largo plazo en economía, a través de elegir la herramienta econométrica adecuada. Tanto para modelos uniecuacionales como multiecuacionales, el nivel de integración de las series y el resultado de la combinación lineal entre ellas pueden dar lugar a los estudios de cointegración.

24.1 Introducción

La relación uniecuacional cointegrada y la expresión multiecuacional de vectores autorregresivos (VAR) expresan las relaciones de largo plazo entre las variables. Los modelos de corrección de error (MEC) y los vectores de corrección de error (VEC) muestran las relaciones de corto plazo que informan acerca de los desequilibrios en el fenómeno bajo estudio.

24.2 Cointegración

En general, si Y_t es $\mathrm{I(d)}$ y X_t es $\mathrm{I(d)}$ donde d es el mismo valor, las series Y_t y X_t pueden estar cointegradas. Si esto es así, el modelo

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \tag{1}$$

No es regresión espuria y no se pierde información valiosa en el largo plazo, lo cual sí sucedería si se toman primeras diferencias.

A (1) se lo conoce como regresión cointegrada y β_1 es el parámetro cointegrado o cointegrante que explicita la relación en el largo plazo de las variables.

Por ejemplo, el ahorro puede definirse como $I - C$ si el ahorro (S) es $\mathrm{I(0)}$ significa que es estacionario. En este contexto, una regresión donde

$$C_t = \beta_1 + \beta_2 I_t + \varepsilon_t \tag{2}$$

Sería cointegrada porque

$$\varepsilon_t = C_t - \beta_1 - \beta_2 I_t \tag{3}$$

Es estacionario y representa el ahorro (variable proxi del ahorro) y es la condición de la relación de largo plazo entre Consumo e Inversión.

La expresión (2) es la regresión cointegrante, β_2 es el parámetro cointegrante que muestra la relación de lago plazo.

¿cómo se trabaja en la práctica?

Primero, se hace la regresión

$$Y_t = \beta_1 + \beta_2 X_{2t} + \varepsilon_t \tag{4}$$

Y_t y X_{2t} son no estacionarios. Al estimar (4) se obtienen los errores e_t

Segundo, se analiza la estacionariedad de los errores de (4) haciendo

$$\Delta e_t = \rho \Delta e_{t-1} \tag{5}$$

Y se plantea la H_0: $raíz\ unitaria \rightarrow\ \rho = 1$

Aceptar la hipótesis nula es sinónimo de No Estacionariedad; mientras que, rechazar la nula significa que los errores son Estacionarios $[\mathrm{I}(0)]$. Se pueden realizar las pruebas de Dickey Fuller (DF), DF Aumentado (DFA), Phillip Perron (PP) u otra.

Tercero, si los residuos son estacionarios significa que β_2 mide la relación en el largo plazo y que podemos, a partir de los errores, encontrar la relación en el corto plazo. Si (4) está cointegrada, entonces

$$e_t = \widehat{Y_t} - \widehat{\beta_1} - \widehat{\beta_2} X_{2t} \tag{6}$$

Es estacionario y puede considerarse el error de equilibrio y utilizarse para relacionar el comportamiento de corto plazo con el largo plazo.

Haciendo

$$\Delta Y_t = \alpha_1 + \alpha_2 \Delta X_{2t} + \alpha_3 e_{t-1} + \varepsilon_t \tag{7}$$

Donde $\Delta Y_t = Y_t - Y_{t-1}$, $\Delta X_t = X_t - X_{t-1}$, $e_{t-1} = Y_{t-1} - \widehat{\beta_1} - \widehat{\beta_2} X_{2t-1}$

En (7), si

$e_{t-1} = 0$ modelo está en equilibrio

$e_{t-1} \neq 0$ modelo no está en equilibrio

Si $\left.\begin{matrix} \Delta X = 0 \\ e_{t-1} > 0 \end{matrix}\right\} Y_{t-1}$ está por encima del valor de equilibrio y se espera que $\alpha_3 < 0$ para restaurar el equilibrio.

Si ΔY está por encima del valor de equilibrio empezará a disminuir en el siguiente periodo para corregir el error de equilibrio.

Dicho de otro modo:

$\alpha_3 < 0 \rightarrow Y_t$ está por encima de su valor de equilibrio, tiene que bajar

$\alpha_3 > 0 \rightarrow Y_t$ está por debajo de su valor de equilibrio, tiene que subir

La expresión (7) se lo conoce como mecanismo de corrección de error (MEC) donde α_2 muestra la relación en el corto plazo y si $\alpha_3 = 0$ a la prueba t, significa que el modelo está en equilibrio de corto y largo plazo.

En la literatura van a encontrar como teorema de representación de Granger: si dos variables están cointegradas, la relación entre las dos se expresa como mecanismo de corrección de error (MEC).

La metodología de trabajo se lo conoce como metodología de Granger y Engle o procedimiento bietápico de Granger y Engle.

24.3 Modelos VAR

Otero (1993, p.318) plantea que la modelización en econometría es un procos complejo que comprende desde el plante formal del problema de interés a la validación de los resultados, pasando por la realización de inferencias estadísticas con datos reales.

Algunos de los supuestos sobre los que descansa la metodología de la econometría tradicional son:

a) Se conoce el orden de causalidad de las variables que entran en estudio
b) Se conoce qué variables hay que omitir en cada ecuación
c) Se ignora la no estacionariedad de las series
d) Los parámetros estimados se suponen constantes
e) El modelo se verifica frente a la realidad peo no se puede verificar frente a otros modelos alternativos

Ignorar estos supuestos da lugar a errores de predicción y en la práctica muchas veces es lo que sucede.

Los dos primeros supuestos son los que presentan las mayores limitaciones. En un sistema multiecuacional, significa especificar de manera correcta el modelo de forma que todos los parámetros estructurales pueden ser estimados; es decir, que cada ecuación esté identificada. Pero hay situaciones en las que, para convertir en identificable una ecuación se agregan variables. Este es precisamente uno de los aspectos que trata de evitar el método de Sims con la modelización VAR (vectores autprregresivos). El origen de estos modelos data de un artículo publicado por Sims en 1980

(Macroeconometría y realidad). Sims afirma que la mayoría de las restricciones que aparecen en los modelos son falsas y postula que:

f) No hay conocimientos teóricos suficientes para clasificar a las variables en endógenas y exógenas
g) No es posible establecer restricciones cero (conocer los coeficientes delas variables omitidas en la ecuación

El objetivo de la propuesta es proporcionar una estrategia de modelización que permita reflejar lo más fielmente posible las regularidades empíricas e iteracciones entre las variables objeto de análisis.

El VAR presenta un sistema de ecuaciones simultáneas en el que cada una de las variables son explicadas por sus propios rezagos y los del resto de variables del sistema. En general, todas las variables son consideradas endógenas.

Los críticos denominan a este enfoque macroeconometría ateórica.

24.3.1 Especificación de un modelo VAR

El planteo original de un modelo VAR parte de la especificación general de un modelo multiecuacional sobre la forma reducida

$$Y_t = X_t \Pi + v_t$$

A partir del cual se construye un nuevo vector Z_t que agrupa todas las variables, tanto endógenas como exógenas

$$Z_t = (Y_t \, X_t)$$

Considerando a Z_t una variable aleatoria multivariante se puede utilizar el Teorema de la descomposición de Wald por el cual toda variable aleatoria se descompone en una parte determinista y otra puramente aleatoria

$$Z_t = D_t + \underbrace{c(\mathrm{L})\varepsilon_t}_{\Delta_t}$$

$\Delta_t = c(\mathrm{L})\varepsilon_t$ es una matriz de retardos aplicados sobre una matriz de componentes aleatorios multivariantes

Descomponiendo a Δ_t en los componentes

a: que recoge los retardos aplicados a las variables

b: que recoge los retardos aplicados a las componentes aleatorias

$$\Delta_t = a^{-1}(L)^* \, b(L)^* \, \varepsilon_t$$

Premultiplicando por $a(L)$

$$a(L)Z_t = b(L)^* \, \varepsilon_t$$

Si se supone $b(L)^*$ constante y se separan los componentes contemporáneos de a, se obtiene una representación autorregresiva pura del modelo multiecuacional

$$A_t = a_0^{-1}(a_1 Z_{t-1} + a_2 Z_{t-2} + \cdots + a_r Z_{t-r} + B \, \varepsilon_t)$$

Esta forma autorregresiva se puede estimar por mínimos cuadrados ordinarios.

Para la especificación del modelo VAR hay que definir el número total de variables que componen el sistema (m) y el número máximo de retardos incluir (r) incluyendo un matriz con términos deterministas D.

De modo que, para cada variable Y_t se plantea una ecuación del tipo

$$Y_{i,t} = D_{ij} + \sum_{i=1}^{m}\sum_{j=1}^{r} \beta_{ij} Y_{it-j} + \varepsilon_{it}$$

Es decir, para cada variable endógena

$$\begin{aligned} Y_{i,t} = D_{ij} + \beta_{11}Y_{1t-1} + \beta_{12}Y_{1t-2} + \cdots + \beta_{1r}Y_{1t-r} + \beta_{21}Y_{2t-1} + \beta_{22}Y_{1t-2} + \cdots \\ + \beta_{2r}Y_{2t-r} + \cdots . + \beta_{m1}Y_{mt-1} + \beta_{m2}Y_{mt-2} + \cdots + \beta_{mr}Y_{mt-r} + \varepsilon_{it} \end{aligned}$$

Las perturbaciones aleatorias de cada ecuación ε_{it} deben comportarse como ruido blanco, por lo tanto las matrices de varianzas y covarianzas son escalares

$$cov(\varepsilon_{it}) = E(\boldsymbol{\varepsilon\varepsilon}') = \sigma^2 \boldsymbol{I}$$

Además, las perturbaciones aleatorias de distintas ecuaciones deben presentar matrices de varianzas y covarianzas constantes paa cada punto muestral

$$cov(\varepsilon_{it}\varepsilon_{jt}) = E(\boldsymbol{\varepsilon_i \varepsilon_j'}) = \Sigma \quad \forall t$$

Donde $\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1m}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots & \sigma_{2m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}^2 & \sigma_{m2}^2 & \cdots & \sigma_{mm}^2 \end{bmatrix}$

24.3.2 Identificación

El procedimiento de identificación en los modelos multiecuacionales informa si todos los parámetros pueden ser estimados. En los VAR, cuando la finalidad es estimación o predicción, no es necesaria la identificación; pero sí lo es si a partir de la estimación van a realizarse simulaciones.

A diferencia de los modelos multiecuacionales estructurales, la identificación se realiza sobre la matriz de varianzas y covarianzas de las perturbaciones porque la simulación se hace a partir de las perturbaciones y se denomina Análisis de Impulso respuesta. Estos consisten en introducir una alteración en la perturbación aleatoria de una ecuación y comprobar el resultado que esta alteración tiene sobre el conjunto del sistema.

La función impulso respuesta es la representación de medias móvilaes asociadas con el modelo estimado y explica las respuestas del sistema a shocks en los componentes del vector de perturbaciones.

La función impulso respuesta traza la respuesta de las variables endógenas en el sistema entre un shock en los errores. El shock puede ser un desvío.

El proceso de identificación en un modelo VAR consiste en ortogonalizar las perturbaciones aleatorias transformando la matriz de varianzas y covarianzas de las perturbaciones en una matriz diagonal.

Generalmente, la ortogonalización de los errores se realiza mediante la descomposición de Choleski, la cual es extensamente usada. La descomposición de la varianza de un VAR se obtiene luego de ortogonalizar el vector de perturbaciones. La descomposición consiste en distribuir la disponibilidad de las correlaciones reflejadas en la matriz de varainzas y covarianzas entre los distintos componentes del vector de perturbaciones.

La descomposición de la varianza de un VAR brinda información acerca de la potencia relativa de innovaciones aleatorias para cada variable endógena.

Loría (…… p.271) muestra que un VAR se puede completar co variables como intercepto, tendencia determinística (@trend) y dummies tracidionales o estacionales.

Un VAR puede representarse como

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Si existe correlación intertemporal de los errores

$$E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0$$

$$E(\varepsilon_{1t}^2) = \sigma_{11}$$

$$E(\varepsilon_{2t}^2) = \sigma_{22}$$

$$E(\varepsilon_{1t}\varepsilon_{2t}) = \sigma_{12}$$

Al ser $\sigma_{12} \neq 0$ se tiene autocorrelación de errores, para eliminarla se debe ortogonalizar la matriz de varianzas y covarianzas. La idea central de la ortogonalización del error consiste en hacer independiente los errores entre ecuaciones y de esta manera usar las ecuaciones por separado para análisis de política.

24.3.3 Estimación

El modelo se estima por mínimos cuadrados ordinarios porque se incluyen variables retardadas como explicativas que no tienen correlación temporal con el término de error

$$C\left(Y_{t-j}\varepsilon_t\right) = 0$$

$$\hat{\beta} = \left(Y'_{t-j}Y_{t-j}\right)^{-1}\left(Y'_{t-j}Y_t\right)$$

Y_t incluye los términos deterministas D_t. El problema que se presenta es definir cuantos retardos tiene el modelo.

El Ratio de verosimilitud propuesto por SIMS se define como

$$(N-k)[log|\Sigma_r| - log|\Sigma_a|] \sim \chi_R^2$$

Donde

N es el total de observaciones

k variables del modelo ampliado (todos los retardos posibles)

Σ es la matriz de productos cruzados de residuos R en el modelo restringido (menor al número de retardos) y a en el modelo ampliado

R es el total de restricciones igual al número de parámetros a estimar

Los criterios de Akaike y Schwarz se definen como

$$\mathrm{AIC} = -2\frac{l}{N} + 2\frac{m}{N}$$

$$\mathrm{SC} = -2\frac{l}{N} + m\frac{\log{(N)}}{N}$$

Donde $\mathrm{l} = -\frac{Nk}{2}(1 + 2log2\Pi) - N\frac{N}{2}log|\Sigma|$, siendo m el número de variables del modelo.

El modelo que presente un menor valor de estos estadísticos es el seleccionado.

24.4 Alternativas al modelo VAR.

El modelo de Sims puede denominarse VAR irrestricto. Blanchard (1989) propone los modelos VAR estructurales o SVAR

$$\mathrm{Y}_t = \mathrm{a}_{10} + \mathrm{a}_{11}\mathrm{Y}_{t-1} + \mathrm{a}_{12}\mathrm{Y}_{t-2} + \mathrm{a}_{13}\mathrm{X}_{t-1} + \mathrm{a}_{14}\mathrm{X}_{t-2} + \varepsilon_{1t}$$
$$\mathrm{X}_t = \mathrm{a}_{20} + \mathrm{a}_{21}\mathrm{Y}_{t-1} + \mathrm{a}_{22}\mathrm{Y}_{t-2} + \mathrm{a}_{23}\mathrm{X}_{t-1} + \mathrm{a}_{24}\mathrm{X}_{t-2} + \varepsilon_{2t}$$

Modelos VAR parciales o PVAR que dan solución a la falta de significatividad de algunas variables

$$\mathrm{Y}_t = \mathrm{a}_{10} + \mathrm{a}_{11}\mathrm{Y}_{t-1} + \mathrm{a}_{12}\mathrm{Y}_{t-2} + \varepsilon_{1t}$$
$$\mathrm{X}_t = \mathrm{a}_{20} + \mathrm{a}_{21}\mathrm{Y}_{t-1} + \mathrm{a}_{22}\mathrm{Y}_{t-2} + \mathrm{a}_{23}\mathrm{X}_{t-1} + \mathrm{a}_{24}\mathrm{X}_{t-2} + \varepsilon_{2t}$$

Modelos vectores de corrección de error VEC

Se originan a partir del concepto de cointegración generalizado en los modelos VAR

$$\Delta\mathrm{Y}_t = \mathrm{a}_{10} + \mathrm{a}_{Y}(\mathrm{Y}_{t-1} - \beta\mathrm{X}_{t-1}) + \mathrm{a}_{11}\Delta\mathrm{Y}_{t-1} + \mathrm{a}_{12}\Delta\mathrm{Y}_{t-2} + \mathrm{a}_{13}\Delta\mathrm{X}_{t-1} + \mathrm{a}_{14}\Delta\mathrm{X}_{t-2} + \varepsilon_{1t}$$
$$\Delta\mathrm{X}_t = \mathrm{a}_{20} + \mathrm{a}_{X}(\mathrm{Y}_{t-1} - \beta\mathrm{X}_{t-1}) + \mathrm{a}_{21}\Delta\mathrm{Y}_{t-1} + \mathrm{a}_{22}\Delta\mathrm{Y}_{t-2} + \mathrm{a}_{23}\Delta\mathrm{X}_{t-1} + \mathrm{a}_{24}\Delta\mathrm{X}_{t-2} + \varepsilon_{2t}$$

El resultado del modelo VEC informa sobre la relación de largo plazo de las variables y sobre los desvíos de corto plazo.

El largo plazo es la relación de las variables en el momento t

El corto plazo incorpora el término de error (coint) y las variables en diferencia.

Para cointegrar, el VAR tiene que tener variables integradas de algún orden. Los residuos que se generan a partir de aquí (para cada ecuación) tienen que ser estacionarios.

Modelos VAR con medias móviles en las perturbaciones VARMA

$$\begin{aligned}
Y_t &= a_{10} + a_{11}Y_{t-1} + a_{12}Y_{t-2} + a_{13}X_{t-1} + a_{14}X_{t-2} + \theta_{11}\varepsilon_{t-1} + \theta_{12}\varepsilon_{1t-2} \\
X_t &= a_{20} + a_{21}Y_{t-1} + a_{22}Y_{t-2} + a_{23}X_{t-1} + a_{24}X_{t-2} + \theta_{21}\varepsilon_{t-1} + \theta_{22}\varepsilon_{1t-2}
\end{aligned}$$

24.5 Test de Johansen para la cointegración

El procedimiento de Johansen (1991) es una generalización multivariada de Dickey Fuller. Parte de la representación autorregresiva general de una matriz multivariante de m variables que vendría dada por

$$Y_t = \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \cdots \Pi_p Y_{t-p} + \varepsilon_t$$

Restando en ambos la matriz Y_{t-1}

$$Y_t - Y_{t-1} = \Pi_1 Y_{t-1} - Y_{t-1} + \Pi_2 Y_{t-2} + \cdots \Pi_p Y_{t-p} + \varepsilon_t$$

$$\Delta Y_t = (\Pi_1 - 1)Y_{t-1} + \Pi_2 Y_{t-2} + \cdots \Pi_p Y_{t-p} + \varepsilon_t$$

Sumando y restando $(\Pi_1 - 1)Y_{t-2}$ a la derecha de la igualdad

$$\Delta Y_t = (\Pi_1 - 1)Y_{t-1} + \Pi_2 Y_{t-2} + (\Pi_1 - 1)Y_{t-2} - (\Pi_1 - 1)Y_{t-2} + \cdots \Pi_p Y_{t-p} + \varepsilon_t$$

Queda

$$\Delta Y_t = (\Pi_1 - 1)(Y_{t-1} - Y_{t-2}) + (\Pi_2 + \Pi_1 - 1)Y_{t-2} + \cdots \Pi_p Y_{t-p} + \varepsilon_t$$

$$\Delta Y_t = (\Pi_1 - 1)\Delta Y_{t-1} + (\Pi_2 + \Pi_1 - 1)Y_{t-2} + \cdots \Pi_p Y_{t-p} + \varepsilon_t$$

Repitiendo el proceso p veces y definiendo las matrices

$$\Omega_i = -\left(1 - \sum_{j=1}^{i} \Pi_i\right)$$

Se puede escribir

$$\Delta Y_t = \Omega_1 \Delta Y_{t-1} + \Omega_2 Y_{t-2} + \cdots \Omega_p Y_{t-p} + \varepsilon_t$$

Donde la matriz Ω contiene los parámetros que definen las relaciones de equilibrio entre las variables incluidas en el sistema. El rango de dicha matriz define el número de retardos de cointegración diferentes que existen entre las m variables

Si $\rho(\Omega) = 0$ todos los vectores son linealmente dependientes y significa que no hay relación de cointegración entre las variables. Mientras que, si $\rho(\Omega) = r < m$, hay r vectores linealmente independientes y significa que hay *r* relaciones de cointegración diferentes.

Si existen las relaciones de cointegración, entonces se pueden descomponer la matriz Ω en el producto de dos submatrices λ y β de forma que

$$\Omega_{mxr} = \lambda\beta'$$

Donde β' contiene los coeficientes de los vectores de cointegración y λ los parámetros de la velocidad de ajuste.

Sustituyendo la matriz Ω en la expresión general del modelo

$$\Delta Y_t = \Omega_1 \Delta Y_{t-1} + \Omega_2 Y_{t-2} + \cdots \lambda\beta' Y_{t-p} + \varepsilon_t$$

Donde el producto $\beta' Y_{t-p}$ es estacionario

En la práctica, el contraste de Johansen consiste en:

1. Determinar el orden óptimo del modelo VAR
2. Estimar el modelo con procedimientos de máximaverosimilitud con restricciones paramétricas
3. Determinar el rango de cointegración en un entorno probabilístico.

Para esto último se definen dos estadísticos basados en los autovalores de la matriz $\hat{\Omega}$ y formulados como:

$$\mathrm{V}_{traza}(r) = -N \sum_{i=r+1}^{m} \ln\left(1 - \hat{v_i}\right)$$

$$\mathrm{V}_{max}(r, r+1) = -N\ln\left(1 - \widehat{v_{r+1}}\right)$$

Donde v son los autovalores, r el número de relaciones de cointegración y N el número total de observaciones

Con V_{traza} se contrasta

$$H_0: n^{\circ}\ vectores\ de\ cointegración\ \leq r$$
$$H_1: n^{\circ}\ vectores\ de\ cointegración\ > r$$

Con V_{max} se contrasta

$$H_0: existencia\ de\ r\ vectores\ de\ cointegración$$
$$H_1: existencia\ de\ r+1\ vectores\ de\ cointegración$$

Los valores de estos estadísticos se comparan con valores críticos estimados por Johansen y Juselius.

Según Perez Lopez (…) en la práctica son más estables las ealciones de cointegración obtenida en el contexto univariante que en el contexto multivariante y hoy en día se usan más debido a la menor fiabilidad de las relaciones derivadas del test de Johansen.

CASOS DE ESTUDIO, PREGUNTAS Y PROBLEMAS

Caso 24.1:

Caso 24.2:

Bibliografía

- IHS Global Inc (2019). “Eviews 11, Users Guide II”. IHS Markit, Irvine (USA).
- Loría, E. (2007). “Econometría con Aplicaciones”. Editorial Pearson Prentice Hall. México.
- Otero, J. (1993). Econometría Series Temporales y Predicción. Editorial AC, Madrid.
- Perez, C. (2006) Econometría de Series de Tiempo. Pearson Prentice Hall. Madrid.
- Perez, C. (2008). Econometría Avanzada. Técnicas y Herramientas”. Pearson Prentice Hall