UNIDAD III: VECTORES AUTOREGRESIVOS (MODELOS VAR)

DR. ROGER ALEJANDRO BANEGAS RIVERO, *PH.D.*

Modelos de Vectores Autoregresivos (VAR)

- Modelos autoregresivos es una propuesta generalizada por Sims (1980).

- Un modelo VAR es en escencia un sistema de modelos de regresión con variables predeterminadas. Ej. Existen más de una variable dependiente.

- El caso más simple es un VAR bivariado:

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \ldots + \beta_{1k} y_{1t-k} + \alpha_{11} y_{2t-1} + \ldots + \alpha_{1k} y_{2t-k} + u_{1t}$$
$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \ldots + \beta_{2k} y_{2t-k} + \alpha_{21} y_{1t-1} + \ldots + \alpha_{2k} y_{1t-k} + u_{2t}$$

Donde u_{it} está iid con E(u_{it})=0, i=1,2; E(u_{1t} u_{2t})=0.

- Este análisis se puede extender a un modelo VAR(g), de tal forma que existen g variables endógenas en g ecuaciones.

Modelos de Vectores Autoregresivos (VAR)
Notación y conceptos

- Un aspecto importante de modelos VARs es la abreviación de su notación. Por ejemplo, considere que existe un vector de variable endógenas con un rezago p=1.

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \alpha_{11} y_{2t-1} + u_{1t}$$
$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \alpha_{21} y_{1t-1} + u_{2t}$$

- Se reescribe como:

o

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

de una manera más compacta se rescribe como VAR(p) o VAR(1):

$$\underset{g\times 1}{y_t} = \underset{g\times 1}{\beta_0} + \underset{g\times g}{\beta_1}\ \underset{g\times 1}{y_{t-1}} + \underset{g\times 1}{u_t}$$

Modelos de Vectores Autoregresivos (VAR)
Notación y conceptos (cont'd)

- Este modelo se puede externder en *p* rezagos para cada variable incluída en cada ecuación:

$$\underset{g\times 1}{y_t} = \underset{g\times 1}{\beta_0} + \underset{g\times g}{\beta_1}\, \underset{g\times 1}{y_{t-1}} + \underset{g\times g}{\beta_2}\, \underset{g\times 1}{y_{t-2}} + \ldots + \underset{g\times g}{\beta_p}\, \underset{g\times 1}{y_{t-p}} + \underset{g\times 1}{u_t}$$

- Se puede extender este caso para un modelo con variables en diferencias (corto plazo) y relaciones de cointegración (largo plazo): el llamado vector con corrector de errores (VECM).

- Los modelos VAR están específicados con variables estacionarias.

Modelos VAR comparado con ecuaciones estructurales

- **Ventajas de los modelos VAR**

  - No necesita especificar cuales son endógenas o exógenas: todas son endógenas.

  - Tiene mayores ventajes que los modelos ARIMA: pueden intervenir otras variables exógenas.

  - Se asume que no existen efectos simultáneos: se puede aplicar de forma separada MCO para cada ecuación.

  - Los pronósticos de modelos VAR son frecuentemente mejores que los modelos estructurales.

- **Problemas con modelos VAR**

  - Modelos VARs son atéoricos (al igual que los modelos ARIMA)

  - ¿Cómo se decide el tamaño óptimo del rezago?

  - Muchos parámetros! Si se tienen g ecuaciones para g variables y se tienen p rezagos en cada ecuación, se tiene que estimar (g+ pg^2) parámetros. e.d. g=3, p=3, nùmero de parámetros = 30.

  - ¿Se debe asegurar que todos los componentes VAR son estacionarios?

  - ¿Còmo se interpretan los coeficientes?

Escogiendo el tamaño óptimo de rezago para un VAR

- 2 enfoques posibles: restricciones entre ecuaciones y criterios de informaciòn.

Restricciones entre ecuaciones.

- En el espírito de modelos VAR(irrestrictro) , cada ecuación debe tener el mismo tamaño de rezago.
- Suponga un modelo bivariado VAR(8) que se estima con base en 8 rezagos para cada ecuación, y se quiere examinar que 4 de 8 son cero de forma conjunta. Esto se realiza mediante la razón de verosimilitud (LR)
- Se denota una matriz residual de varianza - covarianza(dado por $\hat{u}\hat{u}'/T$), como $\hat{\Sigma}$. .La razón de verosimilitud se prueba por:

$$LR = T\left[\log\left|\hat{\Sigma}_r\right| - \log\left|\hat{\Sigma}_u\right|\right]$$

Escogiendo el tamaño óptimo de rezago para un VAR (cont'd)

- Donde $\hat{\Sigma}_r$ es la matriz residual de varianza-covarianza restricta (4 rezagos), $\hat{\Sigma}_u$ es la matriz residual de varianza-covarianza irrestricta (8 rezagos) y T es el tamaño de la muestra.

- La prueba estadística está asintóticamente distribuida con una χ^2 con grados de libertad igual al número total de restricciones.. En el modelo VAR se están Restringiendo 4 rezagos para dos variables en cada una de las dos ecuaciones: = un total de 4 *2 * 2 = 16 restricciones.

- En el caso general se tienen un VAR con g ecuaciones, y se quiere imponer la restricción que q rezagos son iguales a cero; en consecuencia, se tendría g^2q restricciones de forma conjunta.
- Desventajas: conduciendo la prueba LR es tediosa y requiere distribución normal en los residuos.

Criterios de información para el tamaño del rezago VAR

- Se aplican versiones multivariadas de criterios de información, lo cual se puede definir a través de:

$$MAIC = \ln \left|\hat{\Sigma}\right| + 2k'/T$$

$$MSBIC = \ln\left|\hat{\Sigma}\right| + \frac{k'}{T}\ln(T)$$

$$MHQIC = \ln\left|\hat{\Sigma}\right| + \frac{2k'}{T}\ln(\ln(T))$$

k' es el número total de regresores en todas las ecuaciones, que será igual a $g^2k + g$ para g ecuaciones, cada uno con k rezagos de las variables g, más ún termino constante en cada ecuación. Estos valores se construyen para 0, 1, … rezagos (se determina un número máximo rezago preestablecido $\bar{k}$).

¿Modelos VAR incluye términos contemporáneos?

- Si se asume un VAR de la forma:

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \alpha_{11} y_{2t-1} + u_{1t}$$
$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \alpha_{21} y_{1t-1} + u_{2t}$$

- ¿Qué sucede si se añaden términos simultáneos o contemporáneos?

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \alpha_{11} y_{2t-1} + \alpha_{12} y_{2t} + u_{1t}$$
$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \alpha_{21} y_{1t-1} + \alpha_{22} y_{1t} + u_{2t}$$

- Se puede escribir como:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \alpha_{12} & 0 \\ 0 & \alpha_{22} \end{pmatrix} \begin{pmatrix} y_{2t} \\ y_{1t} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

- Esta es la forma primitiva del modelo VAR.

VAR forma primitiva vs. VAR forma estándar

- Se puede ellevar a la parte izquierda:

$$\begin{pmatrix} 1 & -\alpha_{12} \\ -\alpha_{22} & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

o

$$\mathrm{B}\, y_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

- Se puede multiplicar a ambos lados por B^{-1} , lo cual implica:

$$y_t = \mathrm{B}^{-1}\beta_0 + \mathrm{B}^{-1}\beta_1 y_{t-1} + \mathrm{B}^{-1}u_t$$

o

$$y_t = \mathrm{A}_0 + \mathrm{A}_1 y_{t-1} + e_t$$

- Esta es la forma estándar VAR, que se puede estimar mediante MCO.

Significancia en bloque y prueba de causalidad

- Un modelo VAR incluye varios rezagos de las variables, es díficil distinguir cual es significativa y cuál no. Para ilustración considere un modelo bivarido VAR(3):

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \end{pmatrix} + \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \begin{pmatrix} y_{1t-3} \\ y_{2t-3} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

- Este VAR se puede expresar a través de ecuaciones individuales:

$$y_{1t} = \alpha_{10} + \beta_{11} y_{1t-1} + \beta_{12} y_{2t-1} + \gamma_{11} y_{1t-2} + \gamma_{12} y_{2t-2} + \delta_{11} y_{1t-3} + \delta_{12} y_{2t-3} + u_{1t}$$
$$y_{2t} = \alpha_{20} + \beta_{21} y_{1t-1} + \beta_{22} y_{2t-1} + \gamma_{21} y_{1t-2} + \gamma_{22} y_{2t-2} + \delta_{21} y_{1t-3} + \delta_{22} y_{2t-3} + u_{2t}$$

•

•Se está interesado en las pruebas de hipótesis en una forma matricial.

Significancia en bloque y prueba de causalidad(cont'd)

Hypothesis	Implied Restriction
1. Lags of y_{1t} do not explain current y_{2t}	$\beta_{21} = 0$ and $\gamma_{21} = 0$ and $\delta_{21} = 0$
2. Lags of y_{1t} do not explain current y_{1t}	$\beta_{11} = 0$ and $\gamma_{11} = 0$ and $\delta_{11} = 0$
3. Lags of y_{2t} do not explain current y_{1t}	$\beta_{12} = 0$ and $\gamma_{12} = 0$ and $\delta_{12} = 0$
4. Lags of y_{2t} do not explain current y_{2t}	$\beta_{22} = 0$ and $\gamma_{22} = 0$ and $\delta_{22} = 0$

- Se puede aplicar la prueba *F*-test para cada ecuación.
- Lo anterior es la llamada Causalidad de Granger.
- La causalidad de Granger busca las siguientes respuestas"¿Los cambios en y_1 causan cambios en y_2?" Si y_1 causa a y_2, los rezagos de y_1 deberían ser signficativos en y_2. En este caso, se dice que y_1 "causa Granger" y_2.
- Siy_2 causa y_1, los rezagos de y_2 deberían ser significativos para y_1.
- Si ambos rezagos son significativos, se dice que existe " una causalidad bi-direccional"

Funciones de impulso-respuesta

- Los modelos VAR son díficiles de interpretar: una solución es construir funciones de impulso-respuesta y descomposición de varianza
- Las funciones de impulso-respuesta explican la respuesta de cada variable dependiente del VAR a un *shock* en el término de error. Un *shock* único es aplicado a cada variable dependiente del modelo VAR.
- Considere un modelo bivariado VAR(1):

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \alpha_{11} y_{2t-1} + u_{1t}$$
$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \alpha_{21} y_{1t-1} + u_{2t}$$

- Un cambio en u_{1t} cambiará inmediatameente y_1, cambiará y_2 y también y_1 durante el período siguiente.
- Se puede examinar cuánto durará el impacto del *shock* en cada ecuación y sobre todas las variables del sistema.

Descomposición de la varianza

- La descomposición de varianza ofrece una manera diferente de examinar la dinámica VAR. Proporciona información de cambios en la variable dependiente debido a sus propios “*schoks*” versus “*shocks*” de otras variables.

- Sinónimos de *shocks*: choques, innovaciones, perturbaciones o impulsos.

- Lo anterior se consigue al determinar que tanto de los errores pronósticados de la varianza, en *s*-pasos hacia adelante, es explicada por las innovaciones de cada variable incluída en el sistema VAR:

- La descomposición de varianza brinda información acerca de la importancia realativa de cada *shock* entre las variables del modelo VAR.

Funciones de impulso-respuesta y la descomposición de varianza
La ordenación de las variables

- Para realizar funciones de impulso-respuesta y descomposición de la varianza, el orden en la ordenación de las variables es importante.

- Desde que los signos de los coeficientes pueden ser positvos o negativos, estadísticamente significativos o no, la interpretación de modelos VAR es díficil: ¿cómo reacciona una variable frente a cambios, innovaciones o choques de otras variables?
- Se utiliza el análisis de impulso-respuesta en s-pasos hacia adelante.
- En la ordenación de impulso-respuesta Cholesky, se asume que las primera variable es exógena y las que siguien son endógenas.

- La realización de impulso-respuesta mediante una específicación *generalizada* la ordenación de las variables no influye .

Funciones de impulso-respuesta y la descomposición de varianza
La ordenación de las variables

- Lo anterior se basa en que los errores del modelo VAR no están relacionados.
- Sin embargo, lo anterior no es cierto. Los errores están frecuentemente relacionados: tienen un mismo componente..
- Lo que se tiene que hacer es "ortogonalizar" las innovaciones.
- En un VAR bivariado, este problema se puede atribuir a patrones comunes en el modelo VAR.
- En el caso más general, la ortoganilización es más compleja pero la interpretación es la misma.