

VECTORES AUTORREGRESIVOS.

Rafael Bustamante Romani



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS ECONÓMICAS

La **Serie Apuntes de Clase Omega Beta Gamma** tiene por objetivo difundir los materiales de enseñanza generados por los docentes que tienen a su cargo el desarrollo de las asignaturas que forman parte de los Planes de Estudios de las Escuelas Académico-Profesionales de la Facultad de Ciencias Económicas de la Universidad Nacional Mayor de San Marcos. Estos documentos buscan proporcionar a los estudiantes la explicación de algunos temas específicos que son abordados en su formación universitaria.

Encargados de la serie:

Bustamante Romani, Rafael.
rbustamanter@unmsm.edu.pe

Cisneros García, Juan Manuel.
jcisnerosg@unmsm.edu.pe

Facultad de Ciencias Económicas.
Universidad Nacional Mayor de San Marcos.
Calle Germán Amézaga N° 375.
Ciudad Universitaria, Lima 1. Perú.

La **Serie Apuntes de Clase ΩBT** es promovida y desarrollada por un colectivo de docentes del Departamento de Economía de la Universidad Nacional Mayor de San Marcos.

El contenido de cada publicación es íntegramente responsabilidad de cada autor, no representa necesariamente los puntos de vista de los integrantes del colectivo, ni de la Universidad.



Vectores Autorregresivos

Rafael Bustamante Romani[◇]

Resumen

Este documento describe la estimación y análisis de los vectores Autorregresivos (VAR) y los modelos de errores de corrección (VEC). Se consideran a los VAR como una forma reducida que pudo haberse derivado de algún modelo estructural. Esto es, un VAR es una herramienta de análisis econométrico que permite a los datos hablar por ellos mismos, sin que exista necesariamente una teoría económica que guíe o restrinja la estructura de un modelo

Palabras claves: VAR, MA, descomposición de Cholesky, causalidad, test de Sims, exogeneidad.

Clasificación JEL: C32, C40.

[◇] Estudiante del Doctorado en Economía con mención en los Recursos Naturales, Universidad Nacional Autónoma de México. MBA Gerencial, CENTRUM Pontificia Universidad Católica del Perú. Maestría en Economía con mención en Finanzas, Universidad Nacional Mayor de San Marcos. B. Sc. Economía, UNMSM. Profesor del Departamento de Economía de la UNMSM. Investigador asociado al Instituto de Investigaciones FCE - UNMSM. Contacto: rbustamanter@unmsm.edu.pe

I. INTRODUCCIÓN

En los modelos econométricos estructurales (tradicionales), que hacen uso de información en forma de series de tiempo, comúnmente se requiere imponer restricciones a los parámetros involucrados para obtener formas reducidas que puedan ser estimadas con las técnicas estadísticas conocidas; también resulta necesario hacer supuestos acerca de la dinámica del sistema económico, mediante la imposición de restricciones sobre el número de retrasos con que una variable afecta a las demás. Es requisito asimismo, conocer cuáles de las variables involucradas son exógenas y cuáles son endógenas; por otro lado, existe el problema en algunos modelos de que se requiere tener en cuenta las expectativas del comportamiento de algunas variables (lo que ha dado origen en particular a los modelos de expectativas racionales). Este tipo de restricciones han sido subrayadas en especial por Sims (1980) y por Hendry y Richard (1983), entre otros autores de literatura econométrica (Beltrán Barco, 2003).

El enfoque estructural de la modelación de las ecuaciones simultáneas usa la teoría económica para describir las relaciones entre varias variables de interés. El modelo resultante se estima entonces, y probaba la relevancia empírica de la teoría. Desgraciadamente, la teoría económica no es a menudo lo bastante rica para proporcionar una especificación completa de todas las relaciones dinámicas entre las variables. Además, la estimación e inferencia son complicadas por el hecho que las variables endógenas pueden aparecer en la izquierda y derecha de las ecuaciones, recuérdese las estimaciones de ecuaciones simultáneas.

No obstante la arbitrariedad de las restricciones impuestas a priori, ya sea por teoría económica o por necesidades de cómputo, los modelos estructurales han probado ser útiles en la práctica para obtener pronósticos y para realizar análisis de política económica. Este hecho conduce a pensar entonces que son las **formas reducidas** las que realmente importan en la práctica, aun cuando se hayan obtenido con restricciones derivadas de supuestos falsos. Es por este motivo que es

conveniente tener representaciones en forma reducida, aunque no se tenga el modelo estructural completo, y esto es precisamente lo que se logra con un Vector Autorregresivo (VAR): una forma reducida que pudo haberse derivado de algún modelo estructural. Esto es, un VAR es una herramienta de análisis econométrico que permite a los datos hablar por ellos mismos, sin que exista necesariamente una teoría económica que guíe o restrinja la estructura de un modelo. (Beltrán Barco, 2003).

Este documento describe la estimación y análisis de los Vectores autorregresivos (VAR) y los modelos de errores de corrección (VEC).

Los modelos VAR tradicionales son, en cierta forma, una respuesta a la imposición de restricciones a priori que caracteriza a los modelos econométricos keynesianos: En un sistema de ecuaciones simultáneas se requiere imponer restricciones sobre los parámetros de las mismas para garantizar la identificación, y posible estimación, de las ecuaciones que lo conforman. Para ello, además, es indispensable diferenciar entre las variables endógenas y las predeterminadas, es decir, aquellas cuyos valores no son determinados por el modelo en el período actual. Estas últimas pueden ser exógenas o endógenas rezagadas. En cambio en el sistema VAR nos proporciona estimaciones (o predicciones o simulaciones) para las variables endógenas donde el sistema está relacionado a lo largo del tiempo. Igualmente se usa (no sin controversia), para analizar el impacto dinámico de diferentes tipos de variación estocástica, es decir shocks económicos, en las variables del lado derecho de una ecuación.

Finalmente en este enfoque se enfatiza la necesidad de tratar que cada variable endógena dentro del sistema sea función de las variables rezagadas de todas las variables endógenas dentro del sistema. En resumen el vector autoregresivo es

comúnmente usado para pronósticos de sistemas de serie de tiempo interrelacionados y analizar el impacto dinámico de una perturbación aleatoria sobre las demás variables del sistema.

II. ESPECIFICACIÓN

Para especificar un VAR, se requieren las siguientes decisiones:

- La lista de variables endógenas
- El intervalo de retardo del modelo
- La lista de variables exógenas, si fuera necesario

El VAR presenta alternativamente, un sistema de ecuaciones simultáneas en el que cada una de las variables es explicada por sus propios rezagos y los del resto de variables del sistema. Es decir, **no se admiten restricciones a priori** y todas las variables son consideradas endógenas. La única información a priori que se incluye está referida al número de rezagos de las variables explicativas, que se incorporan en cada ecuación a partir del análisis de la data. No obstante, en términos operativos, una correcta especificación del sistema requiere que la determinación de las variables a ser incluidas en él se base en el *conocimiento de un modelo teórico* relevante (Beltrán Barco, 2003).

Supuestos en la estimación de un VAR (Beltrán Barco, 2003):

- Las variables que componen el vector son estacionarios (salvo para casos de cointegración) en cuyo caso existen metodologías alternativas.
- Esto permite que los Test hechos sobre VAR tengan las distribuciones estándar necesarias en la etapa de inferencia.
- Inclusión de variables no estacionarias sujetas a los mismos problemas que el caso univariado: distribuciones no estándar (salvo el caso de cointegración).

Un VAR tiene, en general, la siguiente especificación:

$$y_t = \sum_{i=1}^p \Pi_i y_{t-i} + e_t \quad 1.$$

Donde y_t e y_{t-i} son vectores de orden m (m es el número de variables del sistema) y Π_i es la matriz (cuadrada de orden m) de coeficientes del rezago i de las variables explicativas de las m ecuaciones. De esta forma, se puede observar que deberán estimarse tantas matrices Π_i como rezagos se incluyan en el sistema. Matricialmente, y utilizando una especificación de operadores de rezago **Fuente especificada no válida.:**

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{mt} \end{bmatrix} = \begin{bmatrix} a_{11(L)} & a_{12(L)} & \cdots & a_{1m(L)} \\ a_{21(L)} & \ddots & & a_{2m(L)} \\ \vdots & & & \vdots \\ a_{m1(L)} & & \cdots & a_{mm(L)} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{mt} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{mt} \end{bmatrix} \quad 2.$$

En este sistema existen m variables endógenas y el operador de rezagos L resume los p rezagos del modelo.

Esta es la formulación del VAR en la forma reducida. En este sistema:

$$E \left[e_t e_{t-j}^T \right] = 0 \quad \forall j \neq 0 \quad 3.$$

$$E \left[e_t e_t^T \right] = \Sigma_e \quad \text{Es una matriz simétrica.}$$

Entonces μ tendrá una distribución normal multivariada $N_k(0, \Sigma)$, donde Σ es la matriz de varianza-covarianza:

$$E \left[\begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{mt} \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{mt} \end{bmatrix}^T \right] = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \cdot & \cdot & \cdot & \sigma_{1,m} \\ \sigma_{2,1} & \sigma_2^2 & \cdot & \cdot & \cdot & \sigma_{2,m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \sigma_{m-2,m} \\ \sigma_{m-1,1} & \sigma_{m-1,2} & \cdot & \cdot & \sigma_{m-1}^2 & \sigma_{m-1,m} \\ \sigma_{m,1}^2 & \sigma_{m,2} & \cdot & \cdot & \sigma_{m,m-1} & \sigma_m^2 \end{bmatrix} = \Sigma_e$$

Es decir, no se tiene autocorrelación entre los errores de una misma ecuación pero se observa correlación contemporánea entre los errores de las diferentes ecuaciones.

Características adicionales:

La estimación de esta forma reducida por MCO para cada ecuación del vector es eficiente.

La forma estructural no está identificada salvo supuestos adicionales. Desde que solo los valores rezagados de las variables endógenas aparecen en el lado derecho de las ecuaciones, la simultaneidad no es problema y los estimadores obtenidos por MCO son consistentes siempre y cuando se consideren a las variables rezagadas como determinísticas. Sin embargo aun cuando las innovaciones ε_t pueden estar contemporáneamente correlacionadas los estimadores MCO son eficientes y equivalentes a los MCG debido a que todas las ecuaciones tienen idénticos regresores.

Veamos, por ejemplo, el caso de un VAR (1) con dos variables, de la forma estructural siguiente:

$$y_t = \beta_{10} - \gamma_{12} z_t + \beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \varepsilon_{y_t}$$

4.

$$z_t = \beta_{20} - \gamma_{21} y_t + \beta_{21} y_{t-1} + \beta_{22} z_{t-1} + \varepsilon_{z_t}$$

Donde y_t y z_t son variables endógenas estacionarias, ε_{yt} y ε_{zt} son ruidos blancos que no están correlacionados entre sí. La ecuación 4 sería entonces la forma estructural del sistema. Aunque se tienen endógenas como explicativas. Si se quiere obtener la forma reducida, es decir, expresar las endógenas en función sólo de predeterminadas (rezagos de las endógenas), se debe resolver (Beltrán Barco, 2003):

$$\begin{bmatrix} 1 & \gamma_{12} \\ \gamma_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \quad 5.$$

La ecuación 5 sería entonces la forma estructural del sistema. Aunque se tienen endógenas como explicativas. Si se quiere obtener la forma reducida, es decir, expresar las endógenas en función sólo de predeterminadas (rezagos de las endógenas).

En este sistema si se da que un shock en la variable Y_t expresado en $\varepsilon_{yt} \uparrow$, o un golpe de una desviación estándar en Y_t entonces Z_t también se verá afectado. Además debemos decir lo siguiente:

$-\gamma_{12}$: Es la variable que mide el efecto contemporáneo de un cambio de z_t sobre y_t .

β_{21} : Él es efecto de un cambio en y_{t-1} sobre z_t .

Como se puede observar existe un problema de correlación de z_t con ε_{yt} y de y_t con ε_{zt}

El sistema se puede describir en términos vectoriales como:

$$\Gamma Y_t = B_0 + B_1 Y_{t-1} + \varepsilon_t$$

Siendo Y_t un vector que contiene a y_t y z_t

$$Y_t = \Gamma^{-1} B_0 + \Gamma^{-1} B_1 Y_{t-1} + \Gamma^{-1} \varepsilon_t \quad 6.$$

$$Y_t = A_0 + A_1 Y_{t-1} + e_t$$

Es decir

$$y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{1t} \quad 7.$$

$$z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{2t} \quad 8.$$

La ecuación 7 es la forma reducida del VAR (1) de la ecuación 4. En ella los errores sí están correlacionados debido a que recogen la presencia de y_t, y_{t-1}, z_t como explicativas del VAR original. Así:

$$\Gamma^{-1} = \begin{bmatrix} 1/\Delta & -\gamma_{12}/\Delta \\ -\gamma_{21}/\Delta & 1/\Delta \end{bmatrix} \quad 9.$$

Dónde: $\Delta = 1 - \gamma_{12} \gamma_{21}$
y,

$$\Gamma^{-1} \varepsilon_t = e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}_t \begin{bmatrix} \frac{\varepsilon_{yt} - \gamma_{12} \varepsilon_{zt}}{\Delta} \\ \frac{-\varepsilon_{yt} \gamma_{21} + \varepsilon_{zt}}{\Delta} \end{bmatrix} \quad 10.$$

Ya que:

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

Son procesos ruidos blanco se puede decir que e_{1t} y e_{2t} tienen una media igual a cero y varianza constante y además cada una de ellas es in correlacionada individualmente, en el tiempo con varianza constante pero si existe correlación entre ellas.

$$E(e_{yt} e_{zt}) = E \left[\frac{-\varepsilon_{yt}^2 \gamma_{21} + \varepsilon_{yt} \varepsilon_{zt} + \gamma_{21} \gamma_{12} \varepsilon_{yt} \varepsilon_{zt} - \gamma_{12} \varepsilon_{zt}^2}{\Delta^2} \right] \quad 11.$$

$$E(e_{yt}e_{zt}) = \frac{-\gamma_{21}\sigma_y^2 - \gamma_{12}\sigma_z^2}{\Delta^2} \neq 0 \tag{12}$$

La ecuación anterior es diferente de cero si $\gamma_{21} \neq 0$, $\gamma_{12} \neq 0$, es decir, mientras que y_t, z_t estén presentes en la forma estructural del VAR. Note que cuando se estima un VAR lo que en realidad se observa es la forma reducida.

III. IDENTIFICACIÓN DE UN VAR

El problema consiste en “rescatar” los parámetros de la forma estructural, partir de las estimaciones de la forma reducida.

Sin restricciones adicionales este proceso es imposible (existe una subidentificación). Veamos en un sistema VAR con n variables endógenas.

$$\Gamma Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + B_3 Y_{t-3} + \dots + B_p Y_{t-p} + \varepsilon_t$$

Que es la cantidad de parámetros que se necesitan conocer. Entonces a partir de la forma reducida tiene:

$$Y_t = \pi_0 + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_{t-3} Y_{t-3} + \dots + \pi_{t-p} Y_{t-p} + e_t$$

A partir del cual se puede decir lo siguiente:

$$\# \text{ Parámetros de la forma estructural} - \# \text{ estimaciones de la forma reducida} = \# \text{ restricciones necesarias para la identificación} = n(n-1)/2.$$

Note que en la forma reducida ni los parámetros ni los shocks tienen interpretación económica. Los errores de la forma reducida ya no son ortogonales entre sí, sino combinaciones lineales de innovaciones estructurales.

Los parámetros de la forma reducida pueden ser estimados por (MCO) y estos estimadores son suficientes y consistentes.

Por ejemplo si se tiene un VAR (1):

$$y_t = \beta_{10} - \gamma_{12} z_t + \beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \varepsilon_{yt} \quad 13.$$

$$z_t = \beta_{20} - \gamma_{21} y_t + \beta_{21} y_{t-1} + \beta_{22} z_{t-1} + \varepsilon_{zt} \quad 14.$$

Este VAR contiene 2 variables por lo tanto usando $n(n-1)/2=1$, es decir se necesita una restricción para identificar los coeficientes de la forma estructural.

En un célebre artículo donde Christopher Sims introdujo los VARs a la profesión econométrica (Macroeconomics and Reality, Econometrica 1980) propuso una estrategia de **estimación basada en una representación recursiva** “la teoría económica puede ayudarnos” del VAR (es decir la matriz Γ es triangular).

Él nos dice que no es necesario “volver” a la forma reducida. Otros autores imponen una restricción a la forma estructural y así obtener los parámetros de la forma estructural **Fuente especificada no válida.**

Podemos suponer que $\gamma_{21} = 0$.

$$\begin{bmatrix} 1 & \gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \quad 15.$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \quad 16.$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \quad 17.$$

Reemplazando tenemos:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \beta_{10} - \gamma_{12} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} - \gamma_{12}\beta_{21} & \beta_{12} - \gamma_{12}\beta_{22} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} - \gamma_{12}\varepsilon_{zt} \\ \varepsilon_{zt} \end{bmatrix} \quad 18.$$

Una forma reducida puede corresponder a varias formas estructurales cuando el sistema está sobre identificado.

Se puede estimar la forma reducida por MCO siempre y cuando el VAR sea "estable" (estacionaria en covarianza es decir débilmente estacionaria).

Donde:

$$1) a_{10} = \beta_{10} - \gamma_{12}$$

$$2) a_{20} = \beta_{20}$$

$$3) a_{11} = \beta_{11} - \gamma_{12}\beta_{21}$$

$$4) a_{21} = \beta_{21}$$

$$5) a_{12} = \beta_{12} - \gamma_{12}\beta_{22}$$

$$6) a_{22} = \beta_{22}$$

$$e_{1t} = \varepsilon_{yt} - \gamma_{12}\varepsilon_{zt}$$

$$e_{2t} = \varepsilon_{zt}$$

Recordando que:

$$7) VAR(e_{1t}) = VAR(\varepsilon_{yt} - \gamma_{12}\varepsilon_{zt}) = E(\varepsilon_{yt}^2 - \gamma_{12}^2\varepsilon_{zt}^2 - 2\gamma_{12}\varepsilon_{yt}\varepsilon_{zt})$$

$$8) \text{Var}(e_{2t}) = \text{Var}(\varepsilon_{2t}) = E(\varepsilon_{2t}^2) = \sigma_{2t}^2$$

$$9) \text{COV}(e_{1t}, e_{2t}) = E[(\varepsilon_{1t} - \gamma_{12}\varepsilon_{2t})(\varepsilon_{2t})] = E[\varepsilon_{1t}\varepsilon_{2t} - \gamma_{12}\varepsilon_{2t}^2] = -\gamma_{12}\sigma_{2t}^2$$

Tenemos 9 parámetros estimados:

$$a_{10}, a_{11}, a_{12}, a_{20}, a_{21}, a_{22}, \text{Var}(e_{1t}), \text{Var}(e_{2t}) \text{ y } \text{COV}(e_{1t}, e_{2t})$$

Que pueden ser sustituidos dentro de las nueve ecuaciones de arriba en orden de resolverlo simultáneamente para obtener los parámetros de la forma estructural que son también 9.

$$\gamma_{10}, \gamma_{12}, \gamma_{20}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \text{VAR}(\varepsilon_{1t}); \text{VAR}(\varepsilon_{2t})$$

Note que los estimados de $\{\varepsilon_{1t}\} - y - \{\varepsilon_{2t}\}$ pueden ser recuperados.

Note que esta estrategia de identificación implica que la variable más exógena viene primero.

Cuando se utiliza una identificación recursiva (también llamada de Cholesky) debemos tener en cuenta que un cambio en el orden de las variables cambiará totalmente los resultados, por ello este método bastante frágil y muy poco recomendable.

IV. ESTACIONARIEDAD DEL VAR Y REPRESENTACIÓN MA (∞)

Sea:

$$y_t = \pi_0 + \sum_{i=1}^p \pi_i y_{t-i} + e_t \quad 19.$$

$$Y_t - \pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \dots - \pi_p Y_{t-p} = \pi_0 + e_t$$

$$(I - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \dots - \pi_p L^p) Y_t = \pi_0 + e_t \quad 20.$$

Se puede escribir de la siguiente forma:

$$\Phi(L)Y_t = \pi_0 + e_t$$

Un VAR es estacionario en covarianzas si y solo si las pxn raíces del polinomio formado de la matriz de variables endógenas y endógenas rezagadas $\Phi(L)$ caen fuera del círculo unitario o las pxn raíces características λ_i son menores a uno en valor absoluto.

Sea el VAR de orden 1:

$$Y_t = \pi_0 + \pi_1 Y_{t-1} + e_t$$

$$Y_t = \pi_0 + \pi_1 (\pi_0 + \pi_1 Y_{t-2} + e_{t-1}) + e_t$$

$$Y_t = \pi_0 + \pi_1 (\pi_0 + \pi_1 Y_{t-2} + e_{t-1}) + e_t$$

$$Y_t = (I + \pi_1) \pi_0 + \pi_1^2 Y_{t-2} + \pi_1 e_{t-1} + e_t$$

$$Y_t = (I + \pi_1) \pi_0 + \pi_1^2 (\pi_0 + \pi_1 Y_{t-3} + e_{t-2}) + \pi_1 e_{t-1} + e_t$$

$$Y_t = (I + \pi_1 + \pi_1^2) \pi_0 + \pi_1^3 Y_{t-3} + \pi_1^2 e_{t-2} + \pi_1 e_{t-1} + e_t$$

Generalizando:

$$Y_t = (I + \pi_1 + \pi_1^2 + \pi_1^3 + \dots + \pi_1^n) \pi_0 + \pi_1^{n+1} Y_{t-(n+1)} + \sum_{i=0}^n \pi_1^i e_{t-i}$$

Si n tiende al infinito la convergencia requiere que A_1^{n+1} desaparezca.

La condición de estabilidad requiere que los eigenvalores de la matriz A_1 sean menores a la unidad.

$$|\pi_1 - \lambda I| = 0$$

$$|\lambda_i| < 1$$

Podemos escribir la solución particular como:

$$Y_t = \left(\frac{\pi_0}{I - \pi_1} \right) + \sum_{i=0}^{\infty} \pi_1^i e_{t-i}$$

$$Y_t = (I - \pi_1)^{-1} \pi_0 + \sum_{i=0}^{\infty} \pi_1^i e_{t-i}$$

O alternativamente si:

$$Y_t = \pi_0 + \pi_1 Y_{t-1} + e_t$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

$$y_t = a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{1t}$$

$$z_t = a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{2t}$$

$$y_t - a_{11}y_t L - a_{12}z_t L = a_{10} + e_{1t}$$

$$z_t - a_{21}y_t L - a_{22}z_t L = a_{20} + e_{2t}$$

Resolviendo el sistema se tiene que:

$$(1 - a_{11}L)(1 - a_{22}L) - a_{21}a_{12}L^2 = 0$$

Las raíces L deben encontrarse fuera del círculo unitario.

Además tenemos:

$$I - \pi_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 - a_{11} & a_{12} \\ a_{21} & 1 - a_{22} \end{bmatrix}$$

$$(I - \pi_1)^{-1} = \frac{1}{\Delta} \begin{bmatrix} 1 - a_{22} & -a_{12} \\ -a_{21} & 1 - a_{11} \end{bmatrix}$$

$$\Delta = (1 - a_{22})(1 - a_{11}) - a_{12}a_{21}$$

$$(I - \pi_1)^{-1} \pi_0 = \frac{1}{\Delta} \begin{bmatrix} 1 - a_{22} & -a_{12} \\ -a_{21} & 1 - a_{11} \end{bmatrix} \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix}$$

$$(I - \pi_1)^{-1} \pi_0 = \frac{1}{\Delta} \begin{bmatrix} a_{10}(1 - a_{22}) - a_{20}a_{21} \\ a_{20}(1 - a_{11}) - a_{10}a_{12} \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix}$$

Por lo tanto para todo VAR estacionario se puede obtener una representación MA (00) expresado de la siguiente forma:

$$Y_t = \mu + \varepsilon_t + \psi\varepsilon_{t-1} + \psi\varepsilon_{t-2} + \dots + \psi\varepsilon_{t-p} + \dots = \mu + \psi(L)\varepsilon_t$$

$$\psi(L) = [I + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots]$$

V. ESTIMACIÓN DEL VAR

Trabajando en general con un VAR (p) de la forma **Fuente especificada no válida.:**

$$Y_t = \pi_0 + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \dots + \pi_p Y_{t-p} + e_t \quad 21.$$

Se puede observar que:

1. Se tiene un problema de sobreparametrización: hay que estimar $n \cdot p + n$ parámetros, lo que produce un grave problema de pérdida de grados de libertad. La selección adecuada de los grados de libertad es importante; si p es bastante pequeña, el modelo está mal especificado si p es bastante grande los grados de libertad se desperdician. No obstante, **el objetivo de un VAR es encontrar la interrelación entre las variables y no realizar predicciones de corto plazo**, lo que reduce la importancia del problema.
2. Dado que se trabaja con la forma reducida, los errores de cada ecuación están autocorrelacionados y tienen varianza constante, el mejor método de estimación es aplicar MCO ecuación por ecuación. No obstante, para que sea un estimador eficiente todas las ecuaciones deben tener igual número de rezagos de cada explicativa.

En términos prácticos se recomienda utilizar la siguiente receta¹:

¹ Proporcionado en el Curso de Econometría del 49 Curso de Extensión Universitaria BCRP.

1. Limpiar cada una de las series de cualquier tipo de no estacionariedad. Las variables que componen el vector son estacionarias (salvo para el caso de cointegración), esto permite que los test hechos sobre el VAR tengan distribuciones estándar.
2. Estimar por MCO cada ecuación, individualmente.
3. Determinar el número de rezagos de las variables explicativas que deben permanecer en cada ecuación. Para ello se sugieren aplicar:
 - El Test de Máxima Verosimilitud para el conjunto de ecuaciones. La hipótesis nula de este test es que el sistema tiene un número i de rezagos versus la alternativa de que este número es $i+r$. El estadístico sería:
 - Si alguna de las ecuaciones del VAR tiene regresores que no están incluidas en las otras entonces se puede usar el método SUR que provee de estimadores eficientes de los coeficientes VAR. Así cuando allí existe una buena razón para permitir que la longitud de los rezagos difieran de una ecuación a otra, usar SUR.

El planteamiento formal del Test es el siguiente:

$$(T - c) * \left[\log \left| \hat{\Sigma}_i \right| - \log \left| \hat{\Sigma}_{i+r} \right| \right] \quad 22.$$

Donde:

$\log |\hat{\Sigma}_i|$ = logaritmo del determinante de la matriz de varianzas y covarianzas para el modelo con a rezagos.

T = número de observaciones

c = parámetros del modelo no restringido en cada ecuación = $m \times (r+1)$.

Este test se distribuye χ^2 con grados de libertad igual al número de restricciones en el sistema $q = m^2 \times r$. Este test tiene poco poder para rechazar test sucesivos de restricción de rezagos; por ello el rezago referencial debe ser el de mayor valor en el sistema, es decir, cualquier hipótesis nula debe ser contrastada contra el rezago $(i+r)$.

También se puede usar el LR test:

$$LR = (T) * \left[\log \left| \hat{\Sigma}_i \right| - \log \left| \hat{\Sigma}_{i+r} \right| \right] \quad 23.$$

4. No se debe utilizar el test t ni dar importancia a los signos de los coeficientes, ya que existe una gran multicolinealidad entre las variables de cada ecuación. La magnitud de los coeficientes es un indicador relativo de la significancia de la variable (un coeficiente pequeño generalmente acompaña a una variable poco significativa).

El ratio MV es un test basado en la teoría asintótica que no será muy útil en muestras pequeñas. Además el test del Ratio de MV es solo aplicable cuando un modelo es una versión restringida de la otra. Un test alternativo para determinar la longitud apropiada de los rezagos son las generalizaciones multivariadas del AIC y SBC.

Rezago óptimo - criterios de Akaike y Schwartz: beneficio de mayor verosimilitud vs. Penalización por inclusión de variables adicionales:

$$\text{Akaike AIC} (p) = -2 L/T + (2p/T) \quad 24.$$

$$\text{Schwartz SIC} (p) = -2L/T + p \ln (T)/T \quad 25.$$

Donde

L = logaritmo de la función de máxima verosimilitud (suponiendo distribución normal multivariada).

p = número de parámetros estimados.

T = número de observaciones.

Selección: valor más bajo de criterio de información

5. Llevando a cabo la prueba de exclusión de rezagos para cada rezago dentro del VAR. El estadístico de Wald χ^2 para la prueba de significancia conjunta de todas las variables endógenas en lo que los rezagos son reportados para cada ecuación separadamente y conjuntamente.

RESUMEN:

Supuestos en estimación de VAR:

- Variables que componen el vector son estacionarias (salvo para el caso de cointegración).
- Esto permite que los tests hechos sobre el VAR tengan distribuciones estándar.
- Inclusión de variables no estacionarias sujetas a los mismos problemas que en caso univariado: distribuciones no estándar (salvo cointegración).
- Resumen de comportamiento dinámico de series.
- Forma reducida no permite interpretación estructural.
- Bajo supuesto de orden causal, podemos analizar dinámica y recuperar forma estructural.

VI. LA FUNCIÓN IMPULSO-RESPUESTA Y LA DESCOMPOSICIÓN DE LA VARIANZA

Los coeficientes de regresión estimados por un modelo VAR tienen cierta dificultad de Interpretación, por lo que más acertado es obtener la llamada función

impulso respuesta (FIR) y también la descomposición de la varianza del sistema. Entonces la forma común de evaluar comportamiento dinámico: Es el análisis Impulso-Respuesta que es la respuesta de las variables en VAR ante un shock en una de las variables del modelo.

- ¿Qué nos interesa más, shocks a e_{it} (innovaciones de forma reducida) ó shocks a ε_{it} (innovaciones estructurales)?
- Vimos que e_{it} es en realidad combinación lineal de shocks estructurales $\varepsilon_{it} \rightarrow$ no sabríamos como interpretarlo.
- El objetivo es ver impacto de shocks estructurales.

Se trata de conocer la reacción de las variables del sistema frente a shocks. La FIR representa la reacción de la variable endógena ante un cambio de una de las variables aleatorias (shocks). Así por ejemplo, en nuestro caso, dada una variación del efecto será inmediato en pero también habrá un efecto en los valores futuros de dicha variable y en el valor futuro de las otras variables debido al carácter dinámico interconectado del sistema.

La FIR calcula el efecto presente y futuro en las variables endógenas, ante una variación "shock" expresado en el tamaño de su desviación estándar.

6.1. REPRESENTACIÓN (MA) DE LOS MODELOS VAR

Una forma alternativa de representación del VAR consiste en hacer depender el vector de valores actuales de las variables del valor actual y los infinitos rezagos del vector de errores estructurales:

$$Y_t = \sum_{j=1}^p \pi_j L^j Y_t + e_t \quad 26.$$

$$\left[I - \sum_{i=1}^p \pi_i L^i \right] Y_t = e_t \quad 27.$$

$$\pi(L)Y_t = e_t \quad 28.$$

$$Y_t = \frac{e_t}{\pi(L)} \quad 29.$$

Propiedad en matrices:

$$Y_t = \mu + e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \dots$$

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i e_{t-i} \quad 30.$$

Si aumentamos s periodos adelante tenemos lo siguiente:

$$Y_{t+s} = \mu + e_{t+s} + \psi_1 e_{t+s-1} + \psi_2 e_{t+s-2} + \dots + \psi_s e_t + \psi_{s+1} e_{t+1} + \dots$$

$$Y_{t+s} = \mu + \sum_{i=0}^{\infty} \psi_i e_{t+s-i} \quad 31.$$

Cuál es el efecto s periodos adelante de un shock dado en el instante t ($s < t$)

$$\frac{\partial Y_{t+s}}{\partial e_t} = \psi_s \quad \text{Que viene a ser la matriz de multiplicadores de impacto.}$$

$$\frac{\partial Y_{i,t+s}}{\partial e_{j,t}} = \psi_{ij}(s) \quad \text{Efecto en la variable i de un shock en la variable j.}$$

Asimismo con las funciones impulso respuesta se pueden sustentar empíricamente los mecanismos de transmisión.

Ejemplo: La función impulso respuesta para un VAR (1):

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \quad 32.$$

Obteniendo la versión MA del VAR en forma recursiva:

$$Y_t = \pi_0 + \pi_1 Y_{t-1} + e_t$$

$$Y_t = \pi_0 + \pi_1 (\pi_0 + \pi_1 Y_{t-2} + e_{t-1}) + e_t$$

$$Y_t = \pi_0 + \pi_1 \pi_0 + \pi_1^2 Y_{t-2} + \pi_1 e_{t-1} + e_t$$

$$Y_t = \pi_0 + \pi_1 \pi_0 + \pi_1^2 (\pi_0 + \pi_1 \pi_0 + \pi_1^2 Y_{t-3} + \pi_1 e_{t-2}) + e_t + \pi_1 e_{t-1} + e_t$$

Otra forma de expresarlo:

$$(I - \pi_1 L) Y_t = \pi_0 + e_t$$

$$Y_t = \frac{\pi_0 + e_t}{(I - \pi_1 L)}$$

$$Y_t = (I - \pi_1 L)^{-1} \pi_0 + (I - \pi_1 L)^{-1} e_t$$

$$Y_t = \mu + (I - \pi_1 L)^{-1} B^{-1} \varepsilon_t$$

$$Y_t = \mu + \psi(L) B^{-1} \varepsilon_t$$

$$Y_t = \mu + \phi(L) \varepsilon_t$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix} \quad \text{Para } i = 1, 2, 3, \dots \quad 33.$$

Es una representación MA (∞).

Esta representación pudo ser transformada de tal forma que los valores actuales de las variables sean una función de los valores presentes y pasados de un vector de innovaciones ortogonales: como los errores en (19) no tienen por qué estar no correlacionados, se acostumbra pre multiplicar dicha ecuación por la única matriz triangular (T), con unos en la diagonal principal, que diagonalizará la matriz de covarianzas del error. Así, se obtiene un nuevo modelo con errores ortogonales:

Ejemplo: función de IR para un VAR (1)

Se tiene:

$$\begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{yt} \\ e_{zt} \end{bmatrix} \quad \Sigma_e = \begin{bmatrix} \sigma_y^2 & \sigma_{yz} \\ \sigma_{zy} & \sigma_z^2 \end{bmatrix} \quad 34.$$

Suponga que en $t = 0$ $e_{zt}=1 \rightarrow Z_t$ se incrementa en una unidad. ¿Cómo responde el sistema?

VII. DESCOMPOSICIÓN DE CHOLESKY

Es la modificación que normaliza la varianza de las innovaciones estructurales a la unidad bajo el supuesto de ordenamiento.

Diagonalizando:

$$P Y_t = P \sum_{i=1}^p \pi_i Y_{t-i} + W_t \quad 35.$$

Dónde: $W_t = P e_t$, es el vector de las innovaciones ortogonalizadas, y $D = P \Sigma_e P'$.

Es decir, para cada matriz Σ_e real, simétrica y definida positiva existe una única matriz triangular baja P con unos en la diagonal y una única matriz diagonal D con entradas positivas en la diagonal, tal que:

$$\begin{aligned}\Sigma_e &= P D P' \\ \Sigma_e &= P D P' = P D^{1/2} D^{1/2} P' = V V'\end{aligned}\quad 36.$$

Analizando la matriz de covarianza de los shocks estructurales

$$\begin{aligned}e_t &= B^{-1} \varepsilon_t \\ E(\varepsilon_t \varepsilon_t') &= E(B e_t e_t' B') = B E(e_t e_t') B' = B \Sigma_e B'\end{aligned}$$

Si se quiere obtener un nuevo modelo con errores ortogonales, bastará con hacer $P = B^{-1}$, y además imponemos restricciones a la matriz B de forma tal que B es una matriz triangular inferior con unos diagonal principal:

$$E(W_t W_t') = [B^{-1}] E(e_t e_t') [B^{-1}]' \quad 37.$$

$$= [B^{-1}] \Sigma_e [B^{-1}]' \quad 38.$$

$$= [B^{-1}] B D B' [B^{-1}]^{-1} = D \quad 39.$$

$$= D = \begin{bmatrix} w_{11} & \cdot & \cdot & 0 \\ 0 & w_{22} & \cdot & \cdot \\ \cdot & \cdot & w_{33} & \cdot \\ 0 & \cdot & \cdot & w_{mm} \end{bmatrix} \quad 40.$$

Donde D , la matriz de varianzas y covarianzas de los errores transformados, es una matriz diagonal que garantiza su ortogonalidad. Entonces la matriz de covarianzas de los errores ortogonalizados es la matriz diagonal, esto quiere decir que los mismos tienen todos sus componentes, incorrelacionados cada uno con varianza que vienen a ser los valores de la diagonal principal. El orden de las variables para descomponer matriz de covarianzas es importante, implica que shocks en la segunda variable no tienen efectos contemporáneos sobre la primera y así sucesivamente.

Por lo visto, podemos decir que la descomposición de Cholesky plantea la modificación que normaliza varianza de innovaciones estructurales a la unidad (igual supuesto de ordenamiento). A partir de este modelo transformado se puede obtener la función impulso-respuesta ortogonalizada, calculando el efecto sobre Y_{t+s} de un impulso unitario en w_j . Estos multiplicadores, describen cómo nueva información sobre y_j nos lleva a revisar nuestra predicción de Y_{t+s} , aun cuando la definición implícita de nueva información es diferente para cada variable j .

El orden en que se coloquen las variables en el sistema tendrá un impacto importante sobre los multiplicadores calculados: al elegir un orden recursivo particular de las variables, implícitamente, se responde a un conjunto de preguntas específicas respecto a la predicción; el ordenamiento dependerá de la razón por la cual queremos responder esas preguntas en primer lugar. Asimismo, pueden existir razones teóricas para suponer que una de las variables no se ve afectada por los shocks contemporáneos de las otras.

La importancia del ordenamiento depende de la magnitud de la correlación de los errores no ortogonalizados del sistema. En la práctica, una correlación baja (menor

a 0.2) disminuye la relevancia de un ordenamiento adecuado. Si la correlación es elevada, en cambio, será indispensable probar diferentes ordenamientos y analizar cuánto cambian los resultados. En principio, una buena especificación del sistema debería arrojar resultados muy similares con cualquier ordenamiento utilizado **Fuente especificada no válida.**

Resultados ante distintos ordenamientos diferirán más mientras mayor sea la correlación entre los residuos del VAR.

En resumen podemos plantearnos lo siguiente:

- Teoría económica
- ¿Son funciones impulso respuesta muy distintas bajo distintos ordenamientos?. Contradice la teoría económica.

Asimismo, es posible realizar un análisis de descomposición de la varianza a partir del modelo ortogonalizado. Este consistirá en calcular la contribución de la innovación j sobre el error de predicción del período $t+s$. Es de esperar que en el corto plazo la propia innovación explique la mayor proporción de este error. Cabe resaltar que este análisis también se ve afectado por el ordenamiento de las variables del sistema, por lo que se sugiere probar diferentes ordenamientos, al igual que en el caso de la función impulso- respuesta.

VIII. DESCOMPOSICIÓN DE VARIANZA.

- Se trata de conocer la contribución de la j -ésima innovación ortogonalizada en el error cuadrático medio (ECM) de la predicción s periodos adelante. La descomposición de varianza del error de predicción nos dice que proporción de los movimientos (variabilidad) en una secuencia de uno de

los componentes de Y_t son explicados por sus propios shocks y que proporción lo explican los shocks de los otros componentes.

- Nos mide la proporción (que parte de los movimientos en una de las variables esta explicado por sus propios choques y que parte esta explicado por los choques de las otras variables.
- Una forma fácil de ver cuál es la variable más exógena es la gráfica y comparación de cada una de las series en estudio. Si los movimientos de la serie están explicados por ella misma entonces la variable es exógena o más exógena.

Supóngase que conocemos los coeficientes de π_0 y π_1 y que queremos pronosticar los diversos valores de X_{t+i} condicionados a los valores X_t .

Por ejemplo:

$$\begin{aligned} Y_{t+1} &= \pi_0 + \pi_1 Y_t + e_{t+1} \\ E_t(Y_{t+1}) &= \pi_0 + \pi_1 Y_t \end{aligned} \quad 41.$$

El error de predicción, condicionado a la información disponible en el periodo t , con un periodo de adelanto es:

$$Y_{t+1} - E_t(Y_{t+1}) = e_{t+1} \quad 42.$$

Similarmente:

$$\begin{aligned} Y_{t+2} &= \pi_0 + \pi_1 Y_{t+1} + e_{t+2} \\ Y_{t+2} &= \pi_0 + \pi_1 (\pi_0 + \pi_1 Y_t + e_{t+1}) + e_{t+2} \\ Y_{t+2} &= \pi_0 + \pi_1 \pi_0 + \pi_1^2 Y_t + \pi_1 e_{t+1} + e_{t+2} \\ E_t(Y_{t+2}) &= \pi_0 + \pi_1 \pi_0 + \pi_1^2 Y_t \end{aligned} \quad 43.$$

y el error de predicción asociado es:

$$Y_{t+2} - E_t(Y_{t+2}) = \pi_1 e_{t+1} + e_{t+2}$$

$$\text{Sea } E_t(Y_{t+n}) = \hat{Y}_{t+n/t}$$

Generalizando tenemos:

$$\left(Y_{t+n} - \hat{Y}_{t+n/t} \right) = \pi_1 e_{t+n-1} + \pi_1^2 e_{t+n-2} + \dots + \pi_1^{n-1} e_{t+1} \quad 44.$$

También se puede obtener el error de predicción en términos de VMA dado que el VAR en la forma reducida y el VMA contienen la misma información.

$$Y_t = \mu + \psi(L) B^{-1} \varepsilon_t$$

$$Y_t = \mu + \phi(L) \varepsilon_t$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

$$Y_{t+n} = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t+n-i} \quad 45.$$

Así que el error de predicción será:

$$\text{ECM}(\hat{Y}_{t+n/t}) = \left(Y_{t+n} - \hat{Y}_{t+n/t} \right) = \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i} \quad 31.$$

IX. CAUSALIDAD

Generalmente no resulta fácil determinar la existencia de una relación de causalidad entre dos variables y menos aún su dirección. Para ello se recurre, en la práctica, a la teoría económica aun cuando la verificación de la validez de dicha relación resulta ser poco rigurosa si se pre condiciona la misma a la existencia de una teoría que la apoya. La alternativa econométrica de verificación consiste en general en estimar una regresión entre las variables que se analizan y observar la significancia de los coeficientes obtenidos. Sin embargo, una alta correlación entre

dos variables no asegura una relación causa-efecto entre ellas, ya que la posibilidad de que se haya obtenido una correlación espúrea no debe ser descartada. Como es bien sabido la correlación espúrea se da debido a que el alto valor del coeficiente de correlación se explica principalmente por la presencia de un tercer factor y no por la existencia de una relación con sentido económico entre las variables.

Es por estas razones que se ha desarrollado el concepto de causalidad en econometría, el mismo que se basa en el desarrollo teórico llevado a cabo por Granger². Formalmente podemos definir la llamada causalidad a lo Granger diciendo que: la variable x causa a la variable y si al tomar en cuenta los valores pasados de x se mejoran las predicciones de y . Granger se refiere a las predicciones insesgadas de y que se obtienen a través de la estimación por MCO, midiendo la precisión de las mismas a través de la varianza del error de predicción que para el caso de insesgamiento viene a ser error cuadrático medio (Beltrán Barco, 2003):

$$ECM = E[(\hat{Y} - Y)^2] \quad 46.$$

Así, x causa a y si:

$$ECM(\hat{Y}/\bar{U}) < (\hat{Y}/\bar{U} - \bar{x}) \quad 47.$$

Mientras que X causa instantáneamente a Y si:

$$ECM(\hat{Y}/U) < (\hat{Y}/U - x) \quad 48.$$

Donde U es un set de toda la información disponible pasada y presente, \bar{U} es el mismo set pero que contiene sólo la información pasada, X contiene toda la información pasada y presente de dicha variable y \bar{x} sólo la información pasada.

Asimismo, \hat{y} es el predictor MCO de y .

La posibilidad de hacer operativa esta definición pasa por la necesidad de acotar el set U de toda la información disponible, aun cuando ello implique recurrir a la teoría económica y , de algún modo, perder parte de la objetividad ganada aplicando el concepto mismo de causalidad. Son principalmente tres los test que se utilizan para verificar la existencia de una relación de causalidad entre dos series de tiempo.

9.1. EL TEST DIRECTO

La hipótesis a testear será:

$$H_o : y \rightarrow x \quad 49.$$

Es decir, se requiere examinar si y no causa a x . Para ello se estima la relación entre x , los m primeros rezagos de y_t y los n primeros rezagos de x_t :

$$x_t = \delta + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{i=1}^p \gamma_i x_{t-i} + u_t \quad 50.$$

Luego se verifica la significancia conjunta de los rezagos de y_t ; si ellos no fueran significativos entonces x_t estaría explicado solamente por su propio pasado y por el elemento aleatorio respectivo.

Para que el test arroje un resultado correcto es necesario asegurarse que la especificación del modelo sea la adecuada, de tal forma que el error asociado sea ruido blanco. Por lo mismo se requiere escoger cuidadosamente el valor de p , especialmente en el caso de los rezagos de x , ya que la eliminación de estos últimos, si es que son importantes para explicar el comportamiento de x , inflarían artificialmente la significancia de los rezagos de y llevándonos a obtener una conclusión falsa respecto de la relación de causalidad entre x é y . Cuando una variable y_t no es útil para predecir otra variable x_t , entonces se dice que “ y_t no causa a x_t en el sentido Granger”.

$$E(X_{t+s} / X_t, X_{t-1}, \dots) = E(X_{t+s} / X_t, X_{t-1}, X_{t-2}, \dots; Y_t, Y_{t-1}, Y_{t-2}, \dots) \quad 51.$$

9.2. TEST DE SIMS

Consiste en regresionar x_t en función de los valores pasados y futuros de y_t , de forma tal de testear que y_t no causa x_t sobre la base de la significancia de los coeficientes asociados con los valores futuros de x_t . Dicho de otro modo, si existe una relación entre el valor presente de x_t y los valores futuros de y_t ésta debe expresar una causalidad de x a y y no de y a x , ya que el futuro no puede causar al presente (ver nota 1)². Así, se plantea correr la regresión:

$$x_t = \sum_{\tau=-m}^p \gamma_{\tau} y_{t-\tau} + z_t$$

$$x_t = \sum_{\tau=1}^p \beta_{\tau} y_{t-\tau} + \sum_{\tau=1}^p \gamma_{\tau} y_{t+\tau} + z_t \quad 52.$$

Donde se testear la $H_0 \gamma_{\tau} = 0 \forall \tau = 1, 2, \dots, p$

Lo más importante de la aplicación de este test es que para que arroje un resultado correcto el error de la ecuación debe ser ruido blanco. De no ser así, deberá ser estimada por mínimos cuadrados generalizados, luego de identificar la estructura autorregresiva del error.

² El teorema de Sims que sustenta esta metodología de testeo de la causalidad sostiene que: “cuando (x, y) tienen una representación AR, x puede ser expresado como una función de rezagos y valores actuales de y , con un residuo no correlacionado con ningún valor de y , pasado o futuro, si y sólo si y no causa a x en el sentido Granger”.

REFERENCIAS BIBLIOGRÁFICAS

Beltrán Barco, Arlette. (2003). *Econometría de series de tiempo. Notas de clase de Econometria II*. Lima: Mimeo. Universidad del Pacífico.
<https://econometriaaii.files.wordpress.com/2010/01/beltran.pdf>

Hamilton, J. (1994). *Time series analysis*. Princeton University.

Londoño, W. (2005). *Modelos de ecuaciones múltiples modelos VAR y cointegración*. Medellín: Departamento de Ciencias Básicas, Maestría en Matemáticas Aplicadas. Universidad de EAFIT .
https://repository.eafit.edu.co/bitstream/handle/10784/134/Wbaldo_Londo%C3%B1o_2005.pdf?sequence=3&isAllowed=y

Sims, C. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1-48.