Modelos vectoriales autorregresivos

Tere García Muñoz (tgarciam@ugr.es)
Román Salmerón Gómez (romansg@ugr.es)
Alessio Gaggero (alessiogaggero@ugr.es)

Universidad de Granada
Departamento de Métodos Cuantitativos para la Economía y la Empresa

Tema 6– Econometría 3
Grado en Económicas

Los modelos VAR:

- analizan conjuntamente dos o más series temporales
- son muy útiles cuando existe evidencia de causalidad simultánea entre un grupo de variables y sus relaciones se transmiten a lo largo de un determinado número de períodos.

Modelo VAR con p retardos para dos variables:

$$\begin{aligned} X_t &= \delta_1 + \alpha_{11} X_{t-1} + \cdots + \alpha_{1p} X_{t-p} + \beta_{11} Y_{t-1} + \cdots + \beta_{1p} Y_{t-p} + a_{1t}, \\ Y_t &= \delta_2 + \alpha_{21} X_{t-1} + \cdots + \alpha_{2p} X_{t-p} + \beta_{21} Y_{t-1} + \cdots + \beta_{2p} Y_{t-p} + a_{2t}, \end{aligned}$$

donde

- δ_i representa el término independiente de la ecuación i, con $i = 1, 2$
- α_{ij} y β_{ij} representan los coeficientes de los procesos X e Y, respectivamente, en la ecuación i y el retardo j, con $j = 1, \ldots, p$
- a_{1t} y a_{2t} son dos series independientes e idénticamente distribuidas con media cero.

Así, por ejemplo, si $p = 1$ se tiene:

$$\begin{aligned} X_t &= \delta_1 + \alpha_{11} X_{t-1} + \beta_{11} Y_{t-1} + a_{1t}, \\ Y_t &= \delta_2 + \alpha_{21} X_{t-1} + \beta_{21} Y_{t-1} + a_{2t}. \end{aligned}$$

Expresado matricialmente:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \beta_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \cdot \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}$$

- α_{11}, muestra la dependencia lineal de X_t con X_{t-1} en presencia de Y_{t-1}.
- β_{11}, muestra la dependencia lineal de X_t con Y_{t-1} en presencia de X_{t-1}.
- α_{21}, mide la relación lineal entre Y_t y X_{t-1} en presencia de Y_{t-1}.
- β_{21}, mide la relación lineal entre Y_t y Y_{t-1} en presencia de X_{t-1}.

Caso 1. $\beta_{11} = \alpha_{21} = 0$.

$$\begin{aligned} X_t &= \delta_1 + \alpha_{11} X_{t-1} + a_{1t}, \\ Y_t &= \delta_2 + \beta_{21} Y_{t-1} + a_{2t}. \end{aligned}$$

- X_t e Y_t no están dinámicamente relacionados.
- Cada serie sigue un modelo univariante *AR(1)* y se pueden tratar como tales de acuerdo a lo estudiado en los tres primeros capítulos.
- Se dice que las dos series están desacopladas.

Caso 2. $\beta_{11} = 0$ pero $\alpha_{21} \neq 0$

$$
\begin{aligned}
X_t &= \delta_1 + \alpha_{11} X_{t-1} + a_{1t}, \\
Y_t &= \delta_2 + \alpha_{21} X_{t-1} + \beta_{21} Y_{t-1} + a_{2t}.
\end{aligned}
$$

- X_t no depende del valor pasado Y_{t-1}, pero Y_t si depende del valor pasado X_{t-1}. Relación unidireccional con X_t actuando como una variable *input* y Y_t como una variable *output*. X_t y Y_t tienen una relación de función de transferencia.
- El modelo implica la existencia de causalidad en sentido de Granger entre las dos series con X_t causando Y_t, pero no X_t siendo causada por Y_t.
- La información pasada X_{t-1} mejora la predicción de Y_t, reduciendo la varianza del error de predicción.

De forma análoga, si $\alpha_{21} = 0$ pero $\beta_{11} \neq 0$, entonces Y_t causa X_t pero X_t no causa Y_t.

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \beta_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \cdot \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}$$

Condición de estacionariedad: Las series de un modelo VAR(1) son estacionarias si todos los valores propios de la matriz $\begin{pmatrix} \alpha_{11} & \beta_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix}$ son menores que uno en valor absoluto.

$$\begin{aligned} X_t &= \delta_1 + \alpha_{11} X_{t-1} + \alpha_{12} X_{t-2} + \beta_{11} Y_{t-1} + \beta_{12} Y_{t-2} + a_{1t}, \\ Y_t &= \delta_2 + \alpha_{21} X_{t-1} + \alpha_{22} X_{t-2} + \beta_{21} Y_{t-1} + \beta_{22} Y_{t-2} + a_{2t}. \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} &= \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \beta_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \cdot \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} \\ &\quad + \begin{pmatrix} \alpha_{12} & \beta_{12} \\ \alpha_{22} & \beta_{22} \end{pmatrix} \cdot \begin{pmatrix} X_{t-2} \\ Y_{t-2} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix} \end{aligned}$$

- α_{11}, muestra la dependencia lineal de X_t con X_{t-1} en presencia de X_{t-2},Y_{t-1} e Y_{t-2}.
- β_{21}, mide la relación lineal entre Y_t y Y_{t-1} en presencia de X_{t-1}, X_{t-2} e Y_{t-2} .

$$\begin{aligned}
X_t &= \delta_1 + \alpha_{11} X_{t-1} + \alpha_{12} X_{t-2} + \beta_{11} Y_{t-1} + \beta_{12} Y_{t-2} + a_{1t}, \\
Y_t &= \delta_2 + \alpha_{21} X_{t-1} + \alpha_{22} X_{t-2} + \beta_{21} Y_{t-1} + \beta_{22} Y_{t-2} + a_{2t}.
\end{aligned}$$

$$\begin{pmatrix} X_t \\ Y_t \\ X_{t-1} \\ Y_{t-1} \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \beta_{11} & \alpha_{12} & \beta_{12} \\ \alpha_{21} & \beta_{21} & \alpha_{22} & \beta_{22} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} X_{t-1} \\ Y_{t-1} \\ X_{t-2} \\ Y_{t-2} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \\ 0 \\ 0 \end{pmatrix}$$

$$Z_t = \delta_t + \Phi Z_{t-1} + a_t$$

Dado el contraste:

$$\left.\begin{array}{l} H_0 : \text{el orden del VAR es } p_1 \\ H_1 : \text{el orden del VAR es } p_2 \end{array}\right\},$$

donde $p_1 < p_2$, se tiene que se rechaza la hipótesis nula si el valor experimental:

$$(T - m) \cdot \left(\ln \left(|\widehat{\Sigma}_{p_1}| \right) - \ln \left(|\widehat{\Sigma}_{p_2}| \right) \right),$$

es mayor que el valor teórico de una χ^2_r siendo r el número de parámetros de menos que hay que estimar en el modelo de la hipótesis nula frente al de la hipótesis alternativa, y donde:

- T corresponde al número de observaciones disponible,
- m es el número de parámetros a estimar bajo la hipótesis alternativa
- $|\widehat{\Sigma}_{p_1}|$ y $|\widehat{\Sigma}_{p_2}|$ son los determinantes de las estimaciones de las matrices de varianzas-covarianzas de los residuos de cada modelo.

La forma idónea de actuar es considerar inicialmente que $p_1 = 1$ y $p_2 = 2$ e ir aumentando gradualmente estos valores hasta que no se rechaze la hipótesis nula. En tal caso se tendría el orden idóneo para el modelo.

Para seleccionar el orden adecuado para el VAR es usar los distintos criterios de información: Akaike, Schwarz o Hannan-Quinn. Así, dado un criterio, se elegirá como orden idóneo del VAR el de aquel modelo que presente un menor valor.

Ejemplo. Determinar el orden del modelo VAR teniendo en cuenta:

Orden	AIC	BIC	HQC
1	2.0749	2.3043	2.1622
2	1.7702	2.1526	1.9158
3	1.8080	2.3434	2.0119
4	1.9331	2.6214	2.1952

Orden	AIC	BIC	HQC
1	2.0031	2.1621	2.0627
2	1.5355	1.8536	1.6547
3	1.5251	2.0021	1.7038
4	1.6646	2.3006	1.9028

El objetivo del **análisis de causalidad de Granger** es determinar si los retardos de una variable influyen en otra y, por tanto, son útiles para explicarla.

$$\begin{aligned} X_t &= \delta_1 + \alpha_{11} X_{t-1} + \cdots + \alpha_{1p} X_{t-p} + \beta_{11} Y_{t-1} + \cdots + \beta_{1p} Y_{t-p} + a_{1t}, \\ Y_t &= \delta_2 + \alpha_{21} X_{t-1} + \cdots + \alpha_{2p} X_{t-p} + \beta_{21} Y_{t-1} + \cdots + \beta_{2p} Y_{t-p} + a_{2t}, \end{aligned}$$

se tiene que:

- Si los coeficientes β_{1j} (primera ecuación) son significativamente distintos de cero, mientras que los α_{2j} (segunda ecuación) no lo son, $j = 1, \ldots, p$, se dice que hay causalidad en el sentido de Granger de Y hacia X.
- Si los coeficientes α_{2j} (segunda ecuación) son significativamente distintos de cero, mientras que los β_{1j} (primera ecuación) no lo son, $j = 1, \ldots, p$, se dice que hay causalidad en el sentido de Granger de X hacia Y.
- Si ambos conjuntos de parámetros son estadísticamente significativos (distintos de cero), entonces se dice que hay causalidad en el sentido de Granger bidireccional.
- Si ninguno de los dos conjuntos es estadísticamente significativo (distinto de cero), entonces se dice que no hay ninguna relación de causalidad. En este caso, se deberían analizar las series de forma individual siguiendo un planteamiento unidimensional.

Contrastes F-Snedecor:

$H_0 : \beta_{11} = \beta_{12} = \cdots = \beta_{1p} = 0$
H_1 : Alguno de esos parámetros es distinto de cero

$H_0 : \alpha_{21} = \alpha_{22} = \cdots = \alpha_{2p} = 0$
H_1 : Alguno de esos parámetros es distinto de cero

Ejemplo. Para un modelo VAR(2) se ha obtenido la siguiente información sobre los contrastes de restricciones cero. El número de observaciones utilizadas son 54. ¿Existe relación de causalidad en el sentido de Granger entre ambas variables?:

	Ecuación 1ª variable	Ecuación 2ª variable
Todos los retardos de 1ª variable tienen coeficientes nulos de forma simultánea	$F_{2,49} = 31.295$ p-valor <0.001	$F_{2,49} = 3.7821$ p-valor$=0.03$
Todos los retardos de 2ª variable tienen coeficientes nulos de forma simultánea	$F_{2,49} = 0.086$ p-valor$=0.9175$	$F_{2,49} = 48.405$ p-valor < 0.001

Puesto que los coeficientes del VAR son díficiles de interpretar individualmente, para analizar la respuesta de una variable dependiente se estudia la función de respuesta al impulso.

El objetivo del análisis de la respuesta a un impulso radica en estudiar cómo afecta al análisis realizado una distorsión en los datos. Es decir, calcular los efectos (incremento o disminución) a lo largo del tiempo ocasionados por una distorsión en las perturbaciones del sistema. La distorsión corresponde a un aumento de una desviación estándar en una de las perturbaciones del modelo. Es decir, la distorsión en la ecuación i consiste en la raíz cuadrada del elemento (i, i) de la diagonal principal de $\mathbf{\Sigma_a}$.

Ejemplo. Para el modelo *VAR(1)* para dos variables sin término independiente:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 1{,}168 & -38{,}031 \\ 0{,}0017 & 0{,}7158 \end{pmatrix} \cdot \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \widehat{a}_{1t} \\ \widehat{a}_{2t} \end{pmatrix},$$

con:

$$\widehat{\Sigma} = \begin{pmatrix} 63{,}947 & 0{,}2528 \\ 0{,}2528 & 0{,}001324 \end{pmatrix},$$

Se pide analizar el efecto que tiene una distorsión de una desviación estándar en la perturbación asociada a la primera variable X sobre X e Y en los tres primeros periodos:

Teniendo en cuenta que $\sqrt{63{,}947} = 7{,}9967$:

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} = \begin{pmatrix} 1{,}168 & -38{,}031 \\ 0{,}0017 & 0{,}7158 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 7{,}9967 \\ 0 \end{pmatrix} = \begin{pmatrix} 7{,}9967 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1{,}168 & -38{,}031 \\ 0{,}0017 & 0{,}7158 \end{pmatrix} \cdot \begin{pmatrix} 7{,}9967 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 9{,}3401 \\ 0{,}0136 \end{pmatrix},$$

$$\begin{pmatrix} X_3 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1{,}168 & -38{,}031 \\ 0{,}0017 & 0{,}7158 \end{pmatrix} \cdot \begin{pmatrix} 9{,}3401 \\ 0{,}0136 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 10{,}3923 \\ 0{,}0256 \end{pmatrix}.$$

Se puede observar que un incremento en X implica aumentos en X e Y.

Una de las ventajas de los modelos VAR sobre otro tipo de modelizaciones es su capacidad predictiva. A partir de la expresión :

$$\mathbf{Z}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{Z}_{t-1} + \mathbf{\Phi}_2 \mathbf{Z}_{t-2} + \cdots + \mathbf{\Phi}_p \mathbf{Z}_{t-p} + \mathbf{a}_t,$$

la predicción óptima de $\mathbf{Z}_t$ para k periodos futuros a partir de T observaciones vendrá dada por:

$$\hat{\mathbf{Z}}_T(k) = E[\mathbf{Z}_{T+k} | \mathbf{Z}_T, \ldots, \mathbf{Z}_1] \equiv E_T[\mathbf{Z}_{T+k}].$$

Recordad que:

$$E_T[\mathbf{a}_{T+s}] = 0, \ s > 0,$$

$$E_T[\mathbf{Z}_{T+s}] = \begin{cases} \hat{\mathbf{Z}}_T(s) & s > 0 \\ \mathbf{Z}_{T+s} & s \leq 0 \end{cases}.$$