Modelado Estocástico

Clase 9

Prof. Fernando Grosz

fgrosz@udesa.edu.ar

Modelos VAR (*Vector Autoregressions*)

Hasta acá analizamos modelos autorregresivos univariados, como por ejemplo el proceso AR(*p*). Ahora vamos a considerar un vector columna de k variables diferentes, por ejemplo, el vector $\boldsymbol{y_t} = (y_{1t} \quad y_{2t} \quad y_{3t} \quad \cdots \quad y_{kt})'$ y queremos modelar este vector en términos de valores pasados del mismo. El resultado será un modelo de vectores autorregresivos o proceso VAR. Por ejemplo, un modelo VAR(*p*) tiene la siguiente estructura:

$$\boldsymbol{y_t} = \boldsymbol{m} + \boldsymbol{A_1} \boldsymbol{y_{t-1}} + \boldsymbol{A_2} \boldsymbol{y_{t-2}} + \cdots + \boldsymbol{A_p} \boldsymbol{y_{t-p}} + \boldsymbol{\varepsilon_t}$$

Cada $\boldsymbol{A_i}$ es una matriz de coeficientes de dimensión *k x k* y $\boldsymbol{m}$ es un vector de constantes de dimensión *k x* 1, y $\boldsymbol{\varepsilon_t}$ es un vector de ruidos blanco que satisface que $\mathrm{E}(\boldsymbol{\varepsilon_t}) = \mathbf{0}$ para todo *t* y además,

$$\mathrm{E}(\boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon_s}') = \begin{cases} \boldsymbol{\Omega} & si \quad t = s \\ \mathbf{0} & si \quad t \neq s \end{cases}$$

donde $\boldsymbol{\Omega}$ es una matriz de varianzas covarianzas y se supone definida positiva. Los $\boldsymbol{\varepsilon_t}$ se suponen serialmente no correlacionados pero contemporáneamente pueden estar correlacionados.

Ejemplo: el caso más simple es cuando $k = 2$ y $p = 1$. Nos queda

$$\boldsymbol{y_t} = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \qquad (*)$$

$$= \boldsymbol{m} + \boldsymbol{A} \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

Se puede escribir explícitamente como:

$$y_{1t} = m_1 + a_{11} y_{1t-1} + a_{12} y_{2t-1} + \varepsilon_{1t}$$
$$y_{2t} = m_2 + a_{21} y_{1t-1} + a_{22} y_{2t-1} + \varepsilon_{2t}$$

Cada variable se expresa como una combinación lineal del rezago de sí misma y de los rezagos de las demás. Como uno puede esperarse del caso univariado, el comportamiento de $\boldsymbol{y_t}$ va a depender de la matriz $\boldsymbol{A}$. Sean los autovalores y autovectores de la matriz $\boldsymbol{A}$,

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \boldsymbol{C} = \begin{pmatrix} \vdots & \vdots \\ \boldsymbol{c_1} & \boldsymbol{c_2} \\ \vdots & \vdots \end{pmatrix}$$

Si los autovalores son distintos, entonces los autovectores $\boldsymbol{c_1}$ y $\boldsymbol{c_2}$ van a ser linealmente independientes y $\boldsymbol{C}$ va a ser no singular:

$$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{\Lambda} \quad \text{y} \quad \boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^{-1}$$

Definimos un nuevo vector de variables, $\boldsymbol{z_t} = \boldsymbol{C}^{-1}\boldsymbol{y_t}$ de modo que $\boldsymbol{y_t} = \boldsymbol{C}\boldsymbol{z_t}$. Multiplicando $(*)$ por $\boldsymbol{C}^{-1}$obtenemos que

$$\boldsymbol{C}^{-1}\boldsymbol{y_t} = \boldsymbol{C}^{-1}\boldsymbol{m} + \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{y_{t-1}} + \boldsymbol{C}^{-1}\boldsymbol{\varepsilon_t}$$

$$\boldsymbol{z_t} = \boldsymbol{m}^* + \boldsymbol{\Lambda}\boldsymbol{z_{t-1}} + \boldsymbol{\eta_t}$$

donde $\boldsymbol{m}^* = \boldsymbol{C}^{-1}\boldsymbol{m}$ y donde $\boldsymbol{\eta_t} = \boldsymbol{C}^{-1}\boldsymbol{\varepsilon_t}$ (vector de ruidos blanco). De modo que,

$$z_{1t} = {m_1}^* + \lambda_1 z_{1t-1} + \eta_{1t}$$
$$z_{2t} = {m_2}^* + \lambda_2 z_{2t-1} + \eta_{2t}$$

z_{1t} y z_{2t} son procesos AR(1) univariados y separados y van a ser I(0) si el módulo de los autovalores λ_1 y λ_2 es menor a uno. Van a ser un *random walk* con *drift* cuando los autovalores sean iguales a uno y van a ser explosivos cuando los autovalores sean mayores a uno.

Este último caso vamos a descartarlo y vamos a concentrarnos en los otros dos.

Caso 1: $|\lambda_1| < 1 \ y \ |\lambda_2| < 1$. En este caso, los $\boldsymbol{z_t}$ son integrados de orden cero y como $\boldsymbol{y_t} = \boldsymbol{C z_t}$, los $\boldsymbol{y_t}$ son una combinación lineal de los $\boldsymbol{z_t}$ y por lo tanto también van a ser I(0). También se puede hablar de un equilibrio de largo plazo en este caso. Suponiendo que $\boldsymbol{y_t}$ toma un valor de largo plazo igual a $\overline{\boldsymbol{y}}$, de $(*)$ obtenemos que tomando esperanza que

$$(\boldsymbol{I} - \boldsymbol{A})\,\overline{\boldsymbol{y}} = \boldsymbol{m} \qquad \text{o sea,} \;\; \boldsymbol{\Pi}\overline{\boldsymbol{y}} = \boldsymbol{m} \;\; \text{donde } \boldsymbol{\Pi} = \boldsymbol{I} - \boldsymbol{A}$$

Va a haber una única solución para $\overline{\boldsymbol{y}}$ si la matriz $\boldsymbol{\Pi}$ es no singular.

Recordemos que:

- los autovalores μ_i de la matriz $\mathbf{\Pi}$ satisfacen que $\mu_i = 1 - \lambda_i$
- los autovectores de $\mathbf{\Pi}$ son iguales a los autovectores de $\boldsymbol{A}$

De manera que si $\mathbf{\Pi}$ es no singular, la solución de $\overline{\boldsymbol{y}}$ es única y consecuentemente, los autovalores nos aseguran que las desviaciones del valor de largo plazo de $\overline{\boldsymbol{y}}$ serán transitorias y siempre se retornará a la media.

Caso 2: $\lambda_1 = 1 \;\; y \;\; |\lambda_2| < 1$. En este caso, z_{1t} es un *random walk* con *drift* y z_{2t} es I(0). Entonces, como y_{1t} e y_{2t} son combinaciones lineales de los z_{1t} y z_{2t} tendremos que y_{1t} e y_{2t} serán ambos integrados de orden 1 ya que serán combinaciones lineales de un proceso I(1) y de un I(0). Por lo tanto, no va a tener sentido preguntarse si hay algún equilibrio estático de algún valor de $\bar{y}_1$ y de $\bar{y}_2$. Pero sí, va a tener sentido preguntarse si existe una relación de cointegración entre y_{1t} e y_{2t}.

La segunda fila de $\boldsymbol{z_t} = \boldsymbol{C}^{-1}\boldsymbol{y_t}$ nos queda $z_{2t} = \boldsymbol{c}^{(\mathbf{2})}\boldsymbol{y_t}$ donde $\boldsymbol{c}^{(\mathbf{2})}$ es la segunda fila de la matriz $\boldsymbol{C}^{-1}$. Lo interesante en este caso es que z_{2t} es una combinación lineal de dos variables I(1) (que son y_{1t} y y_{2t}) pero z_{2t} es I(0).

Para ver esto más explícitamente,

$$\boldsymbol{y_t} = \begin{bmatrix} \vdots \\ \boldsymbol{c_1} \\ \vdots \end{bmatrix} z_{1t} + \begin{bmatrix} \vdots \\ \boldsymbol{c_2} \\ \vdots \end{bmatrix} z_{2t}$$

Premultiplicando esta ecuación por el vector fila $\boldsymbol{c^{(2)}}$ nos queda

$$\boldsymbol{c^{(2)}}\boldsymbol{y_t} = z_{2t}$$

Por las propiedades de matrices no singulares: $\boldsymbol{c^{(2)}}\boldsymbol{c_1} = 0$ y $\boldsymbol{c^{(2)}}\boldsymbol{c_2} = 1$.

La relación de cointegración también puede escribirse en términos de la matriz $\mathbf{\Pi}$. Tomando diferencias de:

$$\boldsymbol{y_t} = \boldsymbol{m} + \boldsymbol{A}\boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

obtenemos

$$\Delta\boldsymbol{y_t} = \boldsymbol{m} - \mathbf{\Pi}\boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

Los autovalores de $\mathbf{\Pi}$ son 0 y $1 - \lambda_2$. Por lo tanto, $\mathbf{\Pi}$ es singular y tiene rango igual a uno. Como $\mathbf{\Pi}$ tiene los mismos autovectores que **A**,

$$\mathbf{\Pi} = \mathbf{C}\begin{pmatrix} 0 & 0 \\ 0 & 1-\lambda_2 \end{pmatrix}\boldsymbol{C}^{-1} = \begin{bmatrix} \vdots \\ \boldsymbol{c_2}(1-\lambda_2) \\ \vdots \end{bmatrix}\begin{bmatrix} \cdots & \boldsymbol{c^{(2)}} & \cdots \end{bmatrix}$$

O sea que a $\mathbf{\Pi}$, que tiene rango 1, la hemos factorizado en el producto de un vector columna por un vector fila. Esta factorización se llama "producto exterior" ("*outer product*" en inglés). El vector fila denota la relación de cointegración que recién vimos y el vector columna nos indica los pesos en la relación de cointegración. Es decir,

$$\Delta y_{1t} = m_1 - c_{21}(1-\lambda_2)z_{2t-1} + \varepsilon_{1t}$$
$$\Delta y_{2t} = m_2 - c_{22}(1-\lambda_2)z_{2t-1} + \varepsilon_{2t}$$

Esta última reformulación del VAR está expresada en términos de diferencias y niveles, donde todas estas son I(0). Esta formulación suele llamarse “**Vector Error Correction” (VEC)**, ya que z_{2t-1} mide en cuánto y_{1t} y y_{2t} se desvían de su relación de largo plazo.

Ejemplo (Caso 2): considere el sistema (hacemos $m_1 = m_2 = 0$)

$$y_{1t} = 1.2y_{1t-1} - 0.2y_{2t-1} + \varepsilon_{1t}$$

$$y_{2t} = 0.6y_{1t-1} + 0.4y_{2t-1} + \varepsilon_{2t}$$

Entonces, $\boldsymbol{A} = \begin{pmatrix} 1.2 & -0.2 \\ 0.6 & 0.4 \end{pmatrix}$. Los autovalores de A satisfacen:

$det\begin{pmatrix} 1.2-\lambda & -0.2 \\ 0.6 & 0.4-\lambda \end{pmatrix} = 0$. Resolviendo, $\lambda_1 = 1;\ \lambda_2 = 0.6$

Para encontrar el autovector correspondiente a $\lambda_1 = 1$,

$\begin{pmatrix} 1.2-1 & -0.2 \\ 0.6 & 0.4-1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \implies x_1 = x_2$ o sea, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ es un autovector asociado a $\lambda_1 = 1$

Para encontrar el autovector correspondiente a $\lambda_2 = 0.6$,

$\begin{pmatrix} 1.2-0.6 & -0.2 \\ 0.6 & 0.4-0.6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \implies 0.6x_1 = 0.2x_2$ o sea, $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ es autovector correspondiente a $\lambda_2 = 0.6$

De modo que $C = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$ y $C^{-1} = \begin{pmatrix} 1.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}$

Entonces, la expresión anterior,

$$\Delta y_{1t} = m_1 - c_{21}(1 - \lambda_2) z_{2t-1} + \varepsilon_{1t}$$
$$\Delta y_{2t} = m_2 - c_{22}(1 - \lambda_2) z_{2t-1} + \varepsilon_{2t}$$

nos queda en este ejemplo $(m_1 = m_2 = 0)$:

$$\Delta y_{1t} = 0.2 y_{1t-1} - 0.2 y_{2t-1} + \varepsilon_{1t}$$
$$\Delta y_{2t} = 0.6 y_{1t-1} - 0.6 y_{2t-1} + \varepsilon_{2t}$$

o,

$$\Delta \boldsymbol{y_t} = -\begin{pmatrix} 0.4 \\ 1.2 \end{pmatrix} (-0.5 \quad 0.5) \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

Caso 3: $\lambda_1 = 1 \ \ y \ \lambda_2 = 1$. Este caso es diferente a los anteriores, porque ahora no van a haber dos autovectores linealmente independientes asociados a los autovalores repetidos. Por lo tanto, no va a existir la inversa de la matriz ***C*** que nos permitía diagonalizar la matriz ***A***.

Por ejemplo, si $\boldsymbol{A} = \begin{pmatrix} 0.8 & -0.4 \\ 0.1 & 1.2 \end{pmatrix}$. Se puede mostrar que los autovalores asociados serán $\lambda_1 = 1 \ y \ \lambda_2 = 1$ y el autovector asociado es $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$. Pero nos va a faltar un segundo autovector.

Este caso puede analizarse usando matriz Jordan.

Si bien $\boldsymbol{A}$ no es diagonalizable, es posible encontrar una matriz no singular $\boldsymbol{P}$ tal que $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}$, de modo que $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$ donde $\boldsymbol{J} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ es la matriz Jordan para el autovalor λ con multiplicidad 2. Sea $\boldsymbol{z_t} = \boldsymbol{P}^{-1}\boldsymbol{y_t}$ de manera que $\boldsymbol{y_t} = \boldsymbol{P}\boldsymbol{z_t}$.

Sustituyendo en $(*)$ y simplificando, se obtiene que:

$$\boldsymbol{z_t} = \boldsymbol{J}\boldsymbol{z_{t-1}} + \boldsymbol{m}^* + \boldsymbol{\eta_t}$$

donde $\boldsymbol{m}^* = \boldsymbol{P}^{-1}\boldsymbol{m}$ y donde $\boldsymbol{\eta_t} = \boldsymbol{P}^{-1}\boldsymbol{\varepsilon_t}$ (vector de ruidos blanco). De modo que,

$$z_{1t} = \lambda z_{1t-1} + z_{2t-1} + {m_1}^* + \eta_{1t}$$

$$z_{2t} = \lambda z_{2t-1} + {m_2}^* + \eta_{2t}$$

Entonces, sustituyendo $\lambda = 1$,

$$(1 - L)z_{1t} = z_{2t-1} + {m_1}^* + \eta_{1t}$$

$$(1 - L)z_{2t} = {m_2}^* + \eta_{2t}$$

De modo que z_{2t} es I(1) y multiplicando la primera por $(1-L)$,

$$(1-L)^2 z_{1t} = (1-L)z_{2t-1} + (1-L)\eta_{1t}$$
$$(1-L)^2 z_{1t} = {m_2}^* + \eta_{2t} + \eta_{1t} - \eta_{1t-1}$$

Consecuentemente, z_{1t} es I(2). Y por lo tanto, y_{1t} y y_{2t} son ambas I(2).

```
var infl crec_m, lags(1/2)
```
// Así corro un VAR(2) bivariado

Causalidad en el sentido de Granger (Granger causality): Decimos que x_t causa en el sentido de Granger a y_t si x_t y rezagos de x_t explican a y_t.

Comando de Stata:

```
vargranger
```

```
Granger causality Wald tests
```

Equation	Excluded	chi2	df	Prob > chi2
infl	crec_m	1.0086	2	0.604
infl	ALL	1.0086	2	0.604
crec_m	infl	9.0166	2	0.011
crec_m	ALL	9.0166	2	0.011

En este caso concluiríamos que con una probabilidad de error tipo I de 5%, rechazamos la hipótesis nula de que INF no causa en el sentido de Granger a CREC_BM, o sea que encontramos evidencia estadística de que la tasa de inflación causa en el sentido de Granger a la tasa de crecimiento de la base monetaria o en otras palabras, que los rezagos de la tasa de inflación explican (estadísticamente) la tasa de crecimiento de la base monetaria en “t”.

¿Qué significa esto desde el punto de vista económico?

Vale la pena mencionar que causalidad en el sentido de Granger es un concepto estadístico.