1 Forma estructural y reducida de un modelo de ecuaciones simultaneas.

Tomenos como ejemplo el siguente modelo de ecuaciones simultaneas:

$$
\begin{aligned}
y_{1t} &= \gamma_1 + \alpha_{12} y_{2t} + \alpha_{11} x_{1t} + u_{1t} \\
y_{2t} &= \gamma_2 + \alpha_{21} y_{1t} + \alpha_{22} x_{2t} + u_{2t}
\end{aligned}
$$

Este modelo presenta dos variables endógenasy_{1t} y y_{2t} y dos variables exogenas o predeterminadas ademas del término independiente. Resulta conveniente reescribir el modelo anterior de forma que todas las variables endógenas esten recogidas en el primer término de la identidad:

$$
\begin{aligned}
y_{1t} - \alpha_{12} y_{2t} &= \gamma_1 + \alpha_{11} x_{1t} + u_{1t} \\
y_{2t} - \alpha_{21} y_{1t} &= \gamma_2 + \alpha_{22} x_{2t} + u_{2t}
\end{aligned}
$$

que se puede escribir matricialmente como:

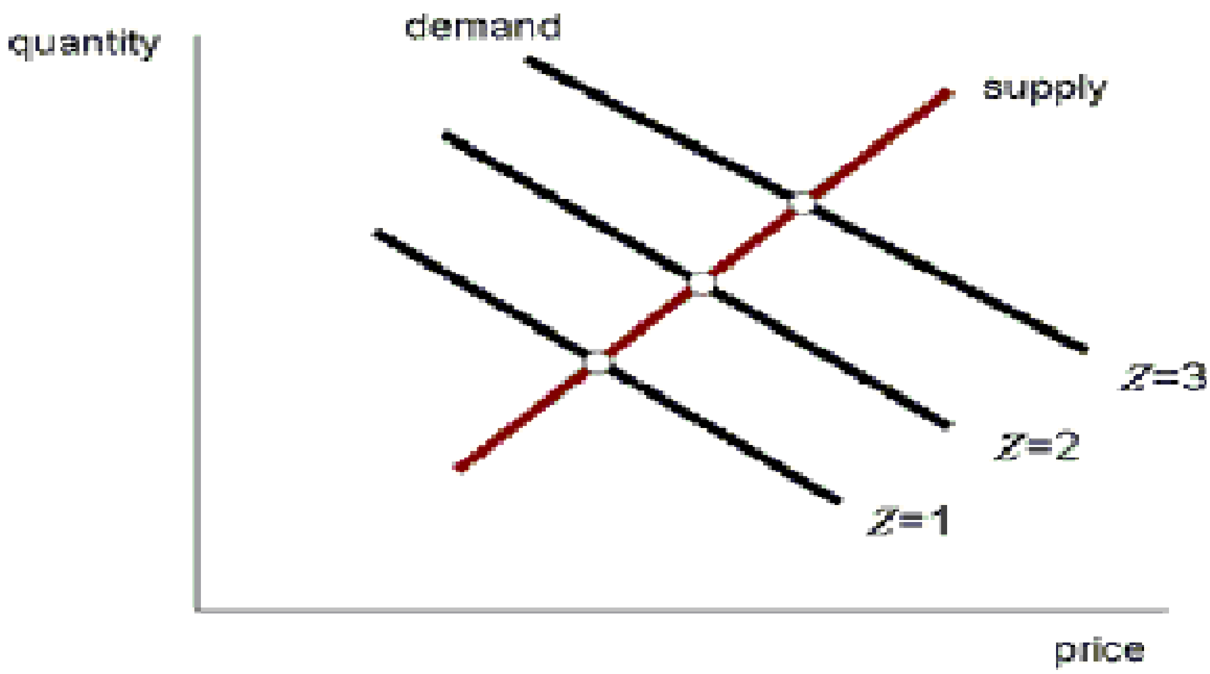
$$
\begin{aligned}
A\boldsymbol{y}_t &= B\boldsymbol{x}_t + \boldsymbol{u}_t \\
\boldsymbol{y}_t' &= [y_{1t}, y_{2t}] \quad \boldsymbol{x}_t' = [1, x_{1t}, x_{2t}] \quad \boldsymbol{u}_t' = [u_{1t}, u_{2t}] \\
A &= \begin{bmatrix} 1 & -\alpha_{12} \\ -\alpha_{21} & 1 \end{bmatrix} \quad B = \begin{bmatrix} \gamma_1 & \alpha_{11} & 0 \\ \gamma_2 & 0 & \alpha_{22} \end{bmatrix}.
\end{aligned} \tag{1}
$$

la anterior representación se conoce con el nombre de forma estructural de un modelo de ecuaciones de simultaneas. Los modelos de ecuaciones simultáneas tienen dos caraterísticas importante, el primer lugar es fácil ver que tanto y_{2t} como y_{1t} estarán correlacionados con los términos de perturbación u_{1t} y u_{2t} respectivamente. Es decir, al igual que los modelos dinámicos tendremos problemas de correlación entre las variables endógenas del modelo que aparecen como explicativas en las ecuaciones del modelo y los términos de perturbación de esas ecuaciones. En segundo lugar existe un problema de identificación que consiste en que aunque creamos que estamos estimando una ecuación del modelo esa ecuación puede no estar identificada. Un ejemplo de modelo en el que las ecuaciones no están identificadas es el modelo de oferta y demanda siguiente:

$$
\begin{aligned}
q_i^D - \alpha_1 p_i &= \gamma_1 + u_{1i} \\
q_i^S - \alpha_2 p_{ti} &= \gamma_2 + u_{2i} \qquad (2) \\
q_i^* &= q_i^D = q_i^S
\end{aligned}
$$

$$
\begin{aligned}
A\boldsymbol{y}_i &= B\boldsymbol{x}_i + \boldsymbol{u}_i \\
\boldsymbol{y}_i' &= [q_i^*, p_i] \quad \boldsymbol{x}_i' = [1] \quad \boldsymbol{u}_i' = [u_{1i}, u_{2i}] \\
A &= \begin{bmatrix} 1 & -\alpha_1 \\ 1 & -\alpha_2 \end{bmatrix} \quad B = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}.
\end{aligned}
$$

Para que la ecuación de un modelo de ecuaciones simultaneas esté identificada se tiene que poder pasar de la forma estructural del modelo $A\boldsymbol{y}_t = B\boldsymbol{x}_t + \boldsymbol{u}_t$ a la forma reducida $\boldsymbol{y}_t = \Pi\boldsymbol{x}_t + \boldsymbol{v}_t$. En el caso del anterior sistema de oferta y demanda, nínguna de las ecuaciones anteriores está identificada, dado que no se puede pasar de la forma reducida a la forma estructural.



Si partimos de la forma estructural del modelo de ecuaciones simultaneas y premultiplicamos por A^{-1} llegamos a la forma reducida del modelo:

$$
\begin{aligned}
A^{-1}A\boldsymbol{y}_t &= A^{-1}B\boldsymbol{x}_t + A^{-1}\boldsymbol{u}_t \\
\boldsymbol{y}_t &= A^{-1}B\boldsymbol{x}_t + A^{-1}\boldsymbol{u}_t \\
\boldsymbol{y}_t &= \Pi\boldsymbol{x}_t + \boldsymbol{v}_t \\
\Pi &= A^{-1}B \quad A\Pi = B \quad \boldsymbol{v}_t = A^{-1}\boldsymbol{u}_t.
\end{aligned}
$$

La forma reducida expresa todas las variables endógenas $\boldsymbol{y}_t$ únicamente en función de las variables exógenas $\boldsymbol{x}_t$ ($\boldsymbol{y}_t = \Pi \boldsymbol{x}_t + \boldsymbol{v}_t$). La forma reducida del modelo, siempre se puede estimar, pero el problema que presenta es que no tiene interpretación económica. La forma estructural $A\boldsymbol{y}_t = B\boldsymbol{x}_t + \boldsymbol{u}_t$ si que tiene interpretación económica pero ha de estar indentificada. Para que las ecuaciones de la forma estructural estén identificas se ha de poder obtener los coeficientes de la forma estructural recogidos en A y en B a partir de los coeficientes de la forma reducida Π. Es decir, resolver el sistema $A\Pi = B$ donde tenemos más incognita los elementos de A y B que coeficientes los elementos de Π o que el número de ecuaciones. Hay que tener en cuenta que la dimensión de A y B es $q \times q$ y $q \times k$ respectivamente. Donde q es el número de variables endógenas del sistema y k es el número de variables, por lo tanto tenemos $q \times q$ y $q \times k$ incognitas y $q \times k$ coeficientes conocidos en las q ecuaciones. En el ejemplo del sistema de demanda y oferta (2) hay cuatro incognitas y dos coeficientes conocidos en dos ecuaciones. En el caso de (1) tenemos seis incognitas y seis coeficientes conocidos (los elementos del Π que será de dimensión 2×3) por lo que las ecuaciones de la forma estructural estarán identificadas. Las codiciones que han de cumplir las ecuaciones de la forma estructural para que estén identificadas se conocen como condiciones de rango y orden ver Johnston y Dinardo para mas detalles(Johnston y Dinardo Econometric methods).

En los modelos de ecuaciones simultaneas, las variables endógenas retardadas se les llama variables predeterminadas y recibien el mismo trato que las exóngenas. Por eso los modelos VAR pueden ser vistos como una forma reducida.

2 Modelos VAR

Los modelos VAR (modelos vectoriales autoregresivios) son una generalización a más de una variable de los procesos o modelos autorregresivos AR(p):

$$
\begin{aligned}
x_t &= c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t \\
\varepsilon_t &\sim iid\left(0, \sigma_\varepsilon^2\right)
\end{aligned}
$$

Si tomamos un proceso autorregresivo de orden 1 AR(1) y sustituimos recursivamente vemos que:

$$
\begin{aligned}
x_t &= c + \phi_1 x_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right) \\
x_t &= c + \phi_1 \left(c + \phi_1 x_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t \\
&= c\left(1 + \phi_1\right) + \phi_1^2 x_{t-1} + \varepsilon_t + \phi_1 \varepsilon_{t-1} \\
&\vdots \\
x_t &= c\left(1 + \phi_1 + \cdots + \phi_1^{t-1}\right) + \phi_1^t x_0 + \varepsilon_t + \phi_1 \varepsilon_{t-1} \\
&\quad + \cdots + \phi_1^{t-1} \varepsilon_1 \\
&= c \sum_{j=0}^{t-1} \phi_1^j + \phi_1^t x_0 + \sum_{j=0}^{t-1} \phi_1^j \varepsilon_{t-j}.
\end{aligned}
$$

Es decir, podemos expresarlo como una proceso Media Mobil en el que la suma $\sum_{j=0}^{t-1} \phi_1^j \varepsilon_{t-j}$ recoge el efecto que tiene un schock ε_t sobre x_t a lo largo del tiempo, es decir:

$$
\begin{array}{llllllll}
\multicolumn{8}{l}{\varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \cdots + \phi_1^{t-1} \varepsilon_1} \\
\frac{\partial x_t}{\partial \varepsilon_t} & \frac{\partial x_{t+1}}{\partial \varepsilon_t} & \frac{\partial x_{t+2}}{\partial \varepsilon_t} & \frac{\partial x_{t+3}}{\partial \varepsilon_t} & \frac{\partial x_{t+4}}{\partial \varepsilon_t} & \frac{\partial x_{t+5}}{\partial \varepsilon_t} & \cdots & \frac{\partial x_{t+j}}{\partial \varepsilon_t} \\
1 & \phi_1 & \phi_1^2 & \phi_1^3 & \phi_1^4 & \phi_1^5 & \cdots & \phi_1^j
\end{array} .
$$

En el caso de un proceso AR(p) se puede ver que:

$$
\begin{aligned}
\phi_p(L)\, x_t &= c + \varepsilon_t \quad \varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right) \\
\phi_p(L) &= \left(1 - \phi_1 L - \cdots - \phi_p L^p\right) \\
x_t &= \frac{c}{\phi_p(1)} + \frac{\varepsilon_t}{\phi_p(L)} \\
x_t &= c' + \psi_\infty(L)\, \varepsilon_t \\
\psi_\infty(L) &= \frac{1}{\phi_p(L)} \\
\psi_\infty(L)\, \phi_p(L) &= 1
\end{aligned}
$$

$$
\begin{aligned}
\left(1 + \psi_1 L + \psi_2 L^2 + \cdots\right)\left(1 - \phi_1 L + \phi_2 L^2 + \cdots\right) &= 1 \\
1 + (\psi_1 - \phi_1) L + (\psi_2 - \phi_1 \psi_1 - \phi_2) L^2 + \cdots &= 1
\end{aligned}
$$

$$
\begin{aligned}
\psi_1 &= \phi_1 \\
\psi_2 &= \phi_1\psi_1 + \phi_2 \\
\psi_3 &= \phi_1\psi_2 + \phi_2\psi_1 + \phi_3 \\
\psi_4 &= \phi_1\psi_3 + \phi_2\psi_2 + \phi_3\psi_1 + \phi_4 \\
&\vdots \\
\psi_j &= \sum_{i=1}^{p} \psi_{j-i}\phi_i.
\end{aligned}
$$

Donde el coeficiente ψ_j recoge el efecto del ε_t sobre x_t tras j periodos, es decir, en x_{t+j}.

En lugar de estar interesados en una única variable x_t, estamos interesados en las posibles relaciones dinámicas existentes entre un conjunto de variables recogidas en un vector X_t de orden $n \times 1$. Por tanto, en dicho vector se recogen un conjunto de n variables x_{1t}, x_{2t}, $\cdots$, x_{nt} de interés $X_t = [x_{1t}, x_{2t}, \cdots, x_{nt}]'$.

Un modelo VAR de orden k responde a:

$$
\begin{aligned}
X_t &= \mu + \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \cdots + \Pi_k X_{t-k} + E_t \qquad (3) \\
X_t &= [x_{1t}, x_{2t}, \cdots, x_{nt}]' \ \ n \times 1 \\
\mu &= [\mu_1, \mu_2, \cdots, \mu_n]' \ \ n \times 1 \\
E_t &= [\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{nt}]' \ \ n \times 1 \\
E_t &\sim iiN_n(0, \Omega_\varepsilon)
\end{aligned}
$$

En el caso de las matrices Π_j de orden $n \times n$ con $j = 1, \cdots, k$, que recogen los coeficientes del modelo, tenemos:

$$
\Pi_j = \begin{bmatrix} \pi_{11}^j & \pi_{12}^j & \cdots & \pi_{1n}^j \\ \pi_{21}^j & \pi_{22}^j & \cdots & \pi_{2n}^j \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{n1}^j & \pi_{n2}^j & \cdots & \pi_{nn}^j \end{bmatrix} \quad j = 1, \cdots, k.
$$

Por lo que el modelo quedaría:

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} + \begin{bmatrix} \pi^1_{11} & \pi^1_{12} & \cdots & \pi^1_{1n} \\ \pi^1_{21} & \pi^1_{22} & \cdots & \pi^1_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \pi^1_{n1} & \pi^1_{n2} & \cdots & \pi^1_{nn} \end{bmatrix} \times \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ \vdots \\ x_{nt-1} \end{bmatrix}$$

$$\cdots + \begin{bmatrix} \pi^k_{11} & \pi^k_{12} & \cdots & \pi^k_{1n} \\ \pi^k_{21} & \pi^k_{22} & \cdots & \pi^k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \pi^k_{n1} & \pi^k_{n2} & \cdots & \pi^k_{nn} \end{bmatrix} \times \begin{bmatrix} x_{1t-k} \\ x_{2t-k} \\ \vdots \\ x_{nt-k} \end{bmatrix} +$$

$$+ \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

Que es equivalente al siguiente sistema de ecuaciones:

$$\begin{aligned} x_{1t} = \; & \mu_1 + \pi^1_{11} x_{1t-1} + \pi^1_{12} x_{2t-1} + \cdots + \pi^1_{1n} x_{nt-1} + \\ & + \pi^2_{11} x_{1t-2} + \pi^2_{12} x_{2t-2} + \cdots + \pi^2_{1n} x_{nt-2} + \\ & \vdots \\ & + \pi^k_{11} x_{1t-k} + \pi^k_{12} x_{2t-k} + \cdots + \pi^k_{1n} x_{nt-k} + \varepsilon_{1t} \end{aligned}$$

$$
\begin{aligned}
x_{2t} \;=\;& \mu_2 + \pi_{21}^1 x_{1t-1} + \pi_{22}^1 x_{2t-1} + \cdots + \pi_{2n}^1 x_{nt-1} + \\
& + \pi_{21}^2 x_{1t-2} + \pi_{22}^2 x_{2t-2} + \cdots + \pi_{2n}^2 x_{nt-2} + \\
& \vdots \\
& + \pi_{21}^k x_{1t-k} + \pi_{22}^k x_{2t-k} + \cdots + \pi_{2n}^k x_{nt-k} + \varepsilon_{2t}
\end{aligned}
$$

$$\vdots$$

$$
\begin{aligned}
x_{nt} \;=\;& \mu_n + \pi_{n1}^1 x_{1t-1} + \pi_{n2}^1 x_{2t-1} + \cdots + \pi_{nn}^1 x_{nt-1} + \\
& + \pi_{n1}^2 x_{1t-2} + \pi_{n2}^2 x_{2t-2} + \cdots + \pi_{nn}^2 x_{nt-2} + \\
& \vdots \\
& + \pi_{n1}^k x_{1t-k} + \pi_{n2}^k x_{2t-k} + \cdots + \pi_{nn}^k x_{nt-k} + \varepsilon_{nt}.
\end{aligned}
$$

Finalmente $E_t \sim iiN_n\left(0, \Omega_\varepsilon\right)$ implica que $E\left[E_t\right] = 0$, $E\left[E_t E_t'\right] = \Omega_\varepsilon$ y al ser identicamente e indepedientemente distribuidos $E\left[E_t E_{t-k}'\right] = 0$ para todo $k = \pm 1, \pm 2, \pm 3, \cdots$

Por ejemplo en el caso de un VAR(1) con tres variables $n = 3$ tenemos:

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{bmatrix} \times \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

$$\begin{aligned} x_{1t} &= \mu_1 + \pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \pi_{13} x_{3t-1} + \varepsilon_{1t} \\ x_{2t} &= \mu_2 + \pi_{21} x_{1t-1} + \pi_{22} x_{2t-1} + \pi_{23} x_{3t-1} + \varepsilon_{2t} \\ x_{3t} &= \mu_3 + \pi_{31} x_{1t-1} + \pi_{32} x_{2t-1} + \pi_{33} x_{3t-1} + \varepsilon_{3t}. \end{aligned}$$

Utilizando el operador de retardos podemos escribir:

$$\begin{aligned} X_t &= \mu + \Pi_1 L X_t + \Pi_2 L^2 X_t + \cdots + \Pi_k L^k X_t + E_t \\ \Pi_k (L) X_t &= \mu + E_t \\ \Pi_k (L) &= \left(I_k - \Pi_1 L - \Pi_2 L^2 - \cdots - \Pi_k L^k \right) \end{aligned}$$

Donde $\Pi_k (L)$ es un polinomio en el operador de retardos pero en lugar de tener coeficientes, tendremos matrices de orden $n \times n$. Para trabajar con los modelos VAR, cuando no se considera la existencia de relaciones de cointegración (relaciones a largo plazo) es necesario comprobar que el modelo sea estable y/o que el vector X_t sea estacionario.

El vector X_t será estacionario en sentido débil (weakly or covariance stationary) si se cumple:

$$
\begin{aligned}
E\left[X_t\right] &= \mu_0 < \infty \quad \forall t \\
VAR\left[X_t\right] &= E\left[\left(X_t - E\left[X_t\right]\right)\left(X_t - E\left[X_t\right]\right)'\right] = \\
&= E\left[\left(X_t - \mu_0\right)\left(X_t - \mu_0\right)'\right] = \mathbf{\Sigma}_0 \quad \forall t \\
COV\left[X_t X_{t-h}\right] &= E\left[\left(X_t - \mu_0\right)\left(X_{t-h} - \mu_0\right)'\right] = \\
&= \mathbf{\Sigma}_h \quad \forall t,\ h.
\end{aligned}
$$

Es decir, que los momentos de primer y segundo orden sean independientes del período de tiempo en el que nos encontremos. La estacionariedad en sentido estricto implica que todos los momentos y por tanto toda la distribución conjunta sea invariante ante desplazamientos en el tiempo. Si el modelo es estacionario podemos tomar esperanzas y obtener:

$$\begin{aligned}
X_t &= \mu + \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \cdots + \Pi_k X_{t-k} + E_t \\
E\left[X_t\right] &= \mu + \Pi_1 E\left[X_{t-1}\right] + \Pi_2 E\left[X_{t-2}\right] + \cdots + \\
&\quad + \Pi_k E\left[X_{t-k}\right] + E\left[E_t\right] \\
\mu_0 &= \mu + \Pi_1 \mu_0 + \Pi_2 \mu_0 + \cdots + \Pi_k \mu_0 \\
\mu_0 &= \mu\left(I_k - \Pi_1 - \Pi_2 - \cdots - \Pi_k\right)^{-1} = \mu \Pi_k\left(1\right)^{-1}.
\end{aligned}$$

De donde se aprecia que si no se cumple la siguiente condición $\Pi_k\left(1\right)^{-1} \neq 0$ el valor esperado $E\left[X_t\right]$ no será finito. Otra manera alternativa de comprobar la estacionariedad es tener presente que:

$$\Pi_k\left(1\right)^{-1} = \frac{adj\left(\Pi_k\left(1\right)\right)}{\left|\Pi_k\left(1\right)\right|}.$$

Por lo tanto, para tener un valor esperado finito, es necesario que $\left|\Pi_k\left(1\right)\right| \neq 0$.

También se puede partir del VAR(1):

$$X_t = \mu + \Pi_1 X_{t-1} + E_t$$

e ir sustituyendo recursivamente y ver que:

$$
\begin{aligned}
X_t &= \mu + \Pi_1 \left(\mu + \Pi_1 X_{t-2} + E_{t-1}\right) + E_t \\
&= \left(I + \Pi_1\right) \mu + \Pi_1 X_{t-2} + E_t + \Pi_1 E_{t-1} \\
&\vdots \\
X_t &= \left(I + \Pi_1 + \cdots + \Pi_1^{t-1}\right) \mu + \Pi_1^t X_0 + \sum_{i=0}^{t-1} \Pi_1^i E_{t-i}
\end{aligned}
$$

donde Π_1^t es la potencia de order t de la matriz cuadrada Π_1 de forma que $\Pi_1^2 = \Pi_1\Pi_1$, $\Pi_1^3 = \Pi_1\Pi_1\Pi_1$, ...si los valores propios de Π_1 tienen modulo inferior a la unidad la secuencia Π_1^i, $i = 0, 1, 2, \ldots$ acabaran tendiendo a zero a medida que $i \rightarrow \infty$, ya que $\Pi_1^i = P\Lambda^i P^{-1} \rightarrow 0$ al tender i a infinito(este rdo se basa $\Pi_1 = P\Lambda^i P^{-1}$ en las propiedades de la descomposición de Jordan de una matrix cuadrada).

En el caso de un VAR(k) poder expresarlo como un VAR(1) utilizando:

$$
\begin{aligned}
\boldsymbol{X}_t &= \boldsymbol{\Pi}\boldsymbol{X}_{t-1} + \boldsymbol{E}_t \qquad (4)\\
\boldsymbol{X}_t &= \begin{bmatrix} X_t \\ X_{t-1} \\ X_{t-2} \\ \vdots \\ X_{t-k-1} \end{bmatrix} \quad \boldsymbol{\Pi} = \begin{bmatrix} \Pi_1 & \Pi_2 & \cdots & \Pi_{k-1} & \Pi_k \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ & & \cdots & I & 0 \end{bmatrix}\\
\boldsymbol{E}_t &= \begin{bmatrix} E_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
\end{aligned}
$$

Por lo tanto para comprobar si un VAR(k) es estacionario o estable basta con mirar si se cumple que el modelo de los valores propios de $\boldsymbol{\Pi}$ en (4) son menores que uno.

3 Estimación de los modelos VAR

Los modelos VAR se pueden estimar por MCO y por MV. En primer lugar veremos la estimación por MCO (mínimos cuadrados ordinarios).

Para la estimación MCO resulta útil la siguiente notación:

$$
\begin{aligned}
\mathsf{Y} &= \begin{bmatrix} X_1 & X_2 & X_3 & \cdots & X_T \end{bmatrix} \quad (n \times T) \\
\mathsf{B} &= \begin{bmatrix} \mu & \Pi_1 & \Pi_2 & \cdots & \Pi_k \end{bmatrix} \quad (n \times (nk+1)) \\
Z_t &= \begin{bmatrix} 1 \\ X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{bmatrix} \quad ((nk+1) \times 1) \\
\mathsf{Z} &= \begin{bmatrix} Z_0 & Z_1 & Z_2 & \cdots & Z_{T-1} \end{bmatrix} \quad ((nk+1) \times T) \\
\mathsf{U} &= \begin{bmatrix} E_1 & E_2 & E_3 & \cdots & E_T \end{bmatrix} \quad (n \times T)
\end{aligned}
$$

$$\mathsf{Y} = \mathsf{B}\mathsf{Z} + \mathsf{U}$$

También se puede escribir utilñizando el operador **vec** que transforma una matrix en un vector (si $A_{m \times n} =$

$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$ entonces $\mathbf{vec}(A) = \begin{bmatrix} a_1' & a_2' & \cdots & a_n' \end{bmatrix}'$ es un vector de $mn \times 1$)

$$\begin{aligned}
\mathbf{vec}(\mathsf{Y}) &= \mathbf{vec}(\mathsf{BZ}) + \mathbf{vec}(\mathsf{U}) \\
\mathbf{y} &= \mathbf{vec}(\mathsf{Y}) \\
\mathbf{u} &= \mathbf{vec}(\mathsf{U}) \\
\boldsymbol{\beta} &= \mathbf{vec}(\mathsf{B}) \\
\mathbf{y} &= \left(\mathsf{Z}' \otimes I_K\right)\boldsymbol{\beta} + \mathbf{u}.
\end{aligned}$$

Donde se utiliza la siguiente propiedad del operador $\mathbf{vec}$ $\mathbf{vec}(XY) = (Y' \otimes I)\,\mathbf{vec}(X) = (I \otimes Y)\,\mathbf{vec}(X)$ e $\otimes$ es el producto de Kronecker. La matriz de varianzas covarianzas de $\mathbf{u}$ responderá a:

$$E\left[\mathbf{u}\mathbf{u}'\right] = \Omega_{\mathbf{u}} = I_T \otimes \Omega_\varepsilon.$$

Los estimadores MCO (bueno MCG) de $\boldsymbol{\beta}$ se pueden obtener de minimizar:

$$\begin{aligned}
S(\boldsymbol{\beta}) &= \mathbf{u}'\left(I_T \otimes \Omega_\varepsilon\right)^{-1}\mathbf{u} = \mathbf{u}'\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\mathbf{u} \\
&= \left(\mathbf{y} - \left(\mathsf{Z}' \otimes I_K\right)\boldsymbol{\beta}\right)'\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\left(\mathbf{y} - \left(\mathsf{Z}' \otimes I_K\right)\boldsymbol{\beta}\right) \\
&= \mathbf{vec}(\mathsf{Y} - \mathsf{BZ})'\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\mathbf{vec}(\mathsf{Y} - \mathsf{BZ}) \\
&= \mathbf{tr}\left[(\mathsf{Y} - \mathsf{BZ})'\,\Omega_\varepsilon^{-1}\,(\mathsf{Y} - \mathsf{BZ})\right],
\end{aligned}$$

donde utilizamos el hecho de que $(X \otimes Y)^{-1} = X^{-1} \otimes Y^{-1}$ y $\mathbf{tr}$ es el operador trza de una matrix (suma de los elementos de su diagonal principal). Que se puede reescribir:

$$
\begin{aligned}
S(\boldsymbol{\beta}) &= \mathbf{y}'\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\mathbf{y} \\
&\quad +\boldsymbol{\beta}'(\mathsf{Z} \otimes I_K)\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\left(\mathsf{Z}' \otimes I_K\right)\boldsymbol{\beta} \\
&\quad -2\boldsymbol{\beta}'(\mathsf{Z} \otimes I_K)\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\mathbf{y} \\
&= \mathbf{y}'\left(I_T \otimes \Omega_\varepsilon^{-1}\right)\mathbf{y} + \boldsymbol{\beta}'\left(\mathsf{ZZ}' \otimes \Omega_\varepsilon^{-1}\right)\boldsymbol{\beta} \\
&\quad -2\boldsymbol{\beta}'\left(\mathsf{Z} \otimes \Omega_\varepsilon^{-1}\right)\mathbf{y}
\end{aligned}
$$

Teniendo en cuenta $(X \otimes Y)' = X' \otimes Y'$ y que $(A \otimes B)(C \otimes D)$ $= AC \otimes BD$. Al calcular las derivadas tenemos:

$$
\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 2\left(\mathsf{ZZ}' \otimes \Omega_\varepsilon^{-1}\right)\boldsymbol{\beta} - \mathbf{2}\left(\mathsf{Z} \otimes \Omega_\varepsilon^{-1}\right)\mathbf{y} \tag{5}
$$

$$
\frac{\partial 2S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'} = 2\left(\mathsf{ZZ}' \otimes \Omega_\varepsilon^{-1}\right). \tag{6}
$$

Al igualar a zero (5) tenemos:

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \left(\mathsf{ZZ}' \otimes \Omega_{\varepsilon}^{-1}\right)\left(\mathsf{Z} \otimes \Omega_{\varepsilon}^{-1}\right)\mathbf{y} \\ &= \left(\left(\mathsf{ZZ}'\right)^{-1} \otimes \Omega\right)\left(\mathsf{Z} \otimes \Omega_{\varepsilon}^{-1}\right)\mathbf{y} \\ &= \left(\left(\mathsf{ZZ}'\right)^{-1}\mathsf{Z} \otimes I_k\right)\mathbf{y}.\end{aligned} \tag{7}$$

Como (6) es una matriz definida positiva podemos concluir que $\hat{\boldsymbol{\beta}}$ es un mínimo. Otra forma de obtener el estimador es escribiendo el modelo como:

$$\boldsymbol{X}_t = \mathsf{B}Z_{t-1} + E_t$$

y postmultiplicando por Z'_{t-1} y tomando esperanzas:

$$\begin{aligned}E\left[\boldsymbol{X}_t Z'_{t-1}\right] &= \mathsf{B}E\left[Z_{t-1}Z'_{t-1}\right] + E\left[E_t Z'_{t-1}\right] \\ E\left[\boldsymbol{X}_t Z'_{t-1}\right] &= \mathsf{B}E\left[Z_{t-1}Z'_{t-1}\right]\end{aligned}$$

Si estimamos los valores esperados $E\left[\boldsymbol{X}_t Z'_{t-1}\right]$ y $E\left[Z_{t-1}Z'_{t-1}\right]$ con:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} Z_{t-1}Z'_{t-1} &= \frac{1}{T}\mathsf{ZZ}' \\ \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{X}_t Z'_{t-1} &= \frac{1}{T}\mathsf{YZ}'\end{aligned}$$

tenemos:

$$
\begin{aligned}
\frac{1}{T} &= \hat{\mathsf{B}}\frac{1}{T}\mathsf{ZZ}' \\
\hat{\mathsf{B}} &= \mathsf{YZ}'\left(\mathsf{ZZ}'\right)^{-1}. \qquad (8)
\end{aligned}
$$

Dado que $E\left[E_t E_t'\right] = \Omega_\varepsilon$ podemos proponer como estimador de Ω_ε:

$$
\begin{aligned}
\tilde{\Omega}_\varepsilon &= \frac{1}{T}\sum_{t=1}^{T}\hat{E}_t\hat{E}_t' = \frac{1}{T}\hat{\mathsf{U}}\hat{\mathsf{U}}' \\
&= \frac{1}{T}\left(\mathsf{Y}-\hat{\mathsf{B}}\mathsf{Z}\right)\left(\mathsf{Y}-\hat{\mathsf{B}}\mathsf{Z}\right)' \\
&= \frac{1}{T}\left(\mathsf{Y}-\mathsf{YZ}'\left(\mathsf{ZZ}'\right)^{-1}\mathsf{Z}\right)\left(\mathsf{Y}-\mathsf{YZ}'\left(\mathsf{ZZ}'\right)^{-1}\mathsf{Z}\right)' \\
&= \frac{1}{T}\mathsf{Y}\left[I-\mathsf{Z}'\left(\mathsf{ZZ}'\right)^{-1}\mathsf{Z}\right]\left[I-\mathsf{Z}'\left(\mathsf{ZZ}'\right)^{-1}\mathsf{Z}\right]'\mathsf{Y}' \\
& \quad \frac{1}{T}\mathsf{Y}\left[I-\mathsf{Z}'\left(\mathsf{ZZ}'\right)^{-1}\mathsf{Z}\right]\mathsf{Y}'
\end{aligned}
$$

pero como dara lugar a un estimador sesgado utilizaremos:

$$
\hat{\Omega}_\varepsilon = \frac{T}{T-kn-1}\tilde{\Omega}_\varepsilon
$$

Es importante tener presente que: $\hat{\Omega}_\varepsilon$ y $\tilde{\Omega}_\varepsilon$ son asintóticamente equivalentes.

También se puede mostrar que los estimadores MV de B son idénticos a los estimadores MCO y que los estimadores MCO $\hat{\mathsf{B}}$ son estimadores consitentes de los parámetros de interés B y que tienen distribución:

$$\begin{aligned}\sqrt{T}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) &= \sqrt{T}\mathbf{vec}\left(\hat{\mathsf{B}}-\mathsf{B}\right)\xrightarrow{d} N\left(\mathbf{0},\Gamma^{-1}\otimes\Omega_{\varepsilon}\right)\\ \Gamma &= p\lim \mathsf{ZZ}'/T.\end{aligned}$$

Ver Lütkepohl (2007). Cada ecuación del VAR puede ser estimada por MCO de forma individual y se obtienen resultados equivalentes tal como demostro Zellner.

Para determinar el orden del VAR, es decir el número de retardos de X_t en (3) se utilizan los criterios de información de Akaike (AIC)), Hannan y Quinn (HQ) y de Schwarz (SC) y también el "Final prediction error"

(FPE) de Lütkepohl:

$$\begin{aligned} AIC(k) &= \ln\left|\tilde{\Omega}_{\varepsilon}\right| + \frac{2}{T}kn^2 \\ HQ(k) &= \ln\left|\tilde{\Omega}_{\varepsilon}\right| + \frac{2\ln(\ln(T))}{T}kn^2 \\ SC(k) &= \ln\left|\tilde{\Omega}_{\varepsilon}\right| + \frac{2\ln(T)}{T}kn^2 \\ FPE(k) &= \left(\frac{T+p^*}{T-p^*}\right)^n \left|\tilde{\Omega}_{\varepsilon}\right| \end{aligned}$$

donde p^* es el número total de parametros estimados en cada ecuación del VAR(k). Normalmente se establece un order máximo del VAR k^* y se estima del modelo VAR para $k = 1, 2, \ldots, k^*$ y se selecciona como orden adecuado para el VAR aquel retardo para el que alcancemos un valor más pequeño del criterio de información utilizado.

4 Diagnóstico del modelo VAR

Para que el modelo esté bien especificado, es importante comprobar que los residuos del modelo cumplen las hipótesis básicas. Por lo tanto debemos comprobar que no hay correlación serial, ausencia de Heteroscedasticidad y Normalidad de los residuos.

Para comprobar la ausencia de correlación serial en los residuos de un VAR(k) existén varias posibilidades desde un punto de vista multivariante Existe un test tipo Q equivalente al test Portmanteu para series temporales univariantes de Box-Pierce y Ljung-Box definido como:

$$
\begin{aligned}
Q_h &= T \sum_{i=1}^{h} \mathbf{tr}\left(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}\right) \sim \chi^2_{n^2(h-k)} \\
\hat{C}_i &= \frac{1}{T} \sum_{t=i+1}^{T} \hat{E}_t \hat{E}_{t-i}'
\end{aligned}
$$

La hipótesis nula (H_0) consisten en la ausencia de correlación serial hasta el orden h y rechazamos H_0 si el estadistico de prueba Q_h es mayor que le valor de tablas de una $\chi^2_{n^2(h-k)}$ al nivel de signficación que se desee trabajar. Existe una corrección del anterior contrate para muestras pequeñas que responde a:

$$Q_h^* = T^2 \sum_{i=1}^{h} \frac{1}{T-i} \mathbf{tr}\left(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}\right) \sim \chi^2_{n^2(h-k)}$$

tal y como se puede apreciar ambos contrastes son equivalentes a nivel asintótico.

Existe otro contraste a de Multiplicadores de Lagrange que sería la extensión del contraste de Breusch Godfrey a nivel del VAR que consiste en contrastar la siguiente hipotesis nula en el modelo:

$$\begin{aligned} H_0 &: B_1 = \cdots = B_h = 0 \\ \hat{E}_t &= \mu + \Pi_1 X_{t-1} + \cdots + \Pi_k X_{t-k} + B_1 \hat{E}_{t-1} + \cdots \quad (9) \\ &+ B_h \hat{E}_{t-h} + V_t \quad (10) \\ LM_h &= T\left(n - \mathbf{tr}\left(\tilde{\Omega}_R^{-1} \tilde{\Omega}_v\right)\right) \sim \chi^2_{n^2 h} \end{aligned}$$

donde $\tilde{\Omega}_R$ y $\tilde{\Omega}_v$ son los estimadores de la matriz de varianzas y covarianzas residual de (9) bajo H_0 y sin restringir respectivamente. Al igual que el caso anterior existe una propuesta de correción del anterior contraste LM_h para muestras pequeñas:

$$\begin{aligned}
FLM_h &= \frac{1-\left(1-R^2\right)^{1/r}}{\left(1-R^2\right)^{1/r}}\frac{Nr-q}{nm} \sim F_{n^2h,[Nr-q]} \\
R^2 &= 1-\left|\tilde{\Omega}_v\right| / \left|\tilde{\Omega}_R\right| \\
m &= nh \\
r &= \left(\left(n^2m^2-4\right) / \left(n^2+m^2-5\right)\right)^{1/2} \\
N &= T-n-m-1/2\,(n-m+1)
\end{aligned}$$

Donde $[A]$ es la parte entera de A.

En cuanto a los contraste de Heteroscedasticidad, existe un contraste de multivariante para comprobar si los errores se comportan como un proceso ARCH que se basa en la siguiente modelo:

$$\begin{aligned}
H_0 &: B_1 = \cdots = B_q = 0 \\
\mathbf{vech}\left(\hat{E}_t \hat{E}_t'\right) &= B_0 + B_1 \mathbf{vech}\left(\hat{E}_{t-1}\hat{E}_{t-1}'\right) + \cdots \\
&\quad + B_q \mathbf{vech}\left(\hat{E}_{t-q}\hat{E}_{t-q}'\right) + V_t \\
VARCH_{LM_q} &= \frac{Tn\,(n-1)}{2} R_m^2 \sim \chi^2_{qn^2(n+1)^2/4} \\
R_m^2 &= 1 - \frac{2\mathbf{tr}\left(\tilde{\Omega}\tilde{\Omega}_0^{-1}\right)}{n\,(n+1)}
\end{aligned}$$

donde $\mathbf{vech}$ es similar al operador $\mathbf{vec}$ pero en lugar de vectorizar siguiendo el orden de las columnas vectoriza siguiendo el orden de las filas. Finalmente se dispone de la extensión multivariante de contraste de Jarque-Bera para contrastar la Normalidad de los residuos el contraste consiste en:

$$
\begin{aligned}
JB_{mv} &= s_3^2 + s_4^2 \sim \chi^2_{2n} \\
s_3^2 &= T\mathbf{b}_1'\mathbf{b}_1/6 \\
s_4^2 &= T\left(\mathbf{b}_2 - 3n\right)'\left(\mathbf{b}_2 - 3n\right)/24 \\
\mathbf{b}_1 &= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \end{bmatrix}' \\
\mathbf{b}_2 &= \begin{bmatrix} b_{21} & b_{22} & \cdots & b_{2n} \end{bmatrix}' \\
b_{1j} &= \frac{1}{T}\sum_t \hat{w}_{jt}^3 \\
b_{2j} &= \frac{1}{T}\sum_t \hat{w}_{jt}^4
\end{aligned}
$$

Donde $\mathbf{b}_1$ y $\mathbf{b}_2$ son los vectores que recogen los momentos de tercer y cuarto orden de los residuos del modelo VAR previamente estandarizados $\hat{w}_{jt}$. La hipótesis nula del contraste JB_{mv} cosnsite en la normalidad de los residuos y el estadistico de prueba es mayor que el valor de tablas de una χ^2_{2n} al nivel de significación que se trabaje rechazaremos la hipótesis nula de normalidad.

Finalmente, también se pueden aplicar los contrastes típicos de series temporales a los residuos de cada una de las ecuaciones del VAR. Se pueden hacer un histograma de los residuos, obtener la FAS y la FAP de los residuos y de los residuos al cuadrado.

5 Función de Respuesta a un impulso

El modelo VAR puede ser expresado en desviaciones con respecto a la media:

$$\begin{aligned}(X_t - \mu) &= \Pi_1 (X_{t-1} - \mu) + \Pi_2 (X_{t-2} - \mu) + \cdots + \\ &\quad + \Pi_k (X_{t-k} - \mu) + E_t.\end{aligned}$$

Si por simplicidad asumimos que el vector X_t está expresado en desviaciones con respecto a la media y también asumimos que los condiciones iniciales son iguales a cero ($X_0 = 0$). Es fácil ver que podemos expresar el modelo

VAR como un modelo Vectorial de medias móviles infinito VMA(∞):

$$
\begin{aligned}
\boldsymbol{\Pi}_k(L) X_t &= E_t \\
X_t &= \boldsymbol{\Pi}_k(L)^{-1} E_t = \boldsymbol{\Psi}_\infty(L) E_t.
\end{aligned}
$$

Donde el polinomio de medias móviles $\boldsymbol{\Psi}_\infty(L)$ resultante de invertir el polinomio de matrices $\boldsymbol{\Pi}_k(L)$ será un polinomio de matrices de orden infinito, cuyos elementos (matrices) podrán ser obtenidos a partir de:

$$
\begin{aligned}
\boldsymbol{\Pi}_k(L)^{-1} &= \boldsymbol{\Psi}_\infty(L) \\
I &= \boldsymbol{\Pi}_k(L) \boldsymbol{\Psi}_\infty(L) \\
I &= \left(I_k - \boldsymbol{\Pi}_1 L - \boldsymbol{\Pi}_2 L^2 - \cdots - \boldsymbol{\Pi}_k L^k\right) \times \\
&\quad \times \left(I_k + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots\right).
\end{aligned}
$$

Realizando el producto y agrupando términos en L^j obtenemos:

$$
\begin{aligned}
I_k &= \left(I_k + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots\right) - \\
&\quad -\boldsymbol{\Pi}_1 \left(I_k + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots\right) L - \\
&\quad -\boldsymbol{\Pi}_2 \left(I_k + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots\right) L^2 - \cdots - \\
&\quad -\boldsymbol{\Pi}_k \left(I_k + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots\right) L^k
\end{aligned}
$$

$$
\begin{aligned}
I_k &= I_k + (\Psi_1 - \Pi_1)\, L + (\Psi_2 - \Pi_1\Psi_1 - \Pi_2)\, L^2 + \\
&\quad + (\Psi_3 - \Pi_1\Psi_2 - \Pi_2\Psi_1 - \Pi_2)\, L^2 + \cdots
\end{aligned}
$$

Igualando términos asociados al operador de retardos con el mismo exponente L^j en ambos lados de la igualdad, obtenemos la forma de calcular las matrices de la representación VMA:

$$
\begin{aligned}
L &: \quad 0 = \Psi_1 - \Pi_1 \quad \Psi_1 = \Pi_1 \\
L^2 &: \quad 0 = \Psi_2 - \Pi_1\Psi_1 - \Pi_2 \quad \Psi_2 = \Pi_1\Psi_1 - \Pi_2 \\
L^3 &: \quad 0 = \Psi_3 - \Pi_1\Psi_2 - \Pi_2\Psi_1 - \Pi_2 \\
\Psi_3 &= \Pi_1\Psi_2 + \Pi_2\Psi_1 + \Pi_2 \\
&\vdots \\
L^j &: \quad 0 = \Psi_j - \Pi_1\Psi_{j-1} - \Pi_2\Psi_{j-2} - \cdots - \Pi_k\Psi_{j-} \\
\Psi_j &= \Pi_1\Psi_{j-1} + \Pi_2\Psi_{j-2} + \cdots + \Pi_k\Psi_{j-k} \\
\text{con} \quad \Psi_0 &= I \quad \Psi_s = 0 \quad s < 0.
\end{aligned}
$$

La representación del modelo VAR(k) en términos de un modelo VMA(∞) nos permite calcular la **función de respuesta a un impulso**.

En un proceso AR(p) estacionario al expresarlo como un proceso MA(∞) podemos analizar como afectan a la variable analizada a medida que pasa el tiempo los shocks (ε_t) a los que se ve expuesta:

$$
\begin{aligned}
x_t &= \phi x_{t-1} + \varepsilon_t \quad (1 - \phi L)\, x_t = \varepsilon_t \\
x_t &= \frac{\varepsilon_t}{(1 - \phi L)} \\
x_t &= \varepsilon_t + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \phi^3 \varepsilon_{t-3} + \cdots
\end{aligned}
$$

Este mismo análisis se puede llevar a cabo para un VAR, pero en este último caso los shocks se transmitirán de unas variables a otras:

$$
\begin{aligned}
X_t &= \boldsymbol{\Psi}_\infty (L)\, E_t = \left(I_k + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \cdots\right)(E_t) \\
&= E_t + \boldsymbol{\Psi}_1 E_{t-1} + \boldsymbol{\Psi}_2 E_{t-2} + \boldsymbol{\Psi}_3 E_{t-3} + \cdots
\end{aligned}
$$

Siendo fácil ver que la interpretación de $\boldsymbol{\Psi}_j$ responde a:

$$
\frac{\partial X_{t+j}}{\partial E_t} = \boldsymbol{\Psi}_j .
$$

Por lo tanto los elementos de las matrices $\boldsymbol{\Psi}_j$ en (11) van recogiendo ecomo se van propagando a lo largo del tiempo una variannción unitaria en en cada uno de los socks recogidos en E_t sobre las variables X_t. Por ejemplo en un VAR con tres variables $n = 3$ tendriamos:

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Psi}_{11}^1 & \boldsymbol{\Psi}_{12}^1 & \boldsymbol{\Psi}_{13}^1 \\ \boldsymbol{\Psi}_{21}^1 & \boldsymbol{\Psi}_{22}^1 & \boldsymbol{\Psi}_{23}^1 \\ \boldsymbol{\Psi}_{31}^1 & \boldsymbol{\Psi}_{32}^1 & \boldsymbol{\Psi}_{33}^1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \\ \varepsilon_{3t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Psi}_{11}^2 & \boldsymbol{\Psi}_{12}^2 & \boldsymbol{\Psi}_{13}^2 \\ \boldsymbol{\Psi}_{21}^2 & \boldsymbol{\Psi}_{22}^2 & \boldsymbol{\Psi}_{23}^2 \\ \boldsymbol{\Psi}_{31}^2 & \boldsymbol{\Psi}_{32}^2 & \boldsymbol{\Psi}_{33}^2 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-2} \\ \varepsilon_{2t-2} \\ \varepsilon_{3t-2} \end{bmatrix} + \cdots$$

donde, $\boldsymbol{\Psi}_{11}^1$, $\boldsymbol{\Psi}_{12}^1$ y $\boldsymbol{\Psi}_{13}^1$ recoge los efectos de una variación unitaria de la innivación asociada x_{1t} transcurrido un periodo sobre x_{1t}, x_{2t} y x_{3t} respectivamente de forma equivalente $\boldsymbol{\Psi}_{31}^2$, $\boldsymbol{\Psi}_{32}^2$ y $\boldsymbol{\Psi}_{33}^2$ recogen el efecto de una variación unitaria de la innivación asociada a x_{3t} sobre x_{1t}, x_{2t} y x_{3t} una vez que han trascurrido dos periodos respectivamente.

El anterior análisis no tiene en cuenta que las innovaciones del modelo VAR pueden pueden estar relacionadas entre si a traves de la matrix de varianzas y covarianzas $E\left[E_t E_t'\right] = \Omega_\varepsilon$ que no tiene porque se diagonal, es decir, con todos elementos que no están en la diagonal principal distintos iguales a cero, si por ejemplo en el caso anterior tuviesemos:

$$\Omega_\varepsilon = \left[\begin{array}{ccc} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{array}\right]$$

Todas las innovaciones estarian relacionadas entre si comteporaneamente. Es decir si el efecto de un scock en ε_{1t} se transmitiria tambien a las otra variables x_{2t} y x_{3t} atraves de ε_{2t} y ε_{3t}, dado que, no son independientes entre si. Para evitar esto problemas se análisis de como se transmiten los efectos de los socks a lo largo del tiempo se hace utilizando la representación VMA(∞), pero utilzando:

$$
\begin{aligned}
X_t &= \Theta_\infty (L) \Xi_t = \left(\Theta_0 + \Theta_1 L + \Theta_2 L^2 + \cdots\right) \Xi_t \\
&= \Theta_0 \Xi_t + \Theta_1 \Xi_{t-1} + \Theta_2 \Xi_{t-2} + \Theta_3 \Xi_{t-3} + \cdots \qquad (12) \\
&= P P^{-1} E_t + \Psi_1 P P^{-1} E_{t-1} + \Psi_2 P P^{-1} E_{t-2} \\
&\quad + \Psi_3 P P^{-1} E_{t-3} + \cdots \\
\Omega_\varepsilon &= P P' \\
\Theta_0 &= P \\
\Theta_i &= P \Theta_i \quad i = 1, 2, \ldots \\
\Xi_t &= P^{-1} E_t
\end{aligned}
$$

De forma que la matriz de varianzas y covarianzas de Ξ_t pasa a ser la matriz indentidad, ya $\Omega_\varepsilon = PP'$ es la descomposición de Cholesky de la matriz Ω_ε donde P es una matriz cuadrada de orden $n \times n$ que es triangular inferior, se cumple que:

$$
\begin{aligned}
E\left[\Xi_t \Xi_t'\right] &= P^{-1} E\left[E_t E_t'\right] \left(P^{-1}\right)' = P^{-1} \Omega_\varepsilon \left(P^{-1}\right)' \\
&= P^{-1} P P' \left(P^{-1}\right)' = I_n.
\end{aligned}
$$

Finalmente el hecho de que $\Theta_0 = P$ sea diagonal inferior, introduce una importante limitación en el análisis dado que el orden es importante a la hora es especficar el VAR, por que la primera variable en VAR es capaz de transmitir socks a las demas variables, y la segunda a todas menos a la primera y asi sucesivamente.

6 Forecast Error Variance Decomposition

Si damos por buena la representación VMA(∞) en podems obtener el error de predicción de h periodos hacia adelante como:

$$
\begin{aligned}
X_{t+h} &= \Theta_0 \Xi_{t+h} + \Theta_1 \Xi_{t+h-1} + \Theta_2 \Xi_{t+h-2} + \Theta_3 \\
\hat{X}_t(h) &= E\left[X_{t+h} | I_h\right] = \Theta_h \Xi_t + \Theta_{h+1} \Xi_{t-1} + \Theta_{h+} \\
X_{t+h} - \hat{X}_t(h) &= \Theta_0 \Xi_{t+h} + \Theta_1 \Xi_{t+h-1} + \cdots \Theta_{h-1} \Xi_{t+1} \\
&= \sum_{j=0}^{h-1} \Theta_j \Xi_{t+h-j}
\end{aligned}
$$

En el caso de una de las variables del VAR tendriamos:

$$\begin{aligned} x_{jt+h} - \hat{x}_{jt+h}(h) &= \sum_{i=0}^{h-1} \left(\Theta_{j1}^{i} \omega_{1,t+h-i} + \cdots + \Theta_{jn}^{i} \omega_{nt+h-} \right. \\ &= \sum_{k=1}^{n} \left(\Theta_{jk}^{0} \omega_{k,t+h} + \cdots + \Theta_{jk}^{h-1} \omega_{k,t+1} \right) \end{aligned}$$

donde $\omega_{k,t+h}$ y Θ_{jk}^{h} son los elementos de Ξ_{t+h} y Θ_{h} respectivamente. Como Ξ_{t} tiene matriz de varianzas y covarianzas igual a la matriz identidad las varianza de todos sus elementos son iguales a uno y sus covarianzas iguales a zero, luego son independientes por tanto:

$$VAR\left[x_{jt+h} - \hat{x}_{jt+h}(h)\right] = \sum_{k=1}^{n} \left(\left(\Theta_{jk}^{0}\right)^{2} + \cdots + \left(\Theta_{jk}^{h-1}\right)^{2} \right)$$

La contribución de la innovacion asociada a x_{jt} al total de la varianza serà $\left(\Theta_{jj}^{0}\right)^{2} + \cdots + \left(\Theta_{jj}^{h-1}\right)^{2}$ o de innivación asociada a otra variable x_{ht} será $\left(\Theta_{jh}^{0}\right)^{2} + \cdots + \left(\Theta_{jh}^{h-1}\right)^{2}$

en porcentaje sería:

$$
\begin{aligned}
C_{jj}(h) &= \frac{\left(\Theta_{jj}^{0}\right)^{2}+\cdots+\left(\Theta_{jj}^{h-1}\right)^{2}}{VAR\left[x_{jt+h}-\hat{x}_{jt+h}(h)\right]} \\
C_{jh}(h) &= \frac{\left(\Theta_{jh}^{0}\right)^{2}+\cdots+\left(\Theta_{jh}^{h-1}\right)^{2}}{VAR\left[x_{jt+h}-\hat{x}_{jt+h}(h)\right]}
\end{aligned}
$$

Que recogen la contribución de la innovacion asociada a la variable x_{jt} sobre esa mismas variable en el caso de $C_{jj}(h)$. Y recoge la la contribución de la innovacion asociada a la variable x_{ht} sobre x_{jt}.

7 Causalidad tipo Granger o el en sentido de Granger

El concepto de causalidad tipo Granger es fácil de implementar mediante los modelos VAR. Este concepto se basa en los dos siguientes supuestos:

- El futuro no puede causar el pasado, es decir, la causalidad estricta sólo se puede dar cuando el pasado puede causar el presente o el futuro.

- Una causa contiene información única que no está disponible en otro lado.

Una variable x_t no está causada en sentido Granger por otra variable y_t si y solo si:

$$
\begin{aligned}
F_x\left(x_t|X_{t-1},Y_{t-1}\right) &= F_x\left(x_t|X_{t-1}\right)\\
X_{t-1} &= \left[x_{t-1},x_{t-2},\cdots,x_1\right]\\
Y_{t-1} &= \left[y_{t-1},y_{t-2},\cdots,y_1\right].
\end{aligned}
$$

En términos de un VAR bi-variante (por ejemplo de orden 2):

$$
\begin{aligned}
\begin{bmatrix} x_t \\ y_t \end{bmatrix} &= \begin{bmatrix} \pi^1_{11} & \pi^1_{12} \\ \pi^1_{21} & \pi^1_{22} \end{bmatrix} \times \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \\
&\quad + \begin{bmatrix} \pi^2_{11} & \pi^2_{12} \\ \pi^2_{21} & \pi^2_{22} \end{bmatrix} \times \begin{bmatrix} x_{t-2} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix}
\end{aligned}
$$

x_t no estará causada en sentido Granger por y_t si $\pi_{12}^1 = 0$ y $\pi_{12}^2 = 0$.

y_t no estará causada en sentido Granger por x_t si $\pi_{21}^1 = 0$ y $\pi_{21}^2 = 0$.

Entonces basta con contrastar las hipótesis $\pi_{12}^1 = \pi_{12}^2 = 0$ y $\pi_{21}^1 = \pi_{21}^2 = 0$ en el anterior modelo.