

Series de Tiempo Económicas

Vectores Autorregresivos

Alejandro Mosiño - *Universidad de Guanajuato*

v. 2026.05.11

Modelos de regresión con series de tiempo

El modelo de regresión lineal

El modelo de regresión lineal en un contexto de series de tiempo es:

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t;$$
$$\varepsilon_t \sim WN(0, \sigma^2)$$

¿Cómo usamos este modelo para hacer predicciones?

Pronóstico condicional (1/2)

Podemos utilizar el modelo de regresión lineal para hacer un pronóstico de y_t *condicional* a los supuestos sobre la variable x_t . Por ejemplo:

$$\mathbb{E}(y_{T+h} | x_{T+h})^* = \beta_1 + \beta_2 x_{T+h}^*$$

donde hemos supuesto que el valor de x_t en el periodo $T + h$ es x_{T+h}^* . Este tipo de análisis se conoce en la práctica como **análisis de escenarios** o **análisis de contingencia**.

Pronóstico condicional (2/2)

Por supuesto, si suponemos que ε_t es un ruido blanco gaussiano, entonces la distribución condicional del pronóstico es:

$$y_{T+h} | x_{T+h}^* \sim \mathcal{N}(y_{T+h} | x_{T+h}^*, \sigma^2)$$

Esto nos permite calcular intervalos de confianza para la predicción.

Incertidumbre en los parámetros (1/2)

Naturalmente, el pronóstico que acabamos de realizar no está exento de incertidumbre. Para ver esto, estimamos el modelo por MCO y lo evaluamos en $\mathbf{x}_{T+h}^*$. En este caso, el pronóstico para el periodo $T + h$ es:

$$\hat{y}_{T+h,T} | \mathbf{x}_{T+h}^* = b_1 + b_2 \mathbf{x}_{T+h}^*,$$

y el residual:

$$\begin{aligned} e_{T+h,T} &= y_{T+h} - \hat{y}_{T+h,T} | \mathbf{x}_{T+h}^* \\ &= (\beta_1 - b_1) + (\beta_2 - b_2) \mathbf{x}_{T+h}^* + \varepsilon_{T+h} \end{aligned}$$

Incertidumbre en los parámetros (2/2)

Si calculamos la varianza del residual:

$$\mathbb{V}ar(e_{T+h,T}) = \mathbb{V}ar(b_1) + x_{T+h}^{*2} \mathbb{V}ar(b_2) + \sigma^2$$

donde los primeros dos términos reflejan la incertidumbre en los parámetros del modelo. Entonces, para construir un intervalo de confianza para el pronóstico que refleje esta incertidumbre:

$$\hat{y}_{T+h} | x_{T+h}^* \sim \square(\hat{y}_{T+h} | x_{T+h}^*, \mathbb{V}ar(b_1) + x_{T+h}^{*2} \mathbb{V}ar(b_2) + \sigma^2)$$

Pronóstico no condicional (1/2)

En ocasiones deseamos encontrar el mejor pronóstico posible para y_t . Esto es, un **pronóstico no condicional**. Para hacer esto, es necesario insertar un pronóstico óptimo para x_t , $x_{T+h,T}$, en la ecuación.

$$y_{T+h,T} = \beta_1 + \beta_2 x_{T+h,T}$$

Naturalmente, no siempre contamos con este valor futuro de x_t (tenemos un modelo para y_t , no para x_t).

Pronóstico no condicional (2/2)

Una solución consiste en estimar y_t usando como regresores los valores pasados de x_t : x_{t-h} , x_{t-h-1} , x_{t-h-2} , Por ejemplo, si deseamos pronosticar un periodo hacia adelante, podemos usar el modelo:

$$y_t = \beta_1 + \beta_2 x_{t-1} + \varepsilon_t$$

Por supuesto, es posible incluir más retardos de x_t . También, es importante resaltar que si el objetivo es pronosticar y_t h periodos hacia adelante, las variables x_t deben retardarse al menos h periodos.

Rezagos distribuidos (1/2)

El modelo anterior puede generalizarse como (nota el cambio de notación):

$$y_t = \phi_0 + \sum_{i=1}^{N_x} \gamma_i x_{t-i} + \varepsilon_t$$

En este caso decimos que y_t depende de los **rezagos distribuidos** de x_t . Pero, ¿cuánto vale N_x ?

Nota. ¿Por qué este modelo tendrá este nombre?

Rezagos distribuidos (2/2)

Para elegir el valor de N_x tenemos varias opciones:

- Usar pruebas de significancia estadística.
- Usar los criterios de información (AIC y BIC).
- Revisar el comportamiento de los residuales.

Los modelos con rezagos distribuidos (si las series son estacionarias) pueden estimarse por MCO o por MCG.

Modelo autorregresivo con rezagos distribuidos (1/2)

Los modelos con rezagos distribuidos pueden generalizarse aun más:

$$y_t = \phi_0 + \sum_{i=1}^{N_y} \phi_i y_{t-i} + \sum_{j=1}^{N_x} \gamma_j x_{t-j} + \varepsilon_t$$

En este modelo hemos agregado retardos de la variable dependiente. Este es un **modelo AR con rezagos distribuidos (ARDL) de orden N_y, N_x** . Este puede estimarse (si las series son estacionarias *y no cointegradas*) por Máxima Verosimilitud, MCO o MCG.

Modelo autorregresivo con rezagos distribuidos (2/2)

Nota

Se puede mostrar que los modelos ARDL pueden expresarse como modelos con *errores ARMA*. Para ver esto, considera el modelo $ARDL(N_y, N_x)$:

$$y_t = \phi_0 + \sum_{i=1}^{N_y} \phi_i y_{t-i} + \sum_{j=1}^{N_x} \gamma_j x_{t-j} + \varepsilon_t$$

El operador de rezagos nos permite reescribir el modelo como:

$$y_t = \frac{\phi_0}{\Phi(L)} + \frac{\Gamma(L)}{\Phi(L)} x_t + \frac{1}{\Phi(L)} \varepsilon_t; \quad \Phi(L) = 1 - \sum_{i=1}^{N_y} \phi_i L^i; \quad \Gamma(L) = \sum_{j=0}^{N_x} \gamma_j L^j.$$

Nota que el término $\frac{1}{\Phi(L)} \varepsilon_t$ induce un $MA(\infty)$ en el error, lo cual es equivalente a un modelo ARMA bajo invertibilidad.

Ejemplo (1/4)

Consideremos el siguiente modelo:

$$u_t = \phi_0 + \phi_1 u_{t-1} + \phi_2 u_{t-2} + \gamma_1 g_{t-1} + \varepsilon_t$$

Donde:

- u_t es la tasa de desempleo en tiempo t ,
- g_t es la tasa de crecimiento del PIB en tiempo t .

Ejemplo (2/4)

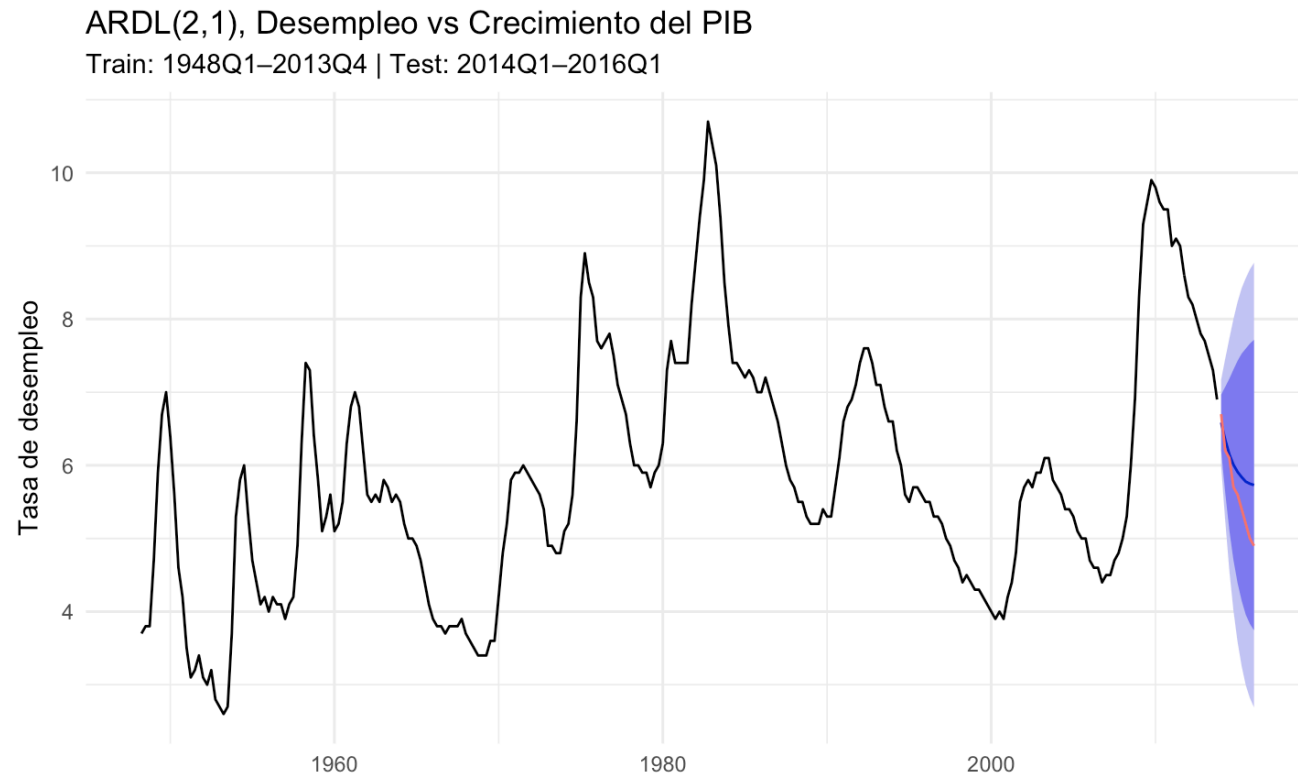
Usando datos de los Estados Unidos tenemos que:

Modelo ARDL(2,1)

term	estimate	std.error	statistic	p.value
(Intercept)	0.3615676	0.0722796	5.002348	0.000001
lag(u, 1)	1.5331023	0.0555538	27.596726	0.000000
lag(u, 2)	-0.5817890	0.0555935	-10.465055	0.000000
lag(g, 1)	-0.0482379	0.0194888	-2.475156	0.013939

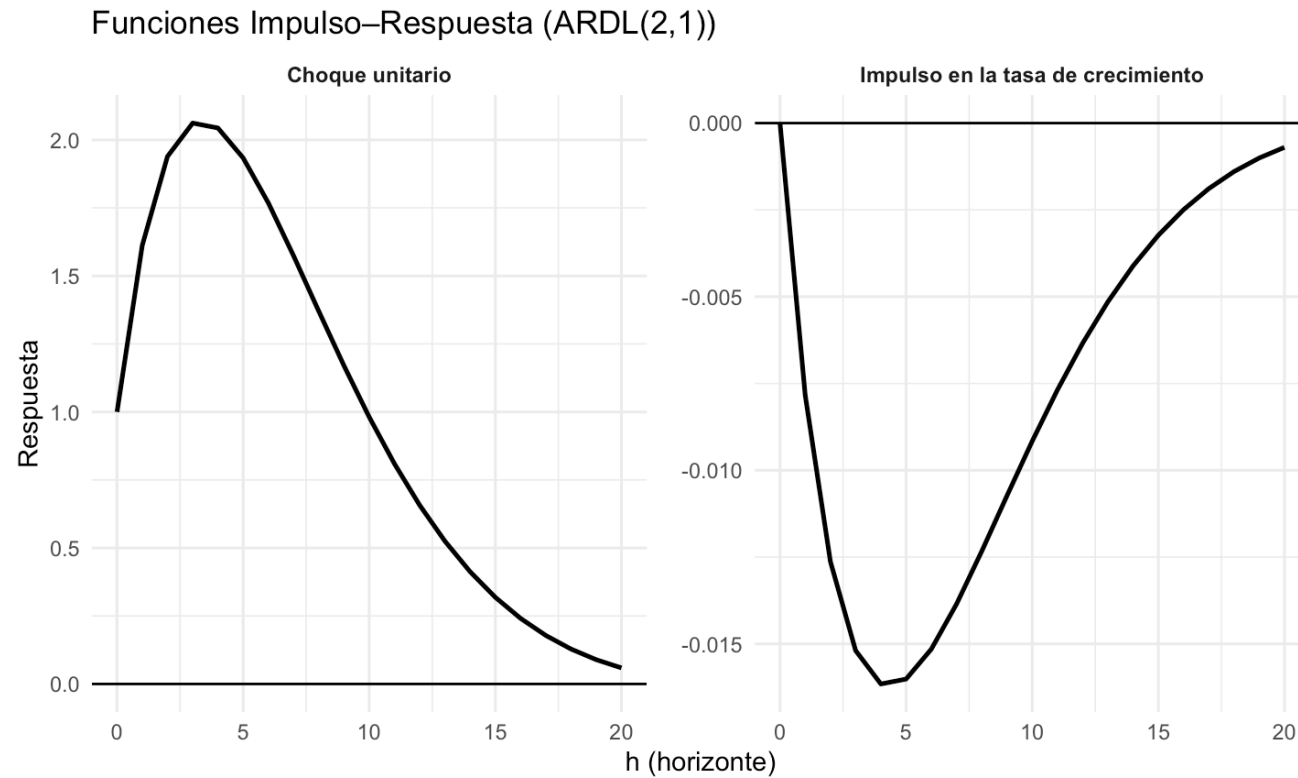
Ejemplo (3/4)

Podemos hacer pronósticos usando modelos autorregresivos con rezagos distribuidos usando las mismas técnicas aprendidas antes (ver modelos ARMA).



Ejemplo (4/4)

La representación ARMA del modelo ARDL nos permite calcular su *función de impulso-respuesta*. En el ejemplo tenemos:



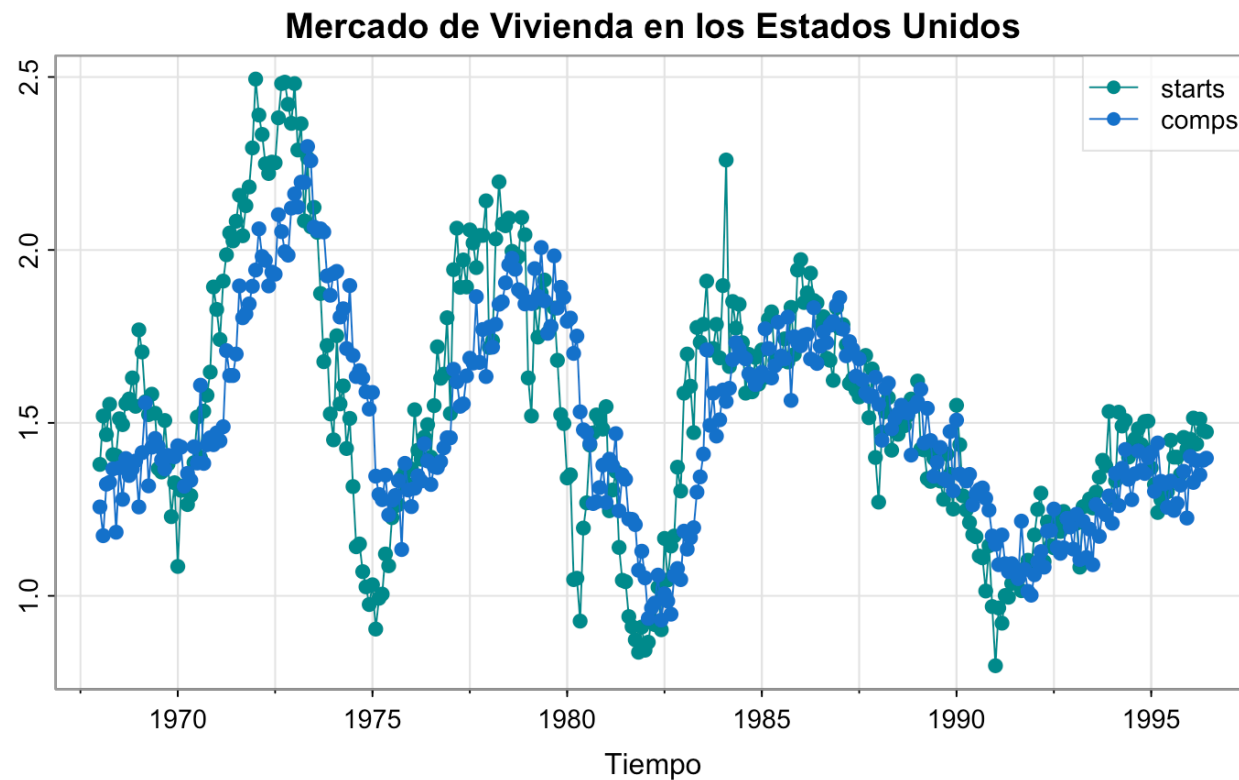
Vectores Autorregresivos

¿Qué son los vectores autorregresivos?

Los **modelos con vectores autorregresivos (VAR)** son una generalización de los *modelos autorregresivos uniecuacionales (AR)* y *de rezagos distribuidos (DL)*. Estos sirven para capturar la evolución y las interdependencias entre múltiples series de tiempo.

Vectores autorregresivos: ejemplo

Como ejemplo, consideramos la relación entre el inicio y el fin de construcción de nuevas viviendas en los Estados Unidos:



Estructura de los VAR (1/5)

Un VAR involucra k variables. En un vector autorregresivo de orden p con k variables, VAR(p), tenemos que estimar k diferentes ecuaciones. En cada ecuación tenemos como regresores p rezagos de la variable dependiente y p rezagos de todas las demás variables del sistema.

Por ejemplo, para un VAR(1) con dos variables, $y_{1,t}$ y $y_{2,t}$, tenemos:

$$y_{1,t} = \phi_{1,1}^1 y_{1,t-1} + \phi_{2,1}^1 y_{2,t-1} + \varepsilon_{1,t}$$

$$y_{2,t} = \phi_{1,1}^2 y_{1,t-1} + \phi_{2,1}^2 y_{2,t-1} + \varepsilon_{2,t}$$

$$\varepsilon_{1,t} \sim WN(0, \sigma_1^2)$$

$$\varepsilon_{2,t} \sim WN(0, \sigma_2^2)$$

y:

$$\text{Cov}(\varepsilon_{1,t}, \varepsilon_{2,\tau}) = \begin{cases} \sigma_{12}; & t = \tau \\ 0; & t \neq \tau \end{cases}$$

Estructura de los VAR (2/5)

Un VAR(1) puede escribirse en una forma más compacta:

$$\mathbf{y}_t = \boldsymbol{\phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

donde:

$$\mathbf{y}_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}; \quad \boldsymbol{\phi}_1 = \begin{pmatrix} \phi_{1,1}^1 & \phi_{2,1}^1 \\ \phi_{1,1}^2 & \phi_{2,1}^2 \end{pmatrix}; \quad \boldsymbol{\varepsilon}_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

Estructura de los VAR (3/5)

Ejercicio

Considera el siguiente modelo VAR(2):

$$\begin{aligned}y_{1,t} &= \phi_{1,1}^1 y_{1,t-1} + \phi_{2,1}^1 y_{2,t-1} + \varepsilon_{1,t} \\y_{2,t} &= \phi_{1,1}^2 y_{1,t-1} + \phi_{2,2}^2 y_{2,t-2} + \varepsilon_{2,t} \\ \varepsilon_{1,t-1} &\sim WN(0, \sigma_1^2) \\ \varepsilon_{2,t-1} &\sim WN(0, \sigma_2^2) \\ \text{Cov}(\varepsilon_{1,t}, \varepsilon_{2,\tau}) &= \begin{cases} \sigma_{12}; & t = \tau \\ 0; & t \neq \tau \end{cases}\end{aligned}$$

Escribe el modelo en forma compacta.

Estructura de los VAR (4/5)

En general, un VAR(p) puede escribirse como:

$$\mathbf{y}_t = \boldsymbol{\phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\phi}_2 \mathbf{y}_{t-2} + \cdots + \boldsymbol{\phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

O como:

$$\boldsymbol{\Phi}(L) \mathbf{y}_t = \boldsymbol{\varepsilon}_t$$

Donde $\boldsymbol{\Phi}(L)$ es un polinomio de orden p :

$$\boldsymbol{\Phi}(L) = \mathbf{I}_k - \boldsymbol{\phi}_1 L - \boldsymbol{\phi}_2 L^2 - \boldsymbol{\phi}_p L^p$$

Estructura de los VAR (5/5)

Además, es fácil deducir que, para un VAR(p):

$$\begin{aligned}\mathbb{E}(\boldsymbol{\varepsilon}_t) &= \mathbf{0} \\ \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_\tau') &= \begin{cases} \boldsymbol{\Omega} & t = \tau \\ \mathbf{0} & t \neq \tau \end{cases}\end{aligned}$$

donde $\boldsymbol{\Omega}$ es una matriz cuadrada de tamaño k , simétrica y definida positiva.

Estacionariedad de los VAR (1/5)

Igual que para los procesos AR(p), un proceso VAR(p) es estacionario en tanto las raíces del polinomio:

$$|\mathbf{I}_k - \phi_1 L - \phi_2 L^2 - \phi_p L^p| = 0$$

se encuentren *fuera del círculo unitario*.

Estacionariedad de los VAR (2/5)

Ejemplo

Consideremos el proceso:

$$\mathbf{y}_t = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim WN(\mathbf{0}, \boldsymbol{\Omega})$$

El polinomio característico del proceso es:

$$|\mathbf{I}_2 - \boldsymbol{\Phi}_1 L| = 0$$

En este caso, es fácil mostrar que raíces de este polinomio son:

$$L_1 = \frac{1}{0.79} > 1, \quad L_2 = \frac{1}{0.51} > 1$$

Estacionariedad de los VAR (3/5)

Nota

La condición de estacionariedad se puede encontrar a partir de los *valores propios de una matriz cuadrada*. Para ver esto, considera un proceso VAR(p):

$$\mathbf{y}_t = \boldsymbol{\phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\phi}_2 \mathbf{y}_{t-2} + \cdots + \boldsymbol{\phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

Definimos la **matriz compañera** como:

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_{p-1} & \boldsymbol{\phi}_p \\ \mathbf{I}_k & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_k & \mathbf{0} \end{pmatrix}$$

Para que este proceso sea estacionario, es necesario que los *valores propios de la matriz compañera* se encuentren *dentro* del círculo unitario.

Estacionariedad de los VAR (4/5)

Ejemplo

Consideremos el proceso:

$$\mathbf{y}_t = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim WN(\mathbf{0}, \boldsymbol{\Omega})$$

Muestra que los valores propios de la matriz $\boldsymbol{\Phi}_1$ son:

$$\lambda_1 = 0.79 < 1, \quad \lambda_2 = 0.51 < 1$$

Estacionariedad de los VAR (5/5)

Ejercicio

Considera el proceso:

$$\mathbf{y}_t = \begin{pmatrix} 0.50 & 0 \\ 0 & 0.40 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} 0.20 & 0 \\ 0 & 0.10 \end{pmatrix} \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim WN(\mathbf{0}, \boldsymbol{\Omega}),$$

1. Encuentra las raíces del polinomio $|\mathbf{I} - \boldsymbol{\Phi}_1 L - \boldsymbol{\Phi}_2 L^2| = 0$.
2. Encuentra la matriz compañera.
3. Encuentra los valores propios de la matriz compañera.
4. ¿Es este proceso estacionario?

Matriz de autocovarianzas de un VAR(p) (1/3)

Considera un proceso VAR(p):

$$\mathbf{y}_t = \sum_{i=1}^p \Phi_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0}, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_\tau') = \begin{cases} \boldsymbol{\Omega}, & t = \tau \\ \mathbf{0}, & t \neq \tau \end{cases}$$

La **función de autocovarianzas** puede calcularse como:

$$\boldsymbol{\Gamma}(h) = \mathbb{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-h} - \boldsymbol{\mu})'], \quad h = 0, 1, 2, \dots$$

Nota. Al igual que en el caso de los procesos AR(p), tenemos estacionariedad si $\boldsymbol{\Gamma}(h)$ no depende de t .

Matriz de autocovarianzas de un VAR(p) (2/3)

Por simplicidad, supongamos que $\boldsymbol{\mu} = \mathbf{0}$. En este caso:

$$\boldsymbol{\Gamma}(h) = \mathbb{E}[\mathbf{y}_t \mathbf{y}_{t-h}'], \quad h = 0, 1, 2, \dots$$

Matriz de autocovarianzas de un VAR(p) (3/3)

Nota

Nota que ahora:

$$\begin{aligned}\mathbf{\Gamma}(h) &= \mathbb{E}[\mathbf{y}_t \mathbf{y}'_{t-h}] \\ &= \mathbb{E}[\mathbf{y}_{t+h} \mathbf{y}'_{t+h-h}] \\ &= \mathbb{E}[\mathbf{y}_{t+h} \mathbf{y}'_t] \\ &\neq \mathbf{\Gamma}(-h)\end{aligned}$$

Sin embargo:

$$\begin{aligned}\mathbf{\Gamma}(h)' &= \mathbb{E}[\mathbf{y}_{t+h} \mathbf{y}'_t]' \\ &= \mathbb{E}[\mathbf{y}_t \mathbf{y}'_{t+h}] \\ &= \mathbf{\Gamma}(-h)\end{aligned}$$

Ecuaciones de Yule-Walker (1/2)

Las **ecuaciones de Yule-Walker** se calculan de manera similar al caso de los procesos $AR(p)$. Estas nos ayudan a sistematizar el cálculo de la matriz de las autocovarianzas de un $VAR(p)$.

Paso 1

Multiplicamos el proceso por $\mathbf{y}'_t$ y calculamos el valor esperado:

$$\mathbb{E}(\mathbf{y}_t \mathbf{y}'_t) = \sum_{i=1}^p \mathbf{\Phi}_i \mathbb{E}(\mathbf{y}_{t-i} \mathbf{y}'_t) + \mathbb{E}(\boldsymbol{\varepsilon}_t \mathbf{y}'_t)$$

Usando la definición de un $VAR(p)$ y las propiedades de $\boldsymbol{\varepsilon}_t$:

$$\boldsymbol{\Gamma}(0) = \sum_{i=1}^p \mathbf{\Phi}_i \boldsymbol{\Gamma}(i)' + \boldsymbol{\Omega}$$

Ecuaciones de Yule-Walker (2/2)

Paso 2

Ahora multiplicamos el proceso por $\mathbf{y}'_{t-h}$ para $h \geq 1$. Obtenemos que:

$$\begin{aligned}\mathbb{E}(\mathbf{y}_t \mathbf{y}'_{t-h}) &= \sum_{i=1}^p \Phi_i \mathbb{E}(\mathbf{y}_{t-i} \mathbf{y}'_{t-h}) + \mathbb{E}(\boldsymbol{\varepsilon}_t \mathbf{y}'_{t-h}) \\ &= \sum_{i=1}^p \Phi_i \mathbb{E}(\mathbf{y}_{t-i} \mathbf{y}'_{t-h})\end{aligned}$$

Esto es:

$$\Gamma(h) = \sum_{i=1}^p \Phi_i \Gamma(h - i)$$

Ecuación de Lyapunov discreta (1/3)

Considera un VAR(1):

$$\mathbf{y}_t = \boldsymbol{\phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

El paso 1 implica:

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}(1)' + \boldsymbol{\Omega}$$

El paso 2 implica:

$$\boldsymbol{\Gamma}(h) = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}(h-1)$$

Ecuación de Lyapunov discreta (2/3)

Entonces:

$$\mathbf{\Gamma}(0) = \mathbf{\Phi}_1 \mathbf{\Gamma}(0) \mathbf{\Phi}_1' + \mathbf{\Omega}$$

Esta se conoce como **ecuación de Lyapunov discreta** y nos permite obtener la matriz de varianzas y covarianzas contemporáneas de un proceso VAR(1) estacionario.

Ecuación de Lyapunov discreta (3/3)

Nota

La ecuación de Lyapunov discreta se extiende de forma natural a procesos VAR(p). Para esto, el modelo puede reescribirse en su **forma estado-espacio**:

$$\xi_t = F \xi_{t-1} + v_t,$$

La covarianza estacionaria de ξ_t satisface:

$$\Gamma_\xi(0) = F \Gamma_\xi(0) F' + Q.$$

El bloque superior izquierdo de $\Gamma_\xi(0)$ corresponde a $\Gamma(0)$ del VAR original.

Ejemplo (1/2)

Ejemplo

Considera el VAR(1):

$$\mathbf{y}_t = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

Utiliza las ecuaciones de Yule–Walker para encontrar que:

$$\boldsymbol{\Gamma}(0) \approx \begin{pmatrix} 2.85 & 1.71 \\ 1.71 & 1.81 \end{pmatrix}.$$

Ejemplo (2/2)

A partir del resultado anterior, podemos encontrar que:

$$\begin{aligned}\Gamma(1) &= \Phi_1 \Gamma(0) \\ &\approx \begin{pmatrix} 2.47 & 1.64 \\ 1.28 & 1.31 \end{pmatrix}\end{aligned}$$

Y:

$$\begin{aligned}\Gamma(2) &= \Phi_1 \Gamma(1) \\ &\approx \begin{pmatrix} 2.01 & 1.39 \\ 1.04 & 1.03 \end{pmatrix}\end{aligned}$$

Representación VMA(∞) (1/3)

Supongamos que las raíces del polinomio $|\Phi(L)| = 0$ se encuentran fuera del círculo unitario. En este caso, el polinomio $\Psi(L) = \Phi(L)^{-1}$ existe, lo que nos permite expresar un proceso VAR(p) en la forma de un proceso VMA(∞):

$$\begin{aligned} \mathbf{y}_t &= \Phi(L)^{-1} \boldsymbol{\varepsilon}_t \\ &= \Psi(L) \boldsymbol{\varepsilon}_t \\ &= \sum_{i=0}^{\infty} \Psi_i \boldsymbol{\varepsilon}_{t-i} \end{aligned}$$

Con:

$$\begin{aligned} \Psi_0 &= \mathbf{I}_k \\ \Psi_1 &= \phi_1 \\ \Psi_2 &= \phi_1 \Psi_1 + \phi_2 \\ &\vdots \\ \Psi_s &= \phi_1 \Psi_{s-1} + \phi_2 \Psi_{s-2} + \cdots + \phi_p \Psi_{s-p} \end{aligned}$$

Representación VMA(∞) (2/3)

Ejemplo

Consideremos un VAR(1) bivariado:

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \mathbf{\Phi}_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix}, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \mathbf{\Omega} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

(Usando las propiedades del operador de rezagos, las cuales se extienden al caso matricial) No es difícil demostrar que la representación VMA(∞) es:

$$\mathbf{y}_t = \sum_{h=0}^{\infty} \mathbf{\Psi}_h \boldsymbol{\varepsilon}_{t-h}, \quad \mathbf{\Psi}_0 = \mathbf{I}_2, \quad \mathbf{\Psi}_h = \mathbf{\Phi}_1^h \quad (h \geq 1).$$

Representación VMA(∞) (3/3)

Ejemplo (continuación)

Específicamente:

$$\Psi_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Psi_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix}$$

$$\Psi_2 = \begin{pmatrix} 0.51 & 0.26 \\ 0.13 & 0.38 \end{pmatrix}$$

$$\Psi_3 = \begin{pmatrix} 0.383 & 0.258 \\ 0.129 & 0.254 \end{pmatrix}$$

$$\Psi_4 \approx \begin{pmatrix} 0.294 & 0.231 \\ 0.116 & 0.178 \end{pmatrix}$$

$\vdots$

Representación VMA(∞): Matriz de autocovarianzas (1/2)

La matriz de autocovarianzas de un proceso VAR(p) también puede obtenerse a partir de su representación VMA(∞). En particular:

$$\begin{aligned}\Gamma(h) &= \mathbb{E}(\mathbf{y}_t \mathbf{y}_{t-h}') \\ &= \mathbb{E} \left[\left(\sum_{j=0}^{\infty} \Psi_j \boldsymbol{\varepsilon}_{t-j} \right) \left(\sum_{i=0}^{\infty} \Psi_i \boldsymbol{\varepsilon}_{t-h-i} \right)' \right] \\ &= \sum_{j=0}^{\infty} \Psi_j \Omega \Psi_{j+h}'\end{aligned}$$

Nota. Esta expresión es equivalente a la obtenida con las ecuaciones de Yule–Walker, pero resulta más práctica si disponemos de las matrices Ψ_h .

Representación VMA(∞): Matriz de autocovarianzas (2/2)

Ejercicio

Considera un VAR(1) bivariado:

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \mathbf{\Phi}_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix}, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \mathbf{\Omega} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

Utiliza la fórmula de $\mathbf{\Gamma}(h)$ para mostrar que esta es equivalente a la encontrada usando el procedimiento de Yule-Walker.

Pista. Quizás sea necesario usar tu computadora.

Función Impulso–Respuesta (1/2)

Consideremos un proceso AR(1):

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \quad |\phi_1| < 1.$$

Como hemos visto, este admite la siguiente representación MA(∞):

$$y_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots$$

La **Función de Impulso-Respuesta (FIR)** muestra cómo un choque en el término de error, ε_t , afecta a una serie de tiempo, y_t , en periodos futuros. En el caso de un AR(1) tenemos:

Periodo	t	$t + 1$	$t + 2$	$t + 3$	...
ε_t	1	0	0	0	...
y_t	1	ϕ_1	ϕ_1^2	ϕ_1^3	...

Función Impulso–Respuesta (2/2)

Considera ahora un VAR(1):

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \mathbf{\Omega}.$$

Como hemos visto, este puede representarse como:

$$\mathbf{y}_t = \boldsymbol{\varepsilon}_t + \mathbf{\Phi}_1 \boldsymbol{\varepsilon}_{t-1} + \mathbf{\Phi}_1^2 \boldsymbol{\varepsilon}_{t-2} + \mathbf{\Phi}_1^3 \boldsymbol{\varepsilon}_{t-3} + \cdots$$

Cada matriz $\mathbf{\Phi}_1^h$ describe cómo un choque en (*cualquiera de los elementos* de) $\boldsymbol{\varepsilon}_t$ afecta la trayectoria de $\mathbf{y}_t$.

Función Impulso–Respuesta: Ejemplo (1/2)

Ejemplo

Consideremos un VAR(1) bivariado:

$$\mathbf{y}_t = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

En este caso:

$$\mathbf{y}_t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{\varepsilon}_t + \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \boldsymbol{\varepsilon}_{t-1} + \begin{pmatrix} 0.51 & 0.26 \\ 0.13 & 0.38 \end{pmatrix} \boldsymbol{\varepsilon}_{t-2} + \begin{pmatrix} 0.383 & 0.258 \\ 0.129 & 0.254 \end{pmatrix} \boldsymbol{\varepsilon}_{t-3} + \dots$$

Función Impulso–Respuesta: Ejemplo (2/2)

Para interpretar, consideremos un choque de una unidad en $\varepsilon_{1,t}$:

	t	$t + 1$	$t + 2$	$t + 3$	...
$\varepsilon_{1,t}$	1	0	0	0	...
$\varepsilon_{2,t}$	0	0	0	0	...
$y_{1,t}$	$(\Psi_0)_{1,1} = 1$	$(\Psi_1)_{1,1} = 0.70$	$(\Psi_2)_{1,1} = 0.51$	$(\Psi_3)_{1,1} = 0.383$	...
$y_{2,t}$	$(\Psi_0)_{2,1} = 0$	$(\Psi_1)_{2,1} = 0.10$	$(\Psi_2)_{2,1} = 0.13$	$(\Psi_3)_{2,1} = 0.129$	...

Por su parte, el efecto de un choque de una unidad en $\varepsilon_{2,t}$:

	t	$t + 1$	$t + 2$	$t + 3$	...
$\varepsilon_{1,t}$	0	0	0	0	...
$\varepsilon_{2,t}$	1	0	0	0	...
$y_{1,\cdot}$	$(\Psi_0)_{1,2} = 0$	$(\Psi_1)_{1,2} = 0.20$	$(\Psi_2)_{1,2} = 0.26$	$(\Psi_3)_{1,2} = 0.258$	...
$y_{2,\cdot}$	$(\Psi_0)_{2,2} = 1$	$(\Psi_1)_{2,2} = 0.60$	$(\Psi_2)_{2,2} = 0.38$	$(\Psi_3)_{2,2} = 0.254$	...

Función Impulso–Respuesta: VAR(p) (1/2)

Para un VAR(p) aplicamos la misma lógica. Tenemos:

$$\mathbf{y}_t = \sum_{i=1}^p \mathbf{\Phi}_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \mathbf{\Omega}.$$

Su representación VMA(∞) es:

$$\mathbf{y}_t = \sum_{h=0}^{\infty} \mathbf{\Psi}_h \boldsymbol{\varepsilon}_{t-h}, \quad \mathbf{\Psi}_0 = \mathbf{I}_k, \quad \mathbf{\Psi}_h = \sum_{i=1}^p \mathbf{\Phi}_i \mathbf{\Psi}_{h-i} \quad (h \geq 1),$$

Las matrices $\mathbf{\Psi}_h$ son las funciones impulso–respuesta (FIR). *Cada una indica cómo un choque actual afecta a las variables endógenas en los siguientes periodos.*

Función Impulso–Respuesta: VAR(p) (2/2)

Ejercicio

Considera el siguiente proceso VAR(2):

$$\mathbf{y}_t = \begin{pmatrix} 0.6 & 0.2 \\ 0.1 & 0.5 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\Omega} = \mathbf{I}_2$$

1. Verifica que el proceso es estable.
2. Encuentra la forma VMA(∞) del proceso.
3. Construye la FIR para un choque de una unidad en $\boldsymbol{\varepsilon}_{1,t}$ y en $\boldsymbol{\varepsilon}_{2,t}$ hasta $h = 3$.

Funciones Impulso–Respuesta: Ortogonalización (1/4)

Considera un VAR(p):

$$\mathbf{y}_t = \sum_{i=1}^p \Phi_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega}.$$

Nota que, *dado que $\boldsymbol{\Omega}$ no es necesariamente diagonal, un un choque unitario en $\boldsymbol{\varepsilon}_{k,t}$ podría afectar a todas las ecuaciones del sistema*. Por lo tanto, las FIR no pueden interpretarse como **choques estructurales puros**.

Funciones Impulso–Respuesta: Ortogonalización (2/4)

Para evitar este problema, podemos construir innovaciones **mutuamente ortogonales**. Esto es:

$$\mathbb{E}(\mathbf{u}_t \mathbf{u}_t') = \mathbf{I}_k,$$

En este caso, cada impulso $\mathbf{u}_{j,t}$ representa un **shock estructural independiente**.

Funciones Impulso–Respuesta: Ortogonalización (3/4)

Para construir innovaciones ortogonales, podemos utilizar la **descomposición de Cholesky**:

$$\mathbf{\Omega} = \mathbf{P}\mathbf{P}',$$

donde $\mathbf{P}$ es una *matriz triangular inferior*.

En este caso:

$$\boldsymbol{\varepsilon}_t = \mathbf{P}\mathbf{u}_t, \quad \Rightarrow \quad \mathbb{E}(\mathbf{u}_t\mathbf{u}_t') = \mathbf{I}_k.$$

Funciones Impulso–Respuesta: Ortogonalización (4/4)

Ahora, sustituimos en la forma VMA(∞) de un VAR(p):

$$\begin{aligned} \mathbf{y}_t &= \sum_{h=0}^{\infty} \mathbf{\Psi}_h \boldsymbol{\varepsilon}_{t-h} \\ &= \sum_{h=0}^{\infty} \mathbf{\Psi}_h \mathbf{P} \mathbf{u}_{t-h} . \end{aligned}$$

De aquí, podemos construir la **FIR ortogonalizada** o **FIR estructural**. Esta es:

$$\mathbf{\Theta}_h = \mathbf{\Psi}_h \mathbf{P} .$$

Ortogonalización: Ejemplo (1/3)

Consideremos nuevamente el VAR(1) bivariado:

$$\mathbf{y}_t = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}.$$

Como primer paso, obtenemos la descomposición de Cholesky de la matriz $\boldsymbol{\Omega}$:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 0.6 & 0.8 \end{pmatrix}, \quad \boldsymbol{\varepsilon}_t = \mathbf{P}\mathbf{u}_t, \quad \mathbb{E}(\mathbf{u}_t \mathbf{u}_t') = \mathbf{I}_2.$$

Ortogonalización: Ejemplo (2/3)

La FIR orthogonalizada no es más que:

$$\mathbf{\Theta}_h = \mathbf{\Psi}_h \mathbf{P}.$$

Por ejemplo:

$$\mathbf{\Theta}_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0.6 & 0.8 \end{pmatrix} = \begin{pmatrix} 0.82 & 0.16 \\ 0.46 & 0.48 \end{pmatrix}.$$

Ortogonalización: Ejemplo (3/3)

Continuando con el procedimiento, no es difícil demostrar que la FIR orthogonalizada es:

Choque estructural en $u_{1,t}$:

	t	$t + 1$	$t + 2$	$t + 3$
$u_{1,t}$	1	0	0	0
$u_{2,t}$	0	0	0	0
$y_{1,t}$	$(\Theta_0)_{1,1} = 1.00$	$(\Theta_1)_{1,1} = 0.82$	$(\Theta_2)_{1,1} = 0.64$	$(\Theta_3)_{1,1} = 0.51$
$y_{2,t}$	$(\Theta_0)_{2,1} = 0.60$	$(\Theta_1)_{2,1} = 0.46$	$(\Theta_2)_{2,1} = 0.33$	$(\Theta_3)_{2,1} = 0.24$

Choque estructural en $u_{2,t}$:

	t	$t + 1$	$t + 2$	$t + 3$
$u_{1,t}$	0	0	0	0
$u_{2,t}$	1	0	0	0
$y_{1,t}$	$(\Theta_0)_{1,2} = 0$	$(\Theta_1)_{1,2} = 0.16$	$(\Theta_2)_{1,2} = 0.21$	$(\Theta_3)_{1,2} = 0.19$
$y_{2,t}$	$(\Theta_0)_{2,2} = 0.80$	$(\Theta_1)_{2,2} = 0.48$	$(\Theta_2)_{2,2} = 0.36$	$(\Theta_3)_{2,2} = 0.28$

Descomposición de la varianza del error de pronóstico

(1/4)

La **Descomposición de la varianza del error de pronóstico (FEVD)** (Forecast Error Variance Decomposition) mide *qué proporción de la varianza del error de pronóstico de cada variable se debe a los diferentes choques estructurales del sistema.*

Para calcular esta, partimos de la representación VMA(∞) ortogonalizada:

$$\mathbf{y}_t = \sum_{h=0}^{\infty} \mathbf{\Theta}_h \mathbf{u}_{t-h}, \quad \mathbb{E}(\mathbf{u}_t \mathbf{u}_t') = \mathbf{I}_k,$$

donde:

$$\mathbf{\Theta}_h = \mathbf{\Psi}_h \mathbf{P}, \quad \mathbf{P} \mathbf{P}' = \mathbf{\Omega}.$$

Descomposición de la varianza del error de pronóstico

(2/4)

El error de pronóstico a k pasos para $\mathbf{y}_t$ es:

$$\mathbf{e}_{t+k|t} = \mathbf{y}_{t+k} - \mathbb{E}_t(\mathbf{y}_{t+k}) = \sum_{h=0}^{k-1} \mathbf{\Theta}_h \mathbf{u}_{t+k-h}.$$

Su matriz de varianza-covarianza es:

$$\mathbb{E}(\mathbf{e}_{t+k|t} \mathbf{e}_{t+k|t}') = \sum_{h=0}^{k-1} \mathbf{\Theta}_h \mathbf{\Theta}_h'.$$

Cada elemento diagonal corresponde a la varianza del error de pronóstico de cada variable a k pasos.

Descomposición de la varianza del error de pronóstico

(3/4)

Para cada variable y_i , la varianza de su error de pronóstico a h pasos es:

$$\text{Var}(e_{i,t+h}|t) = \sum_{s=0}^{h-1} \sum_{m=1}^k (\Theta_s)_{im}^2.$$

La contribución del choque estructural u_j a dicha varianza es:

$$\text{Var}_{(i \leftarrow j)}(e_{i,t+h}|t) = \sum_{s=0}^{h-1} (\Theta_s)_{ij}^2.$$

Descomposición de la varianza del error de pronóstico

(4/4)

Entonces, la proporción de la varianza del error de pronóstico atribuible al error u_j es:

$$\text{FEVD}_{ij}(h) = \frac{\sum_{s=0}^{h-1} (\Theta_s)_{ij}^2}{\sum_{s=0}^{h-1} \sum_{m=1}^k (\Theta_s)_{im}^2}.$$

En general, la FEVD:

- Cuantifica *qué tanto de la incertidumbre futura* de cada variable se debe a los diferentes shocks estructurales.
- Las proporciones $\text{FEVD}_{ij}(k)$ para un i fijo suman 1.
- A medida que $k \rightarrow \infty$, la descomposición converge al efecto de largo plazo de cada choque sobre la variable i .

FEVD: Ejemplo (1/6)

Consideremos el VAR(1) bivariado:

$$\mathbf{y}_t = \begin{pmatrix} 0.7 & 0.2 \\ 0.1 & 0.6 \end{pmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}.$$

Donde, como hemos visto:

$$\boldsymbol{\Theta}_0 = \begin{pmatrix} 1 & 0 \\ 0.6 & 0.8 \end{pmatrix}, \quad \boldsymbol{\Theta}_1 = \begin{pmatrix} 0.82 & 0.16 \\ 0.46 & 0.48 \end{pmatrix}, \quad \boldsymbol{\Theta}_2 = \begin{pmatrix} 0.636 & 0.208 \\ 0.328 & 0.356 \end{pmatrix}, \quad \boldsymbol{\Theta}_3 = \begin{pmatrix} 0.505 & 0.190 \\ 0.244 & 0.277 \end{pmatrix}.$$

FEVD: Ejemplo (2/6)

Para $h = 1$, usando:

$$\Theta_0 = \begin{pmatrix} 1 & 0 \\ 0.6 & 0.8 \end{pmatrix},$$

la FEVD se calcula como:

$$\text{FEVD}_{ij}(1) = \frac{(\Theta_0)_{ij}^2}{\sum_{m=1}^k (\Theta_0)_{im}^2}, \quad k = 2.$$

FEVD: Ejemplo (3/6)

Entonces:

$$\begin{aligned}\text{FEVD}_{1 \leftarrow 1}(1) &= \frac{1^2}{1^2 + 0^2} \\ &= 1.00\end{aligned}$$

$$\begin{aligned}\text{FEVD}_{1 \leftarrow 2}(1) &= \frac{0^2}{1^2 + 0^2} \\ &= 0.00\end{aligned}$$

$$\begin{aligned}\text{FEVD}_{2 \leftarrow 1}(1) &= \frac{0.6^2}{0.6^2 + 0.8^2} \\ &= 0.36\end{aligned}$$

$$\begin{aligned}\text{FEVD}_{2 \leftarrow 2}(1) &= \frac{0.8^2}{0.6^2 + 0.8^2} \\ &= 0.64\end{aligned}$$

FEVD: Ejemplo (4/6)

Para $h = 2$, usamos las matrices:

$$\Theta_0 = \begin{pmatrix} 1 & 0 \\ 0.6 & 0.8 \end{pmatrix}, \quad \Theta_1 = \begin{pmatrix} 0.82 & 0.16 \\ 0.46 & 0.48 \end{pmatrix},$$

Entonces:

$$\text{FEVD}_{ij}(2) = \frac{\sum_{s=0}^1 (\Theta_s)_{ij}^2}{\sum_{s=0}^1 \sum_{m=1}^k (\Theta_s)_{im}^2}, \quad k = 2.$$

FEVD: Ejemplo (5/6)

Sustituyendo:

$$\text{FEVD}_{1 \leftarrow 1}(2) = \frac{1.6724}{1.6980} \\ \approx 0.985$$

$$\text{FEVD}_{1 \leftarrow 2}(2) = \frac{0^2 + 0.16^2}{1.6980} \\ \approx 0.015$$

$$\text{FEVD}_{2 \leftarrow 1}(2) = \frac{0.5716}{1.4420} \\ \approx 0.397$$

$$\text{FEVD}_{2 \leftarrow 2}(2) = \frac{0.8^2 + 0.48^2}{1.4420} \\ \approx 0.603$$

FEVD: Ejemplo (6/6)

Ejercicio

Muestra que la FEVD para y_1 es:

	$t + 1$	$t + 2$	$t + 3$	$t + 4$
$u_{1,t}$	1.00	0.985	0.968	0.957
$u_{2,t}$	0.00	0.015	0.032	0.043

Y para y_2 :

	$t + 1$	$t + 2$	$t + 3$	$t + 4$
$u_{1,t}$	0.36	0.397	0.405	0.408
$u_{2,t}$	0.64	0.603	0.595	0.592

Estimación de un VAR(p) (1/3)

Consideremos el modelo:

$$\mathbf{y}_t = \mathbf{c} + \sum_{i=1}^p \mathbf{\Phi}_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \text{i.i.d. } \square(\mathbf{0}, \mathbf{\Omega}).$$

Para simplificar, definimos:

$$\mathbf{x}_t = \begin{bmatrix} 1 & \mathbf{y}_{t-1} & \cdots & \mathbf{y}_{t-p} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \mathbf{c} \\ \mathbf{\Phi}_1 \\ \vdots \\ \mathbf{\Phi}_p \end{bmatrix}.$$

Esto nos permite reescribir el modelo como:

$$\mathbf{y}_t = \mathbf{x}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$$

Estimación de un VAR(p) (2/3)

Si suponemos que los errores son normales (multivariados):

$$\varepsilon_t \sim \text{i.i.d. } \square(\mathbf{0}, \mathbf{\Omega}), \quad t = p + 1, \dots, T$$

y que las observaciones $(\mathbf{y}_{-p+1}, \dots, \mathbf{y}_0)$ son conocidas, la función de verosimilitud (condicional) es:

$$L(\boldsymbol{\beta}, \mathbf{\Omega}) = (2\pi)^{-\frac{nk}{2}} |\mathbf{\Omega}|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=p+1}^T (\mathbf{y}_t - \mathbf{x}_t \boldsymbol{\beta})' \mathbf{\Omega}^{-1} (\mathbf{y}_t - \mathbf{x}_t \boldsymbol{\beta}) \right\}; \quad n = T - p$$

La función de log-verosimilitud (condicional) es:

$$\ell(\boldsymbol{\beta}, \mathbf{\Omega}) = -\frac{nk}{2} \ln(2\pi) - \frac{n}{2} \ln |\mathbf{\Omega}| - \frac{1}{2} \sum_{t=p+1}^T (\mathbf{y}_t - \mathbf{x}_t \boldsymbol{\beta})' \mathbf{\Omega}^{-1} (\mathbf{y}_t - \mathbf{x}_t \boldsymbol{\beta}).$$

Estimación de un VAR(p) (3/3)

A partir de lo anterior, podemos encontrar que el estimador de máxima verosimilitud es:

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{t=p+1}^T \mathbf{x}_t' \mathbf{x}_t \right)^{-1} \left(\sum_{t=p+1}^T \mathbf{x}_t' \mathbf{y}_t \right).$$

Este estimador coincide con el *estimador de Mínimos Cuadrados Ordinarios (MCO) ecuación por ecuación*.

Por otra parte, el estimador de máxima verosimilitud para la matriz de varianzas y covarianzas de los errores es:

$$\widehat{\boldsymbol{\Omega}} = \frac{1}{n} \sum_{t=p+1}^T \widehat{\boldsymbol{\varepsilon}}_t \widehat{\boldsymbol{\varepsilon}}_t', \quad \widehat{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \mathbf{x}_t \widehat{\boldsymbol{\beta}}.$$

Causalidad de Granger (1/3)

La **causalidad de Granger** evalúa si los valores pasados de una variable y_{2t} ayudan a predecir a otra variable y_{1t} , más allá de la información contenida en los rezagos de y_{1t} misma.

Formalmente, decimos que y_{2t} *no causa a y_{1t} en el sentido de Granger si:*

$$\mathbb{E}(y_{1t} \mid \Omega_{t-1}) = \mathbb{E}(y_{1t} \mid \Omega_{t-1} \setminus y_{2,t-1}, y_{2,t-2}, \dots),$$

donde Ω_{t-1} es el conjunto de toda la información disponible hasta $t - 1$.

Causalidad de Granger (2/3)

Para un VAR(p) bivariado tenemos:

$$y_{1t} = c_1 + \sum_{i=1}^p \phi_{11,i} y_{1,t-i} + \sum_{i=1}^p \phi_{12,i} y_{2,t-i} + \varepsilon_{1t},$$
$$y_{2t} = c_2 + \sum_{i=1}^p \phi_{21,i} y_{1,t-i} + \sum_{i=1}^p \phi_{22,i} y_{2,t-i} + \varepsilon_{2t}.$$

Entonces:

- y_{2t} no causa a y_{1t} en el sentido de Granger *si y solo si* $\phi_{12,1} = \dots = \phi_{12,p} = 0$.
- y_{1t} no causa a y_{2t} en el sentido de Granger *si y solo si* $\phi_{21,1} = \dots = \phi_{21,p} = 0$.

Causalidad de Granger (3/3)

Podemos verificar la causalidad de Granger empíricamente. Para esto, planteamos:

$$H_0 : \phi_{ij,1} = \phi_{ij,2} = \cdots = \phi_{ij,p} = 0, \quad i \neq j.$$

Usamos la prueba F:

$$F = \frac{(SSR_R - SSR_{UR})/q}{SSR_{UR}/(T - k)},$$

donde:

- SSR_R : suma de residuos al cuadrado del modelo restringido (sin rezagos de y_2)
- SSR_{UR} : suma de residuos del modelo no restringido
- $q = p$: número de restricciones.

Vectores autorregresivos: Ejemplo

Estimamos un VAR(4) para el inicio y el fin de construcción de nuevas viviendas en los Estados Unidos.

```
##  
## Call:  
## lm(formula = y ~ -1 + ., data = datamat)  
##  
## Coefficients:  
## starts.l1  comps.l1  starts.l2  comps.l2  starts.l3  comps.l3  starts.l4  
##  0.655743  0.007276  0.235371 -0.118989  0.138673 -0.021141  0.011877  
##  comps.l4      const  
##  0.002711  0.131220
```

```
##  
## Call:  
## lm(formula = y ~ -1 + ., data = datamat)  
##  
## Coefficients:  
## starts.l1  comps.l1  starts.l2  comps.l2  starts.l3  comps.l3  starts.l4  
##  0.06802  0.20486  0.05966  0.21775  0.03358  0.10504  0.09309  
##  comps.l4      const  
##  0.18231  0.04524
```

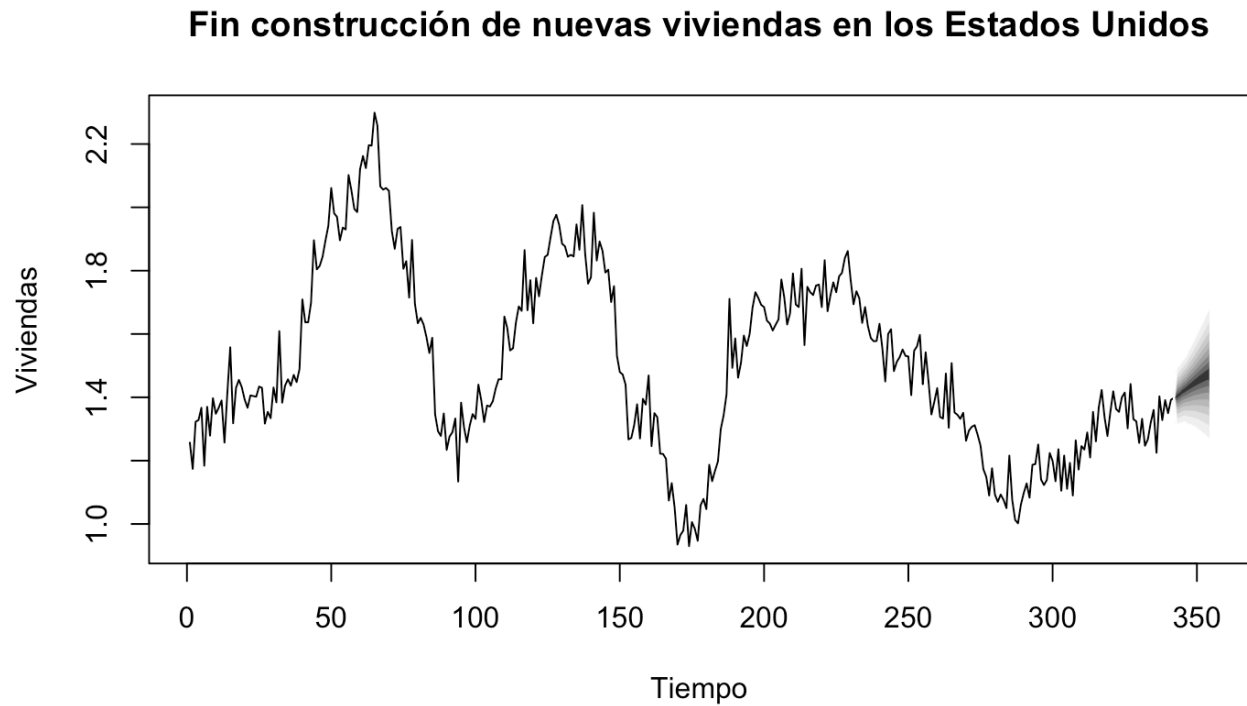
Vectores autorregresivos: predicción (1/2)

Para predecir el futuro usando un modelo VAR(p), usamos exactamente las mismas técnicas que para los modelos ARMA(p, q). Usamos la muestra 1968:01-1991:12 para estimar el modelo y pronosticamos el periodo 1992:01-1996:06. Para el inicio de construcción de nuevas viviendas tenemos:



Vectores autorregresivos: predicción (2/2)

Para el fin de construcción de nuevas viviendas tenemos:



Vectores autorregresivos: causalidad de Granger (1/2)

La siguiente tabla muestra el resultado del test de Granger para el inicio de construcción de viviendas:

```
## $Granger
##
##   Granger causality H0: comps do not Granger-cause starts
##
## data:  VAR object var_fit
## F-Test = 2.3066, df1 = 4, df2 = 658, p-value = 0.05685
##
##
## $Instant
##
##   H0: No instantaneous causality between: comps and starts
##
## data:  VAR object var_fit
## Chi-squared = 11.478, df = 1, p-value = 0.0007043
```

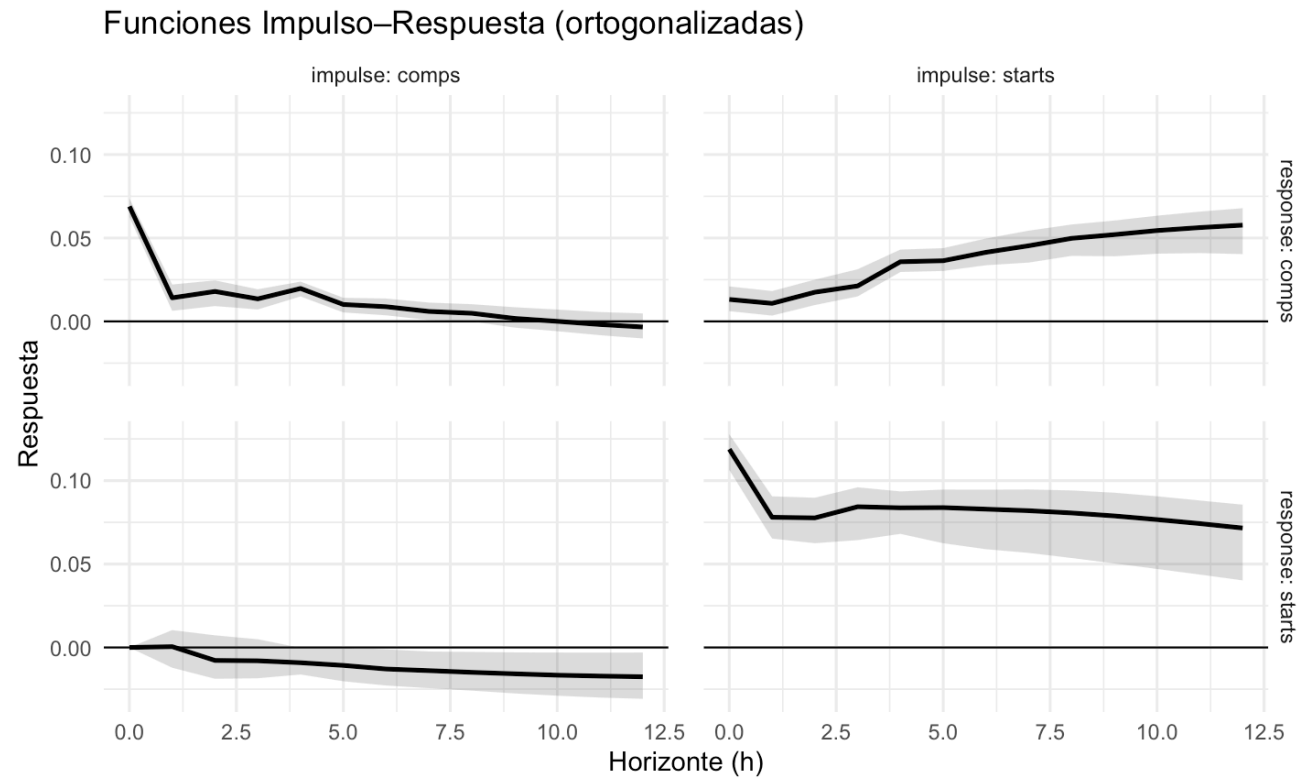
Vectores autorregresivos: causalidad de Granger (2/2)

El test de Granger para el fin de construcción de viviendas:

```
## $Granger
##
## Granger causality H0: starts do not Granger-cause comps
##
## data: VAR object var_fit
## F-Test = 32.125, df1 = 4, df2 = 658, p-value < 2.2e-16
##
##
## $Instant
##
## H0: No instantaneous causality between: starts and comps
##
## data: VAR object var_fit
## Chi-squared = 11.478, df = 1, p-value = 0.0007043
```

Funciones de impulso-Respuesta

Las funciones de impulso-respuesta para nuestro ejemplo son:



Descomposición de la varianza

Otra forma de caracterizar la dinámica de un VAR —muy relacionada a las funciones impulso-respuesta, es la **descomposición de la varianza**. Esta nos permite analizar el error de predicción en cada variable a un determinado horizonte de tiempo. Este error se atribuye a la incertidumbre acerca de la evolución futura en cada una de las variables.

