



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
 FACULTAD DE CIENCIAS ECONÓMICAS Y  
 ADMINISTRATIVAS  
 INSTITUTO DE ECONOMÍA  
 TEORÍA ECONOMETRICA III, EAE3513-1.

## Tarea 2.

Oscar Herrera

Segundo Semestre de 2024.

### Pregunta 1.1.a

#### Problemas econométricos de estimaciones de VAR reducido.

El principal problema econométrico al obtener estimadores de un VAR en forma reducida tiene relación con el método de estimación generalmente empleado: **Máxima Verosimilitud**. Esta metodología asume una forma funcional para la función de densidad conjunta de la muestra, lo que a su vez requiere conocer los parámetros de la función de densidad. Debido a que trabajamos con un conjunto de variables en el sistema, las funciones de densidad imponen la necesidad de conocer la matriz de varianza-covarianza de shocks contemporáneos de este sistema<sup>1</sup>, es decir, se requiere estimar la matriz varianza-covarianza de los residuos en forma reducida.

Por ejemplo, sea el siguiente VAR(1) estacionario en covarianza:

$$\underbrace{\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix}}_{\equiv \mathbf{x}_t} = \underbrace{\begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}}_{\equiv \Phi} \underbrace{\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix}}_{\equiv \mathbf{x}_{t-1}} + \underbrace{\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}}_{\equiv \mathbf{u}_t}$$

con  $\mathbf{u}_t \sim N(\mathbf{0}, \Omega)$  y  $\Omega = \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_{1,2}}^2 \\ \sigma_{u_{2,1}}^2 & \sigma_{u_2}^2 \end{bmatrix}$  la matriz varianza-covarianza de los residuos contemporáneos en forma reducida. Entonces, la función conjunta de verosimilitud condicional en la primera observación está dada por:

$$\begin{aligned} f(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1 | \mathbf{x}_0, \theta) &= \prod_{t=1}^T f(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1 | \mathbf{x}_0, \theta) \\ &= \prod_{t=1}^T N(\Phi \mathbf{x}_{t-1}, \Omega) \end{aligned}$$

<sup>1</sup>Es decir, si tenemos un sistema de orden  $k$ , se requiere conocer la covarianza entre los shocks de  $x_{i,t}$  y  $x_{1,t}$ , de  $x_{i,t}$  y  $x_{2,t}$ , de  $x_{i,t}$  y  $x_{3,t}$ , ...,  $x_{i,t}$  y  $x_{k,t}$ ,  $\forall i \in (1, 2, \dots, k)$  &  $\forall t \in (1, 2, \dots, T)$ . Al trabajar con VARs estacionarios, esta matriz es la misma  $\forall t$ .

con  $\theta = [\phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}, \mathbf{\Omega}]$ .

Luego, tomando logaritmo a la función de verosimilitud obtenemos la *log-likelihood*:

$$\begin{aligned}\mathcal{L}(\theta) &= \mathcal{L}(\phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}, \mathbf{\Omega}) = \sum_{t=1}^T \log N(\Phi \mathbf{x}_{t-1}, \mathbf{\Omega}) \\ &= -\frac{Tn}{2} \ln(2\pi) - \frac{T}{2} \ln |\mathbf{\Omega}| - \frac{1}{2} \sum_{t=1}^T \mathbf{u}_t' \mathbf{\Omega}^{-1} \mathbf{u}_t\end{aligned}$$

Luego, es fácil ver que para obtener el set de estimadores  $\{\hat{\phi}_{11}, \hat{\phi}_{12}, \hat{\phi}_{21}, \hat{\phi}_{22}\}$  se requiere conocer o estimar la matriz varianza-covarianza  $\mathbf{\Omega}$ .

Se puede mostrar (Soto, 2024) que, a partir de la función *log-likelihood*, el estimador para  $\mathbf{\Omega}$  está dado por:

$$\hat{\mathbf{\Omega}} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{u}}_t' \hat{\mathbf{u}}_t$$

Sin embargo, la estimación de residuos es a su vez una función de los coeficientes estimados (simultaniedad), ya que:

$$\mathbf{u}_t = \mathbf{x}_t - \Phi \mathbf{x}_{t-1} \quad \Leftrightarrow \quad \hat{\mathbf{u}}_t = \mathbf{x}_t - \hat{\Phi} \mathbf{x}_{t-1}$$

Por lo tanto, la estimación de un VAR en forma reducida requiere estimar simultáneamente los residuos, su matriz varianza-covarianza y los coeficientes. En otras palabras, la estimación para  $\Phi$  es el resultado de maximizar la función objetivo para una cierta volatilidad  $\mathbf{\Omega}$ . De no estimar simultáneamente la matriz  $\mathbf{\Omega}$ , entonces el sistema *log-likelihood* se indetermina y no tiene solución.

En la práctica, dado un *guess* inicial de los coeficientes (usualmente la estimación resultante por OLS), entonces los residuos terminan siendo una función de los coeficientes. Luego, los residuos son re-calculados y derivados a partir de un proceso iterativo:

- Dado un *guess* inicial se obtienen los residuos.
- Estos residuos se usan para re-estimar los coeficientes.
- Esta nueva de estimación de coeficientes mapea una nueva estimación de residuos.
- Así sucesivamente hasta encontrar convergencia/maximización de  $\mathcal{L}(\hat{\theta})$ .

Recalcar, en vista del el algoritmo iterativo en cuestión esta vez, que la no estimación de  $\Omega$  implica la imposibilidad de evaluar la función objetivo *log-likelihood* en el proceso iterativo, y por ende, la indeterminación del sistema como ya fue mencionado.

Todo lo anterior se traduce a los **problemas econométricos** de *i)* dependencia de supuestos fuertes (asunción de una función de densidad conjunta específica), *ii)* incertidumbre paramétrica en torno a  $\hat{\Omega}$ , y *iii)* convergencia computacional costosa.

Otros problemas econométricos asociados a la estimación del VAR reducido son la elección del orden  $p$  del VAR( $p$ ) y, luego de escoger el orden, el posterior aseguramiento de la estabilidad del sistema una vez estimado.

En el primer caso, la problemática tiene mayor relación con la complejidad comuptacional. La elección del orden autorregresivo es usualmente implementada a través de la múltiple estimación de sistemas, una estimación para cada orden, y el subsiguiente cálculo de criterios informativos para cada estimación, con el fin de poder determinar el orden adecuado. Este cálculo puede ser costoso en si mismo.

En el segundo caso, la problemática tiene mayor relación con decisiones procedimentales que recaen en el investigador. Una vez escogido el largo de un VAR, ya sea mediante criterios de información o basado en teoría económica, nada garantiza que la estimación resultante produzca un VAR estable. En dicho caso, se requieren decisiones adicionales fundamentadas en el criterio del investigador.

## Recuperación del VAR estructural a partir del VAR reducido.

Podemos decir que siempre y cuando el DGP subyacente presente una relación de simultaneidad, entonces no podremos recuperar el VAR estructural a partir del reducido, independiente del orden de este<sup>2</sup>. Tal como establece Soto (2024), “Nunca será posible “identificar” un modelo de ecuaciones **simultáneas** dinámico. Este resultado es general, no importando el número de variables incluidas en el VAR ni su orden (número de rezagos)”. Adicionalmente, se puede mostrar que el aumento del número de variables o del número de rezagos del sistema, aumentará la cantidad de parámetros a estimar a una velocidad mayor que el aporte adicional de la nueva información, por lo que este *approach* agravaría el problema.

Lo anterior se cumple **si y sólo si** el VAR estructural presenta **simultaniedad**. En otras palabras, existe un caso particular en el que podremos recuperar el VAR estructural mediante el reducido, esto es cuando el VAR estructural mismo no presenta simultaneidad. En este caso, se puede mostrar (a continuación) que el VAR reducido corresponde al VAR estructural.

## VAR estructural con simultaneidad.

Siguiendo el approach de Soto (2024), sea el siguiente VAR(1) estructural con relación de simultaneidad entre las variables endógenas  $y_t$  y  $z_t$ :

---

<sup>2</sup>Para todo  $n \geq 2$ .

$$\begin{aligned}
y_t &= \beta_{10} - \beta_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_t^y \\
z_t &= \beta_{20} - \beta_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_t^z
\end{aligned}$$

Reordenado términos correspondientes al periodo  $t$  en el lado izquierdo:

$$y_t + \beta_{12}z_t = \beta_{10} + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_t^y \quad (1)$$

$$z_t + \beta_{21}y_t = \beta_{20} + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_t^z \quad (2)$$

Compactando (1) y (2) en notación matricial:

$$\underbrace{\begin{bmatrix} 1 & \beta_{12} \\ \beta_{21} & 1 \end{bmatrix}}_{\equiv B} \underbrace{\begin{bmatrix} y_t \\ z_t \end{bmatrix}}_{\equiv x_t} = \underbrace{\begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}}_{\equiv \Gamma_0} + \underbrace{\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}}_{\equiv \Gamma_1} \underbrace{\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix}}_{\equiv x_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}}_{\equiv \varepsilon_t}$$

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$

Pre-multiplicando por  $B^{-1}$  a ambos lados:

$$x_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 x_{t-1} + B^{-1}\varepsilon_t \quad (3)$$

Luego, podemos definir (3) como:

$$x_t = A_0 + A_1 x_{t-1} + e_t \quad (4)$$

donde  $A_0 = B^{-1}\Gamma_0$ ,  $A_1 = B^{-1}\Gamma_1$ , y  $B^{-1}\varepsilon_t = e_t$ .

Notar que (4) es de la forma:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_t^1 \\ e_t^2 \end{bmatrix}$$

Luego, para recuperar el VAR estructural se debe resolver:

$$\begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} = \begin{bmatrix} 1 & \beta_{12} \\ \beta_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}; \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & \beta_{12} \\ \beta_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$$

Notar que la estimación del VAR reducido entrega los siguientes parámetros:  $\{\hat{a}_{10}, \hat{a}_{20}, \hat{a}_{11}, \hat{a}_{12}, \hat{a}_{21}, \hat{a}_{22}, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\sigma}_{1,2}^2\}$ , mientras que la identificación del VAR estructural depende de los siguientes parámetros:  $\{\beta_{10}, \beta_{20}, \beta_{12}, \beta_{21}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, \sigma_y^2, \sigma_z^2\}$ . Como se poseen más incógnitas que variables, no es posible resolver el sistema que recupera el VAR estructural.

### VAR estructural sin simultaneidad.

Sea el siguiente VAR(1) estructural sin relación de simultaneidad entre las variables endógenas  $y_t$  y  $z_t$ :

$$\begin{aligned} y_t &= \beta_{10} + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_t^y \\ z_t &= \beta_{20} + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_t^z \end{aligned}$$

Compactando en notación matricial:

$$\underbrace{\begin{bmatrix} y_t \\ z_t \end{bmatrix}}_{\equiv x_t} = \underbrace{\begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}}_{\equiv \Gamma_0} + \underbrace{\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}}_{\equiv \Gamma_1} \underbrace{\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix}}_{\equiv x_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}}_{\equiv \varepsilon_t}$$

$$x_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$

Es fácil ver que este caso particular el VAR reducido (sin simultaneidad) corresponde a la vez al VAR estructural.

En la práctica, los casos de interés a estimar corresponden a VARs con simultaneidad toda vez que frecuentemente estimaciones de VARs reducido **arrojan** correlación contemporánea de los shocks del sistema. Es decir, para determinar si en la *praxis* nos encontramos con el caso particular donde el VAR reducido es a la vez el VAR estructural, se debe evaluar si los shocks reducidos contemporáneos **no presentan** correlación. Sin embargo, empíricamente este es un caso particular **bastante improbable**. Por ende, en la práctica no podremos recuperar el VAR estructural a partir del reducido.

## Pregunta 1.1.b

En línea con lo mencionado en la pregunta anterior, podemos pensar el **problema de simultaneidad** como el hallazgo de correlación contemporánea de los shocks una vez estimado el sistema. Así, la resolución del problema de simultaneidad —shocks estructurales propios de una variable que afectan a todo el sistema— corresponde al mapeo de shocks reducidos a shocks de un VAR estructural, donde por definición estos últimos no presentan correlación contemporánea. Por lo tanto, un VAR identificado à la Sims (Cholesky), elimina el problema de simultaneidad ya que permite mapear de un VAR reducido a un VAR estructural, a través de supuestos que se traducen en la imposición de restricciones adicionales al sistema.

### Identificación de Cholesky para VAR(1) con n=2.

A continuación, se muestra como Cholesky resuelve para un VAR bivariado. Sin embargo, la extensión para VARs de dimensiones superiores resulta natural.

Sea el VAR(1) estructural (3) (o equivalentemente, (4)):

$$x_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1x_{t-1} + B^{-1}\varepsilon_t \quad (3)$$

$$x_t = A_0 + A_1x_{t-1} + e_t \quad (4)$$

Para resolver la simultaneidad se debe resolver:  $B^{-1}\varepsilon_t = e_t \implies B \cdot \varepsilon_t = e_t$ . Una vez estimado el VAR, el problema se condensa en encontrar la matriz  $B$ .

Por un lado, la matriz varianza-covarianza de los residuos reducidos corresponde a:

$$\Sigma_e = \mathbb{E}[e_te_t'] = \mathbb{E}[B\varepsilon_t(B\varepsilon_t)'] = B\mathbb{E}[\varepsilon_t\varepsilon_t']B' = B\Sigma_\varepsilon B'$$

Por otro lado, por definición los shocks estructurales no están correlacionados. Asumimos generalmente  $\Sigma_\varepsilon = \mathbf{I}_n$ . Usando la definición y el supuesto usual obtenemos:

$$B\Sigma_\varepsilon B' = BB' = \Sigma_e$$

$$BB' = \Sigma_e$$

Por ende, queda demostrado que la matriz de impacto  $B$  del VAR reducido depende directamente de la matriz varianza-covarianza de los shocks reducidos. El problema entonces, una vez estimada la matriz varianza-covarianza de los residuos reducidos, se reduce a encontrar la matriz  $B$  que satisface  $BB' = \Sigma_e$ .

Ahora bien, usando descomposición de Cholesky para la matriz varianza-covarianza del VAR(1) con  $n = 2$ :

$$\Sigma_e = BB' \implies \begin{bmatrix} \sigma_1^2 & \sigma_{1,2}^2 \\ \sigma_{1,2}^2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{2,1} \\ b_{1,2} & b_{2,2} \end{bmatrix}$$

Reescribiendo el sistema:

$$\begin{cases} \sigma_1^2 = b_{11}^2 + b_{12}^2 \\ \sigma_{12}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_{12}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_2^2 = b_{21}^2 + b_{22}^2 \end{cases}$$

$$\iff \begin{cases} \sigma_1^2 = b_{11}^2 + b_{12}^2 \\ \sigma_{12}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_2^2 = b_{21}^2 + b_{22}^2 \end{cases}$$

Notar que por la simetría de  $\Sigma_e$  el sistema consiste en 3 ecuaciones y 4 incógnitas; no tiene solución (o en estricto rigor, posee infinitas soluciones).

La **identificación por descomposición de Cholesky** consiste en descomponer la matriz  $\Sigma_e$  en un producto de matriz triangular inferior con su transpuesta (matriz triangular superior) (Vicondoa, 2023):  $\Sigma_e = PP' \implies P = B$ . Esto, ya que sabemos que las matrices de varianzas cumplen con esta caracterización.

Esta caracterización impone restricciones adicionales contemporáneas (*zero contemporaneous restrictions*), lo que reduce las incógnitas del sistema. Por ende, el sistema tendrá solución única:

$$\Sigma_e = BB' \implies \begin{bmatrix} \sigma_1^2 & \sigma_{1,2}^2 \\ \sigma_{1,2}^2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} b_{1,1} & 0 \\ b_{2,1} & b_{2,2} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{2,1} \\ 0 & b_{2,2} \end{bmatrix}$$

Lo que nos lleva al siguiente sistema:

$$\begin{cases} \sigma_1^2 = b_{11}^2 \\ \sigma_{12}^2 = b_{11}b_{21} \\ \sigma_2^2 = b_{21}^2 + b_{22}^2 \end{cases}$$

Se puede mostrar que la solución al sistema es la siguiente:

$$\begin{cases} b_{11} = \sigma_1 \\ b_{21} = \frac{\sigma_{12}^2}{\sigma_1} \\ b_{22} = \sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}^2}{\sigma_1}\right)^2} \end{cases}$$

Por lo tanto, la identificación del VAR estructural y la resolución del problema de simultaneidad<sup>3</sup> es la siguiente:

$$\begin{aligned} B^{-1}\varepsilon_t &= \begin{bmatrix} \sigma_1 & 0 \\ \frac{\sigma_{12}^2}{\sigma_1} & \sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}^2}{\sigma_1}\right)^2} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} = e_t = \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix} \\ BB^{-1}\varepsilon_t &= \varepsilon_t = \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} = Be_t = \begin{bmatrix} \sigma_1 & 0 \\ \frac{\sigma_{12}^2}{\sigma_1} & \sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}^2}{\sigma_1}\right)^2} \end{bmatrix} \begin{bmatrix} e_t^1 \\ e_t^2 \end{bmatrix} \\ \varepsilon_t &= \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \frac{\sigma_{12}^2}{\sigma_1} & \sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}^2}{\sigma_1}\right)^2} \end{bmatrix} \begin{bmatrix} e_t^1 \\ e_t^2 \end{bmatrix} \end{aligned}$$

Así, queda demostrado que Cholesky elimina el problema de simultaneidad.

### Identificación de Cholesky para VAR(1) con n=3.

A continuación, se muestra como Cholesky resuelve para un VAR “tri-variado”. Nuevamente, la extensión para VARs de dimensiones superiores resulta natural.

Siguiendo el *approach* de Soto (2024), sea el siguiente VAR para  $y_t$ ,  $z_t$  y  $w_t$ :

$$\begin{bmatrix} y_t \\ z_t \\ w_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \\ w_{t-1} \end{bmatrix} + \begin{bmatrix} e_t^1 \\ e_t^2 \\ e_t^3 \end{bmatrix}$$

Por el teorema de la representación de Wold, todo VAR(1) estable puede ser representado como un proceso  $VMA(\infty)$  (Vicondoa, 2023):

---

<sup>3</sup>Entendida esta como el *mapping* a shocks ortogonales.



$$x_t = \Phi^t x_0 + \sum_{j=0}^{t-1} \Phi^j u_{t-j}$$

$$x_t = \Phi^\infty x_{t-\infty} + \sum_{j=0}^{\infty} \Phi^j u_{t-j}$$

Luego, la representación media móvil del VAR(1) de  $y_t$ ,  $z_t$  y  $w_t$  es:

$$\begin{bmatrix} y_t \\ z_t \\ w_t \end{bmatrix} = \begin{bmatrix} \bar{a}_{10} \\ \bar{a}_{20} \\ \bar{a}_{30} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^i \begin{bmatrix} e_{t-i}^1 \\ e_{t-i}^2 \\ e_{t-i}^3 \end{bmatrix}$$

Imponemos la descomposición de Cholesky:

$$\begin{bmatrix} y_t \\ z_t \\ w_t \end{bmatrix} = \begin{bmatrix} \bar{a}_{10} \\ \bar{a}_{20} \\ \bar{a}_{30} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^i \begin{bmatrix} \beta_{11} & 0 & 0 \\ \beta_{21} & \beta_{22} & 0 \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \\ \varepsilon_t^w \end{bmatrix}$$

Así, queda resuelta la simultaneidad para un VAR(1) con  $n = 3$ .

Lo anterior se extiende a dimensiones superiores:  $n > 3$ . La extensión corresponde a añadir un mayor número de *zero contemporaneous restrictions*, con el fin de establecer una solución única al problema de *mapping*.

Notar que a medida que se imponen mayor número de restricciones, mayor es la importancia del orden del VAR, el cual recae en los supuestos establecidos por parte del investigador.

## Identificación de Cholesky para VAR(p).

Se puede mostrar que todo VAR( $p$ ) puede ser reescrito como un VAR(1) mediante la “*companion representation*” (Vicondoa, 2023).

Siguiendo el ejemplo de un VAR(2) de “SVAR Models” (Vicondoa, 2023):

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11}^1 & \phi_{12}^1 \\ \phi_{21}^1 & \phi_{22}^1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^2 & \phi_{22}^2 \end{bmatrix} \begin{bmatrix} y_{t-2} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\text{Demand}} \\ \varepsilon_t^{\text{MonPol}} \end{bmatrix}$$

Reescribiendo el VAR(2):

$$\begin{bmatrix} y_t \\ r_t \\ y_{t-1} \\ r_{t-1} \end{bmatrix} = \underbrace{\begin{bmatrix} \phi_{11}^1 & \phi_{12}^1 & \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^1 & \phi_{22}^1 & \phi_{21}^2 & \phi_{22}^2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{\text{Companion matrix } \tilde{\Phi}} \begin{bmatrix} y_{t-1} \\ r_{t-1} \\ y_{t-2} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\text{Demand}} \\ \varepsilon_t^{\text{MonPol}} \end{bmatrix}.$$

Entonces el VAR(2) corresponde a su vez a un VAR(1) con companion matrix  $\tilde{\Phi}$ :

$$\tilde{x}_t = \tilde{\Phi} \tilde{x}_{t-1} + \tilde{B} \varepsilon_t.$$

Luego, toda vez que un VAR( $p$ ) arbitrario puede ser representado como VAR(1), habiendo demostrado que para todo VAR(1) Cholesky resuelve el problema de simultaneidad, entonces Cholesky resuelve el **problema** de simultaneidad para todo VAR( $p$ ).

## Pregunta 1.1.c

La afirmación es parcialmente correcta. Se entiende por variable exógena como una variable que sólo se ve afectada por shocks estructurales propios contemporáneos, i.e., ningún otro shock la afecta a través de la matriz de impacto. Por lo tanto, según esta definición, la variable será exógena sólo si **toda** la incertidumbre de su predicción **1 periodo hacia adelante** proviene de innovaciones propias. En periodos sucesivos, si la variable es determinada por rezagos de las otras variables en la matriz dinámica, entonces su predicción será afectada por shocks aleatorios del pasado de otras variables (i.e., la predicción de la variable exógena estará afectada por si misma para  $h > 1$  si no se determina con lags de otras variables).

## Forecast Error Variance Decomposition de una variable exógena.

Sea el Forecast Error el cambio de la variable entre  $t-1$  and  $t+h$  a causa de realizaciones de shocks estructurales que no se podrían haber predecido. Adicionalmente, en línea con la notación usada en “SVAR Models, Vicondoa, 2023”, sea  $x_t$  un vector  $n \times 1$  y  $\mathbf{B}$  la matriz de impacto. Entonces, el error en  $t$  estará dado por:

$$x_t - \mathbb{E}[x_t \mid \Omega_{t-1}] = (\Phi x_{t-1} + \mathbf{B} \varepsilon_t) - (\Phi x_{t-1}) = \mathbf{B} \varepsilon_t$$

El error en  $t+1$  estará dado por:

$$x_{t+1} - \mathbb{E}[x_{t+1} \mid \Omega_{t-1}] = (\Phi x_t + \mathbf{B} \varepsilon_{t+1}) - (\Phi x_t) = \Phi \mathbf{B} \varepsilon_t + \mathbf{B} \varepsilon_{t+1}$$

El error en  $t + 2$  estará dado por:

$$\begin{aligned}
x_{t+2} - \mathbb{E}[x_{t+2} \mid \Omega_{t-1}] &= (\Phi x_{t+1} + \mathbf{B}\varepsilon_{t+2}) - \Phi x_{t+1} \\
&= \Phi(\Phi x_t + \mathbf{B}\varepsilon_{t+1}) + \mathbf{B}\varepsilon_{t+2} - \Phi^2 x_t \\
&= \Phi(\Phi(\Phi x_{t-1} + \mathbf{B}\varepsilon_t) + \mathbf{B}\varepsilon_{t+1}) + \mathbf{B}\varepsilon_{t+2} - \Phi^3 x_{t-1} \\
&= \Phi^3 x_{t-1} + \Phi^2 \mathbf{B}\varepsilon_t + \Phi \mathbf{B}\varepsilon_{t+1} + \mathbf{B}\varepsilon_{t+2} - \Phi^3 x_{t-1} \\
&= \Phi^2 \mathbf{B}\varepsilon_t + \Phi \mathbf{B}\varepsilon_{t+1} + \mathbf{B}\varepsilon_{t+2}
\end{aligned}$$

Generalizando para  $h + 1$  periodos adelante (siguiendo Vicondoa, 2023):

$$x_{t+j} - \mathbb{E}[x_{t+j} \mid \Omega_{t-1}] = \sum_{i=0}^j \Phi^{j-i} \mathbf{B}\varepsilon_{t+i}$$

Luego, *variance of forecast error* del SVAR (siguiendo Vicondoa, 2023) está dada por:

$$\begin{aligned}
\text{Var}(\text{FE}_{t+j}) &= \text{Var}(x_{t+j} - \mathbb{E}[x_{t+j} \mid \Omega_{t-1}]) \\
&= \text{Var}\left(\sum_{i=0}^j \Phi^{j-i} \mathbf{B}\varepsilon_{t+i}\right)
\end{aligned}$$

Considerando un *Forecast error* para un VAR  $n = 2$  con  $x_t = [y_t ; r_t]$  (siguiendo el ejemplo de Vicondoa, 2023):

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11}^1 & \phi_{12}^1 \\ \phi_{21}^1 & \phi_{22}^1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^2 & \phi_{22}^2 \end{bmatrix} \begin{bmatrix} y_{t-2} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\text{Demand}} \\ \varepsilon_t^{\text{MonPol}} \end{bmatrix}$$

Luego, retomando la expresión del SVAR:

$$\begin{aligned}
\text{FE}_{t+j} &= \begin{bmatrix} x_{t+j}^1 - \mathbb{E}[x_{t+j}^1 \mid \Omega_{t-1}] \\ x_{t+j}^2 - \mathbb{E}[x_{t+j}^2 \mid \Omega_{t-1}] \end{bmatrix} \\
&= \sum_{i=0}^j \Phi^{j-i} \mathbf{B}\varepsilon_{t+i} \\
&= \sum_{i=0}^j \Phi^{j-i} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{t+i}^1 \\ \varepsilon_{t+i}^2 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^j \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}^{j-i} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{t+i}^1 \\ \varepsilon_{t+i}^2 \end{bmatrix} \\
&= \sum_{i=0}^j \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}^{j-i} \begin{bmatrix} b_{11}\varepsilon_{t+i}^1 + b_{12}\varepsilon_{t+i}^2 \\ b_{21}\varepsilon_{t+i}^1 + b_{22}\varepsilon_{t+i}^2 \end{bmatrix}
\end{aligned}$$

Luego, para  $x_t$ :

$$\text{Var}(x_t - \mathbb{E}[x_t \mid \Omega_{t-1}]) = \text{Var}(\mathbf{B}\boldsymbol{\varepsilon}_t) = \mathbf{B}'\text{Var}(\boldsymbol{\varepsilon}_t)\mathbf{B}'$$

Para la primera fila, suponiendo  $\mathbb{E}[\varepsilon_{t+i}\varepsilon_{t+i}'] = \mathbf{I}, \forall i$ :

$$\begin{aligned}
\text{Var}(x_t^1 - \mathbb{E}[x_t^1 \mid \Omega_{t-1}]) &= \text{Var}(b_{11}\varepsilon_t^1 + b_{12}\varepsilon_t^2) \\
&= b_{11}^2 \text{Var}(\varepsilon_t^1) + b_{12}^2 \text{Var}(\varepsilon_t^2) \\
&= b_{11}^2 + b_{12}^2
\end{aligned}$$

Luego, calculando la descomposición de varianza para cada variable, se obtiene:

$$\begin{aligned}
\text{VD } x_0^{1,\varepsilon^1} &= \frac{b_{11}^2}{b_{11}^2 + b_{12}^2}, & \text{VD } x_0^{1,\varepsilon^2} &= \frac{b_{12}^2}{b_{11}^2 + b_{12}^2} \\
\text{VD } x_0^{2,\varepsilon^1} &= \frac{b_{21}^2}{b_{21}^2 + b_{22}^2}, & \text{VD } x_0^{2,\varepsilon^2} &= \frac{b_{22}^2}{b_{21}^2 + b_{22}^2}
\end{aligned}$$

Luego, en  $t+1$ :

$$\begin{aligned}
\text{FE}_{t+1} &= \Phi\mathbf{B}\boldsymbol{\varepsilon}_t + \mathbf{B}\boldsymbol{\varepsilon}_{t+1} \\
&= \Phi \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{t+1}^1 \\ \varepsilon_{t+1}^2 \end{bmatrix} \\
&= \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} b_{11}\varepsilon_t^1 + b_{12}\varepsilon_t^2 \\ b_{21}\varepsilon_t^1 + b_{22}\varepsilon_t^2 \end{bmatrix} + \begin{bmatrix} b_{11}\varepsilon_{t+1}^1 + b_{12}\varepsilon_{t+1}^2 \\ b_{21}\varepsilon_{t+1}^1 + b_{22}\varepsilon_{t+1}^2 \end{bmatrix} \\
&= \begin{bmatrix} \phi_{11}(b_{11}\varepsilon_t^1 + b_{12}\varepsilon_t^2) + \phi_{12}(b_{21}\varepsilon_t^1 + b_{22}\varepsilon_t^2) \\ \phi_{21}(b_{11}\varepsilon_t^1 + b_{12}\varepsilon_t^2) + \phi_{22}(b_{21}\varepsilon_t^1 + b_{22}\varepsilon_t^2) \end{bmatrix} + \begin{bmatrix} b_{11}\varepsilon_{t+1}^1 + b_{12}\varepsilon_{t+1}^2 \\ b_{21}\varepsilon_{t+1}^1 + b_{22}\varepsilon_{t+1}^2 \end{bmatrix}
\end{aligned}$$

Luego, la varianza del error y “*Variance decomposition of forecast error*” en  $x_{t+1}^1$  es:

$$\begin{aligned}\text{Var}(x_{t+1}^1 - \mathbb{E}[x_{t+1}^1 | \Omega_{t-1}]) &= (\phi_{11}b_{11} + \phi_{12}b_{21})^2 + (\phi_{11}b_{12} + \phi_{12}b_{22})^2 + b_{11}^2 + b_{12}^2 \\ \text{VD } x_{t+1}^{1,\varepsilon^1} &= \frac{(\phi_{11}b_{11} + \phi_{12}b_{21})^2 + b_{11}^2}{(\phi_{11}b_{11} + \phi_{12}b_{21})^2 + (\phi_{11}b_{12} + \phi_{12}b_{22})^2 + b_{11}^2 + b_{12}^2} \\ \text{VD } x_{t+1}^{1,\varepsilon^2} &= \frac{(\phi_{11}b_{12} + \phi_{12}b_{22})^2 + b_{12}^2}{(\phi_{11}b_{11} + \phi_{12}b_{21})^2 + (\phi_{11}b_{12} + \phi_{12}b_{22})^2 + b_{11}^2 + b_{12}^2}\end{aligned}$$

Por otro lado, para  $x_{t+1}^2$ :

$$\begin{aligned}\text{Var}(x_{t+1}^2 - \mathbb{E}[x_{t+1}^2 | \Omega_{t-1}]) &= (\phi_{21}b_{11} + \phi_{22}b_{21})^2 + (\phi_{21}b_{12} + \phi_{22}b_{22})^2 + b_{21}^2 + b_{22}^2 \\ \text{VD } x_{t+1}^{2,\varepsilon^1} &= \frac{(\phi_{21}b_{11} + \phi_{22}b_{21})^2 + b_{21}^2}{(\phi_{21}b_{11} + \phi_{22}b_{21})^2 + (\phi_{21}b_{12} + \phi_{22}b_{22})^2 + b_{21}^2 + b_{22}^2} \\ \text{VD } x_{t+1}^{2,\varepsilon^2} &= \frac{(\phi_{21}b_{12} + \phi_{22}b_{22})^2 + b_{22}^2}{(\phi_{21}b_{11} + \phi_{22}b_{21})^2 + (\phi_{21}b_{12} + \phi_{22}b_{22})^2 + b_{21}^2 + b_{22}^2}\end{aligned}$$

Con *Cholesky decomposition*, es decir, con  $b_{12} = 0$ :

$$\text{FE}_{t+j} = \sum_{i=0}^j \Phi^{j-i} \mathbf{B} \varepsilon_{t+i} = \sum_{i=0}^j \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}^{j-i} \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{t+i}^1 \\ \varepsilon_{t+i}^2 \end{bmatrix}$$

Así, para la variable exógena se obtiene:

$$\begin{aligned}\text{VD } x_0^{\varepsilon^1} &= \frac{b_{11}^2}{b_{11}^2} = 1 \\ \text{VD } x_0^{\varepsilon^2} &= 0\end{aligned}$$

$$\begin{aligned}\text{VD } x_1^{\varepsilon^1} &= \frac{(\phi_{11}b_{11} + \phi_{12}b_{21})^2 + b_{11}^2}{(\phi_{11}b_{11} + \phi_{12}b_{21})^2 + (\phi_{11}b_{12} + \phi_{12}b_{22})^2 + b_{11}^2 + b_{12}^2} \\ \text{VD } x_1^{\varepsilon^2} &= \frac{(\phi_{12}b_{22})^2 + b_{12}^2}{(\phi_{11}b_{11} + \phi_{12}b_{21})^2 + (\phi_{11}b_{12} + \phi_{12}b_{22})^2 + b_{11}^2 + b_{12}^2}\end{aligned}$$

Podemos ver que esta variable estará afecta por sus shocks, pero por la dependendencia dinámica de otras variables estará afecta a shocks de otras variables a medida que pasa el tiempo. Por otro lado, la otra variable se verá expuesta a más shocks.

$$\begin{aligned} \text{VD } x_0^{\varepsilon^1} &= \frac{b_{21}^2}{b_{21}^2 + b_{22}^2} \\ \text{VD } x_0^{\varepsilon^2} &= \frac{b_{22}^2}{b_{21}^2 + b_{22}^2} \\ \text{VD } x_1^{\varepsilon^1} &= \frac{(\phi_{21}b_{11} + \phi_{22}b_{21})^2 + b_{21}^2}{(\phi_{21}b_{11} + \phi_{22}b_{21})^2 + (\phi_{21}b_{12} + \phi_{22}b_{22})^2 + b_{21}^2 + b_{22}^2} \\ \text{VD } x_1^{\varepsilon^2} &= \frac{(\phi_{22}b_{22})^2 + b_{22}^2}{(\phi_{21}b_{11} + \phi_{22}b_{21})^2 + (\phi_{21}b_{12} + \phi_{22}b_{22})^2 + b_{21}^2 + b_{22}^2} \end{aligned}$$

### Pregunta 1.1.d

En línea con las preguntas anteriores, una vez teniendo el VAR reducido tendremos:

$$\begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} = \begin{bmatrix} 1 & \beta_{12} \\ \beta_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}; \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & \beta_{12} \\ \beta_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}; \quad \begin{bmatrix} 1 & \beta_{12} \\ \beta_{21} & 1 \end{bmatrix}^{-1} \varepsilon_t = e_t$$

$$\begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}; \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}; \quad \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \varepsilon_t = e_t$$

Con la restricción propuesta obtendríamos:

$$\begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} = \begin{bmatrix} b_{1,1} & b_{1,2} \\ 0 & b_{2,2} \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}; \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{1,1} & b_{1,2} \\ 0 & b_{2,2} \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}; \quad \begin{bmatrix} b_{1,1} & b_{1,2} \\ 0 & b_{2,2} \end{bmatrix} \varepsilon_t = e_t$$

Así, las incógnitas serían:  $\{\beta_{10}, \beta_{20}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}\}$ . Todo el resto de variables serían conocidas. Esto conforma un sistema identificable. Por lo tanto, la restricción es suficiente para lograr la identificación.

### Pregunta 1.2.a

Sea un  $VAR(p)$  con el siguiente error de pronóstico óptimo  $h$  periodos adelante:

$$\mathbf{e}_t^f(h) = \Psi_{h-1}\varepsilon_{t+1} + \Psi_{h-2}\varepsilon_{t+2} + \cdots + \Psi_1\varepsilon_{t+h-1} + \varepsilon_{t+h}$$

Con error de cuadrático medio de proyección, con  $\mathbb{E}[\varepsilon_t \varepsilon_t^\top] = \Sigma$ , dado por:

$$\mathbb{E}[\mathbf{e}_t^f(h) \mathbf{e}_t^f(h)^\top] = \Psi_{h-1} \Sigma \Psi_{h-1}^\top + \Psi_{h-2} \Sigma \Psi_{h-2}^\top + \cdots + \Psi_1 \Sigma \Psi_1^\top + \Sigma$$

Desarrollando la expresión para el error de pronóstico traspuesto:

$$\begin{aligned} \mathbf{e}_t^f(h)^\top &= \left[ \Psi_{h-1} \varepsilon_{t+1} + \Psi_{h-2} \varepsilon_{t+2} + \cdots + \Psi_1 \varepsilon_{t+h-1} + \varepsilon_{t+h} \right]^\top \\ &= (\Psi_{h-1} \varepsilon_{t+1})^\top + (\Psi_{h-2} \varepsilon_{t+2})^\top + \cdots + (\Psi_1 \varepsilon_{t+h-1})^\top + \varepsilon_{t+h}^\top \\ &= \varepsilon_{t+1}^\top \Psi_{h-1}^\top + \varepsilon_{t+2}^\top \Psi_{h-2}^\top + \cdots + \varepsilon_{t+h-1}^\top \Psi_1^\top + \varepsilon_{t+h}^\top \end{aligned}$$

Luego, desarrollando la expresión para la multiplicación:

$$\begin{aligned} \mathbf{e}_t^f(h) \cdot \mathbf{e}_t^f(h)^\top &= \Psi_{h-1} \varepsilon_{t+1} (\varepsilon_{t+1}^\top \Psi_{h-1}^\top + \varepsilon_{t+2}^\top \Psi_{h-2}^\top + \cdots + \varepsilon_{t+h}^\top) \\ &\quad + \Psi_{h-2} \varepsilon_{t+2} (\varepsilon_{t+1}^\top \Psi_{h-1}^\top + \varepsilon_{t+2}^\top \Psi_{h-2}^\top + \cdots + \varepsilon_{t+h}^\top) \\ &\quad + \cdots \\ &\quad + \Psi_1 \varepsilon_{t+h-1} (\varepsilon_{t+1}^\top \Psi_{h-1}^\top + \varepsilon_{t+2}^\top \Psi_{h-2}^\top + \cdots + \varepsilon_{t+h}^\top) \\ &\quad + \varepsilon_{t+h} (\varepsilon_{t+1}^\top \Psi_{h-1}^\top + \varepsilon_{t+2}^\top \Psi_{h-2}^\top + \cdots + \varepsilon_{t+h}^\top) \end{aligned}$$

Lo anterior puede ser reescrito como:

$$\mathbf{e}_t^f(h) \mathbf{e}_t^f(h)^\top = \sum_{i=1}^h \sum_{j=1}^h \Psi_{h-i} \varepsilon_{t+i} \varepsilon_{t+j}^\top \Psi_{h-j}^\top,$$

Tomando esperanza a la expresión se obtiene:

$$\mathbb{E}[\mathbf{e}_t^f(h) \mathbf{e}_t^f(h)^\top] = \sum_{i=1}^h \sum_{j=1}^h \Psi_{h-i} \mathbb{E}[\varepsilon_{t+i} \varepsilon_{t+j}^\top] \Psi_{h-j}^\top.$$

Como  $\varepsilon_{t+i}$  sigue un proceso de ruido blanco, tenemos que:

$$\mathbb{E}[\varepsilon_{t+i}\varepsilon_{t+j}^\top] = \begin{cases} \Sigma, & \text{si } i = j, \\ \mathbf{0}, & \text{si } i \neq j. \end{cases}$$

Luego:

$$\mathbb{E}[\mathbf{e}_t^f(h)\mathbf{e}_t^f(h)^\top] = \sum_{i=1}^h \Psi_{h-i}\Sigma\Psi_{h-i}^\top.$$

Desarrollando la sumatoria, obtenemos la expresión final:

$$\mathbb{E}[\mathbf{e}_t^f(h)\mathbf{e}_t^f(h)^\top] = \Psi_{h-1}\Sigma\Psi_{h-1}^\top + \Psi_{h-2}\Sigma\Psi_{h-2}^\top + \cdots + \Psi_1\Sigma\Psi_1^\top + \Sigma$$

■

## Pregunta 1.2.b

Sea el VAR univariado:

$$y_t = 2y_{t-1} - \frac{3}{2}y_{t-2} + \varepsilon_t$$

Reexpresando el *VAR* en forma matricial:

$$\underbrace{\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}}_{\equiv \mathbf{y}_t} = \underbrace{\begin{bmatrix} 2 & -\frac{3}{2} \\ 1 & 0 \end{bmatrix}}_{\Theta} \underbrace{\begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix}}_{\equiv \mathbf{y}_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{bmatrix}}_{\varepsilon_t}$$

Resolviendo recursivamente:

$$\begin{aligned} \mathbf{y}_t &= \Theta \mathbf{y}_{t-1} + \varepsilon_t \\ &= \Theta (\Theta \mathbf{y}_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \Theta^2 \mathbf{y}_{t-2} + \Theta \varepsilon_{t-1} + \varepsilon_t \\ &= \Theta^3 \mathbf{y}_{t-3} + \Theta^2 \varepsilon_{t-2} + \Theta \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \\ \mathbf{y}_t &= \underbrace{\Theta^{t+1} \mathbf{y}_0}_{\rightarrow 0} + \Theta^t \varepsilon_t + \Theta^{t-1} \varepsilon_{t-1} + \cdots + \underbrace{\Theta^3}_{\equiv \Theta_3} \varepsilon_{t-3} + \underbrace{\Theta^2}_{\equiv \Theta_2} \varepsilon_{t-2} + \underbrace{\Theta}_{\equiv \Theta_3} \varepsilon_{t-1} + \varepsilon_t \end{aligned}$$



Reexpresando:

$$\mathbf{y}_t = \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + \Theta_3 \varepsilon_{t-3} + \Theta_4 \varepsilon_{t-4} + \dots$$

■

## Pregunta 1.2.c

Para determinar condiciones estacionariedad del  $VAR(1)$ , se determinará su media incondicional constante, y su estabilidad (siguiendo Sección 7: VAR Models, Vicondoa, 2023).

Sea el  $VAR(1)$ :

$$\mathbf{y}_t = \mu + \Phi_1 \mathbf{y}_{t-1} + \varepsilon_t$$

Tomando media incondicional para obtener la media para la estacionariedad, y sea  $\mathbb{E}[\mathbf{y}_t] \equiv \mu_y, \forall t$ :

$$\mu_y = \mu + \Phi_1 \mu_y.$$

Reordenando:

$$\mu_y = (\mathbf{I} - \Phi_1)^{-1} \mu.$$

Notar que la matriz  $(\mathbf{I} - \Phi_1)$  debe ser invertible, por lo tanto:

$$\det(\mathbf{I} - \Phi_1) \neq 0.$$

Luego, sea  $\Gamma_0 = \text{Cov}(\mathbf{y}_t, \mathbf{y}_t)$ , tomando operador de varianza incondicional a  $\mathbf{y}_t = \mu + \Phi_1 \mathbf{y}_{t-1} + \varepsilon_t$ , el  $VAR$  para que sea estable y que sus shocks se disipen en el tiempo debe satisfacer:

$$\Gamma_0 = \Phi_1 \Gamma_0 \Phi_1^\top + \Sigma_\varepsilon,$$

con  $\Sigma_\varepsilon = \text{Cov}(\varepsilon_t)$ .

Luego, para la estabilidad del sistema y que sus shocks se disipen a través del tiempo, se requiere que todos los *eigenvalues* de  $\Phi_1$  se encuentren dentro del círculo unitario en módulo.

$$\det(\Phi_1 - \lambda \mathbf{I}) = 0.$$

$$\implies |\lambda_i| < 1, \quad \forall i.$$

con  $\lambda_i$ , los *eigenvalues* de  $\Phi_1$ .

### Pregunta 1.3.a

Sea el invertible  $VAR(1)$  invertible:

$$(1) \quad \mathbf{y}_t^1 = \theta_1^1 \mathbf{y}_{t-1}^1 + \theta_1^2 \mathbf{y}_t^2 + \varepsilon_t^1,$$

$$(2) \quad \mathbf{y}_t^2 = \theta_2^1 \mathbf{y}_{t-1}^2 + \theta_2^2 \mathbf{y}_t^1 + \varepsilon_t^2$$

con  $\mathbb{E}[\varepsilon_t \varepsilon_t'] = \mathbf{I}$ .

Reexpresando en forma matricial:

$$y_t^1 = \Theta_1^1 y_{t-1}^1 + 0 \cdot y_t^1 + \Theta_1^2 y_t^2 + \varepsilon_t^1$$

$$y_t^2 = \Theta_2^1 y_{t-1}^2 + \Theta_2^2 \cdot y_t^1 + 0 \cdot y_t^2 + \varepsilon_t^2$$

$$\Leftrightarrow \underbrace{\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix}}_{\equiv \mathbf{X}_t} = \underbrace{\begin{bmatrix} 0 & \Theta_1^2 \\ \Theta_2^1 & 0 \end{bmatrix}}_{\equiv \mathbf{A}_0} \underbrace{\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix}}_{\equiv \mathbf{X}_t} + \underbrace{\begin{bmatrix} \Theta_1^1 & 0 \\ 0 & \Theta_2^2 \end{bmatrix}}_{\equiv \mathbf{A}_1} \underbrace{\begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix}}_{\equiv \mathbf{X}_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}}_{\equiv \boldsymbol{\varepsilon}_t = \mathbf{I} \boldsymbol{\varepsilon}_t}$$

Desarrollando algebraicamente:

$$\begin{aligned} \mathbf{I} \mathbf{X}_t &= \mathbf{A}_0 \mathbf{X}_t + \mathbf{A}_1 \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \\ \underbrace{(\mathbf{I} - \mathbf{A}_0) \mathbf{X}_t}_{\equiv \mathbf{B}^{-1}} &= \mathbf{A}_1 \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \\ \mathbf{B}^{-1} \mathbf{X}_t &= \mathbf{A}_1 \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \\ \implies \mathbf{X}_t &= \mathbf{B}^{-1} \mathbf{A}_1 \mathbf{X}_{t-1} + \mathbf{B}^{-1} \boldsymbol{\varepsilon}_t \end{aligned}$$

Luego, desarrollando la expresión para el *forecast error* a un periodo adelante:

$$\begin{aligned}
\mathbf{X}_{t+1} - \mathbb{E}[\mathbf{X}_{t+1} \mid \Omega_t] &= \mathbf{BA}_1 \mathbf{X}_t + \mathbf{B}\boldsymbol{\varepsilon}_{t+1} - \mathbb{E}[\mathbf{BA}_1 \mathbf{X}_t \mid \Omega_t] \\
&= \mathbf{BA}_1 \mathbf{X}_t - \mathbf{BA}_1 \mathbf{X}_t + \mathbf{B}\boldsymbol{\varepsilon}_{t+1} \\
&= \mathbf{B}\boldsymbol{\varepsilon}_{t+1}
\end{aligned}$$

Luego, desarrollando la expresión para la *forecast error variance* a un periodo adelante:

$$\begin{aligned}
\text{Var}(\mathbf{X}_{t+1} - \mathbb{E}[\mathbf{X}_{t+1} \mid \Omega_t]) &= \text{Var}(\mathbf{B}\boldsymbol{\varepsilon}_{t+1}) \\
&= \mathbf{B}\text{Var}(\boldsymbol{\varepsilon}_{t+1})\mathbf{B}^\top \\
&= \mathbf{B}\mathbb{E}[\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}_{t+1}']\mathbf{B}^\top \\
&= \mathbf{B}\mathbf{I}\mathbf{B}^\top \\
&= \mathbf{BB}^\top \\
&= \Sigma_u
\end{aligned}$$

Luego, desarrollando la expresión para el *forecast error* a 2 periodos adelante:

$$\begin{aligned}
\mathbf{X}_{t+2} - \mathbb{E}[\mathbf{X}_{t+2} \mid \Omega_t] &= \mathbf{BA}_1 \mathbf{X}_{t+1} + \mathbf{B}\boldsymbol{\varepsilon}_{t+2} - \mathbf{BA}_1 \mathbb{E}[\mathbf{X}_{t+1} \mid \Omega_t] \\
&= \mathbf{BA}_1 (\mathbf{X}_{t+1} - \mathbb{E}[\mathbf{X}_{t+1} \mid \Omega_t]) + \mathbf{B}\boldsymbol{\varepsilon}_{t+2} \\
&= \mathbf{BA}_1 \mathbf{B}\boldsymbol{\varepsilon}_{t+1} + \mathbf{B}\boldsymbol{\varepsilon}_{t+2}
\end{aligned}$$

Desarrollando la expresión para el *forecast error* a 3 periodos adelante:

$$\begin{aligned}
\mathbf{X}_{t+3} - \mathbb{E}[\mathbf{X}_{t+3} \mid \Omega_t] &= \mathbf{BA}_1 \mathbf{X}_{t+2} + \mathbf{B}\boldsymbol{\varepsilon}_{t+3} - \mathbf{BA}_1 \mathbb{E}[\mathbf{X}_{t+2} \mid \Omega_t] \\
&= \mathbf{BA}_1 (\mathbf{X}_{t+2} - \mathbb{E}[\mathbf{X}_{t+2} \mid \Omega_t]) + \mathbf{B}\boldsymbol{\varepsilon}_{t+3} \\
&= \mathbf{BA}_1 \mathbf{BA}_1 \mathbf{B}\boldsymbol{\varepsilon}_{t+1} + \mathbf{BA}_1 \mathbf{B}\boldsymbol{\varepsilon}_{t+2} + \mathbf{B}\boldsymbol{\varepsilon}_{t+3}
\end{aligned}$$

Desarrollando la expresión para el *forecast error* a  $h$  periodos adelante:

$$\begin{aligned}
\mathbf{X}_{t+h} - \mathbb{E}[\mathbf{X}_{t+h} \mid \Omega_t] &= \mathbf{BA}_1 \mathbf{X}_{t+h-1} + \mathbf{B}\varepsilon_{t+h} - \mathbb{E}[\mathbf{BA}_1 \mathbf{X}_{t+h-1} \mid \Omega_t] \\
&= \mathbf{BA}_1 \mathbf{X}_{t+h-1} - \mathbf{BA}_1 \mathbb{E}[\mathbf{X}_{t+h-1} \mid \Omega_t] + \mathbf{B}\varepsilon_{t+h} \\
&= \mathbf{BA}_1 (\mathbf{X}_{t+h-1} - \mathbb{E}[\mathbf{X}_{t+h-1} \mid \Omega_t]) + \mathbf{B}\varepsilon_{t+h}
\end{aligned}$$

$$\mathbf{X}_{t+h} - \mathbb{E}[\mathbf{X}_{t+h} \mid \Omega_t] = \sum_{i=1}^h (\mathbf{BA}_1)^{h-i} \mathbf{B}\varepsilon_{t+i}$$

**Forecast Error Variance** a  $h$  periodos adelante.

$$\begin{aligned}
\text{Var} \left( \sum_{i=1}^h (\mathbf{BA}_1)^{h-i} \mathbf{B}\varepsilon_{t+i} \right) &= \sum_{i=1}^h \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right] \text{Var}(\varepsilon_{t+i}) \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right]^\top \\
&= \sum_{i=1}^h \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right] \underbrace{\mathbb{E}[\varepsilon_{t+i} \varepsilon_{t+i}']}_{\mathbf{I}} \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right]^\top \\
&= \sum_{i=1}^h \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right] \mathbf{I} \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right]^\top \\
&= \sum_{i=1}^h \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right] \left[ (\mathbf{BA}_1)^{h-i} \mathbf{B} \right]^\top
\end{aligned}$$

■

### Pregunta 1.3.b

Sea la siguiente representación del mercado internacional del cobre:

$$\begin{aligned}
q_t^{Supply} &= \alpha_1 + \alpha_2 p_t + \alpha_3 p_{t-1} + \varepsilon_t^s \\
q_t^{Demand} &= \beta_1 + \beta_2 p_t + \beta_3 q_{t-1} + \varepsilon_t^d \\
q_t^{Supply} &= q_t^{Demand}
\end{aligned}$$

Reexpresando:

$$q_t = \alpha_1 + \underbrace{\Theta_1}_0 q_t + \alpha_2 p_t + \underbrace{\Theta_2}_0 q_{t-1} + \alpha_3 p_{t-1} + \varepsilon_t^s$$

$$q_t = \beta_1 + \underbrace{\Theta_3}_0 q_t + \beta_2 p_t + \beta_3 q_{t-1} + \underbrace{\Theta_4}_0 p_{t-1} + \varepsilon_t^d$$

Compactando la expresión en forma matricial:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 0 & \alpha_2 \\ 0 & \beta_2 \end{bmatrix} \begin{bmatrix} q_t \\ p_t \end{bmatrix} + \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ q_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix} \begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ q_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{bmatrix}$$

$$\begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}^{-1} \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ q_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{bmatrix}$$

Inviertiendo la matriz the matrix  $\begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}$ :

$$\begin{aligned} \begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}^{-1} &= \frac{1}{(1 - \beta_2)} \begin{bmatrix} 1 - \beta_2 & \alpha_2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \frac{\alpha_2}{1 - \beta_2} \\ 0 & \frac{1}{1 - \beta_2} \end{bmatrix} \end{aligned}$$

Lo anterior implica que la *dynamic matrix* está dada por:

$$\begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix} = \frac{1}{1 - \beta_2} \begin{bmatrix} 1 - \beta_2 & \alpha_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix}$$

$$\begin{aligned}
\begin{bmatrix} 1 & -\alpha_2 \\ 0 & 1 - \beta_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix} &= \frac{1}{1 - \beta_2} \begin{bmatrix} 1 - \beta_2 & \alpha_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \alpha_3 \\ \beta_3 & 0 \end{bmatrix} \\
&= \frac{1}{1 - \beta_2} \begin{bmatrix} \alpha_2 \beta_3 & (1 - \beta_2) \alpha_3 \\ \beta_3 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{\alpha_2 \beta_3}{1 - \beta_2} & \frac{(1 - \beta_2) \alpha_3}{1 - \beta_2} \\ \frac{\beta_3}{1 - \beta_2} & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{\alpha_2 \beta_3}{1 - \beta_2} & \alpha_3 \\ \frac{\beta_3}{1 - \beta_2} & 0 \end{bmatrix}
\end{aligned}$$

Así, el *VAR* del mercado internacional del cobre estará dado por:

$$\begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} 1 & \frac{\alpha_2}{1 - \beta_2} \\ 0 & \frac{1}{1 - \beta_2} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \frac{\alpha_2 \beta_3}{1 - \beta_2} & \alpha_3 \\ \frac{\beta_3}{1 - \beta_2} & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ q_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & \frac{\alpha_2}{1 - \beta_2} \\ 0 & \frac{1}{1 - \beta_2} \end{bmatrix} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{bmatrix}$$

Finalmente, es fácil ver que el coeficiente asociado a la persistencia del precio en equilibrio (i.e., el efecto de  $p_{t-1}$  sobre  $p_t$ ) es  $\frac{\beta_3}{1 - \beta_2}$ .

### Pregunta 1.3.c

Sea el SVAR:

$$\begin{aligned}
y_t^1 &= ay_{t-1}^1 + by_{t-1}^2 - \theta_1 y_t^2 + \varepsilon_t^1, \\
y_t^2 &= cy_{t-1}^1 + dy_{t-1}^2 - \theta_2 y_t^1 + \varepsilon_t^2,
\end{aligned}$$

con

$$\varepsilon_t = \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} \sim \mathcal{N}(0, \mathbf{I})$$

Reexpresando:

$$\begin{aligned}y_t^1 &= ay_{t-1}^1 + by_{t-1}^2 + 0 \cdot y_t^1 - \theta_1 y_t^2 + \varepsilon_t^1, \\y_t^2 &= cy_{t-1}^1 + dy_{t-1}^2 - \theta_2 y_t^1 + 0 \cdot y_t^2 + \varepsilon_t^2,\end{aligned}$$

Compactando la expresión en forma matricial:

$$\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} 0 & -\theta_1 \\ -\theta_2 & 0 \end{bmatrix} \begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}$$

$$\left[ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -\theta_1 \\ -\theta_2 & 0 \end{bmatrix} \right] \begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix} + \mathbf{I}_2 \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \theta_1 \\ \theta_2 & 1 \end{bmatrix} \begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix} + \mathbf{I}_2 \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}$$

$$\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} 1 & \theta_1 \\ \theta_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix} + \begin{bmatrix} 1 & \theta_1 \\ \theta_2 & 1 \end{bmatrix}^{-1} \mathbf{I}_2 \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}.$$

Invirtiendo  $\begin{bmatrix} 1 & \theta_1 \\ \theta_2 & 1 \end{bmatrix}$ :

$$\begin{bmatrix} 1 & \theta_1 \\ \theta_2 & 1 \end{bmatrix}^{-1} = \frac{1}{\theta_2 \theta_1 - 1} \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix} \equiv \alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix}$$

con  $\theta_2 \theta_1 \neq 0$ .

Así, la matriz de impacto está dada por:

$$\alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix}$$

Luego, el VAR reducido que es posible estimar de forma consistente mediante Máxima Verosimilitud está dado por:

$$\begin{aligned} \begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} &= \alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix} + \alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} \\ \Rightarrow \begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} &= \alpha \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \end{bmatrix} + \alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} \end{aligned}$$

donde:

$$\alpha \begin{bmatrix} 1 & -\theta_1 \\ -\theta_2 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} \equiv \boldsymbol{\xi}_t = \mathbf{B}^{-1} \varepsilon_t.$$

Luego, es fácil ver que el supuesto de identificación de *shocks* originales donde  $\mathbf{B}^{-1}$  es matriz triangular con  $\mathbb{E}[] = \mathbf{B}^{-1} \mathbf{B}^{-1'}$ , sólo se puede cumplir si  $\theta_1 = 0$  o  $\theta_2 = 0$ .

### Pregunta 1.3.d

Sea el siguiente  $VAR(1)$  estructural:

$$\begin{aligned} x_{t,1} &= \alpha_1 + \phi_{11}x_{t-1,1} + \phi_{12}x_{t-1,2} + \phi_{13}x_{t-1,3} + w_{t,1}, \\ x_{t,2} &= \alpha_2 + \phi_{21}x_{t-1,1} + \phi_{22}x_{t-1,2} + \phi_{23}x_{t-1,3} + w_{t,2}, \\ x_{t,3} &= \alpha_3 + \phi_{31}x_{t-1,1} + \phi_{32}x_{t-1,2} + \phi_{33}x_{t-1,3} + w_{t,3}. \end{aligned}$$

con  $w_{t,1}, w_{t,2}$ , y  $w_{t,3}$  ruidos blancos independientes entre sí.

Reescribiendo el  $VAR$  en forma matricial:

$$\begin{bmatrix} x_{t,1} \\ x_{t,2} \\ x_{t,3} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} x_{t-1,1} \\ x_{t-1,2} \\ x_{t-1,3} \end{bmatrix} + \begin{bmatrix} w_{t,1} \\ w_{t,2} \\ w_{t,3} \end{bmatrix}$$

$$\mathbf{x}_t = \boldsymbol{\alpha} + \boldsymbol{\Phi} \mathbf{x}_{t-1} + \mathbf{w}_t$$



Para determinar estacionaridad se requieren imponer restricciones al determinante de la matriz dinámica, tal como en la pregunta 1.2.b En particular, se requiere que los *eigenvalues* de  $\Phi$  sean menores a 1 en módulo.

Para encontrar los *eigenvalues* se requiere encontrar la solución al polinomio característico:

$$\det(\mathbf{I} - \Phi\lambda) = 0$$

Luego, desarrollando:

$$\det(\Phi - \lambda\mathbf{I}) = \det \begin{bmatrix} \phi_{11} - \lambda & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} - \lambda & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} - \lambda \end{bmatrix}$$

Podemos establecer que el polinomio establece una ecuación cúbica en  $\lambda$  para encontrar las soluciones. Sin embargo, se puede establecer que las raíces de la ecuación, para que el VAR sea estacionario (estable), deben cumplir:

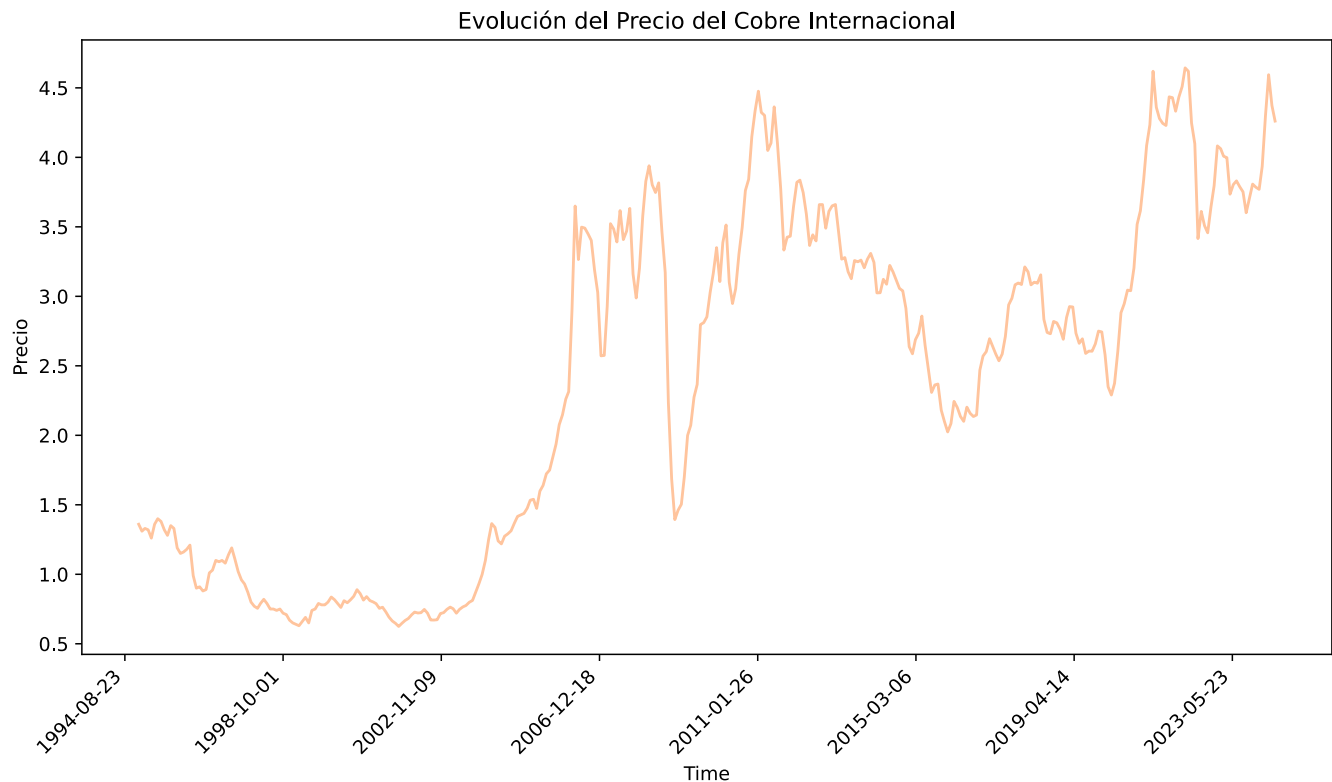
$$|\lambda_1| < 1, \quad |\lambda_2| < 1, \quad |\lambda_3| < 1.$$

Si lo anterior se cumple, el *VAR* es estacionario.

## Pregunta 2

### Pregunta 2.a

A continuación, se despliega la evolución del precio del cobre.



A simple vista, se puede establecer que es muy probable que la serie sea no estacionaria, Por lo tanto, se ha sometido a prueba a la serie mediante el *Augmented Dickey-Fuller test* implementado vía comando de python:

```
# (A Little hardcoded):
```

```
Serie = "Precio del Cobre Internacional"
```

```
print(f"Augmented Dickey-Fuller test for \"{Serie}\":")
result = adfuller(df['copper'].dropna(), autolag='AIC')
print(f"ADF Statistic: {result[0]}")
print(f"p-value: {format(result[1], 'f')}")
# print("Critical Values:")
# for key, value in result[4].items():
#     print(f"    {key}: {value}")
if result[1] < 0.05:
    print(f"\ \"{Serie}\" is likely stationary (reject the null hypothesis).")
```

```

else:
    print(f"\n{Serie}\" is likely non-stationary (fail to reject the null hypothesis).")
print()

```

El diseño de las hipótesis del test está dado por:

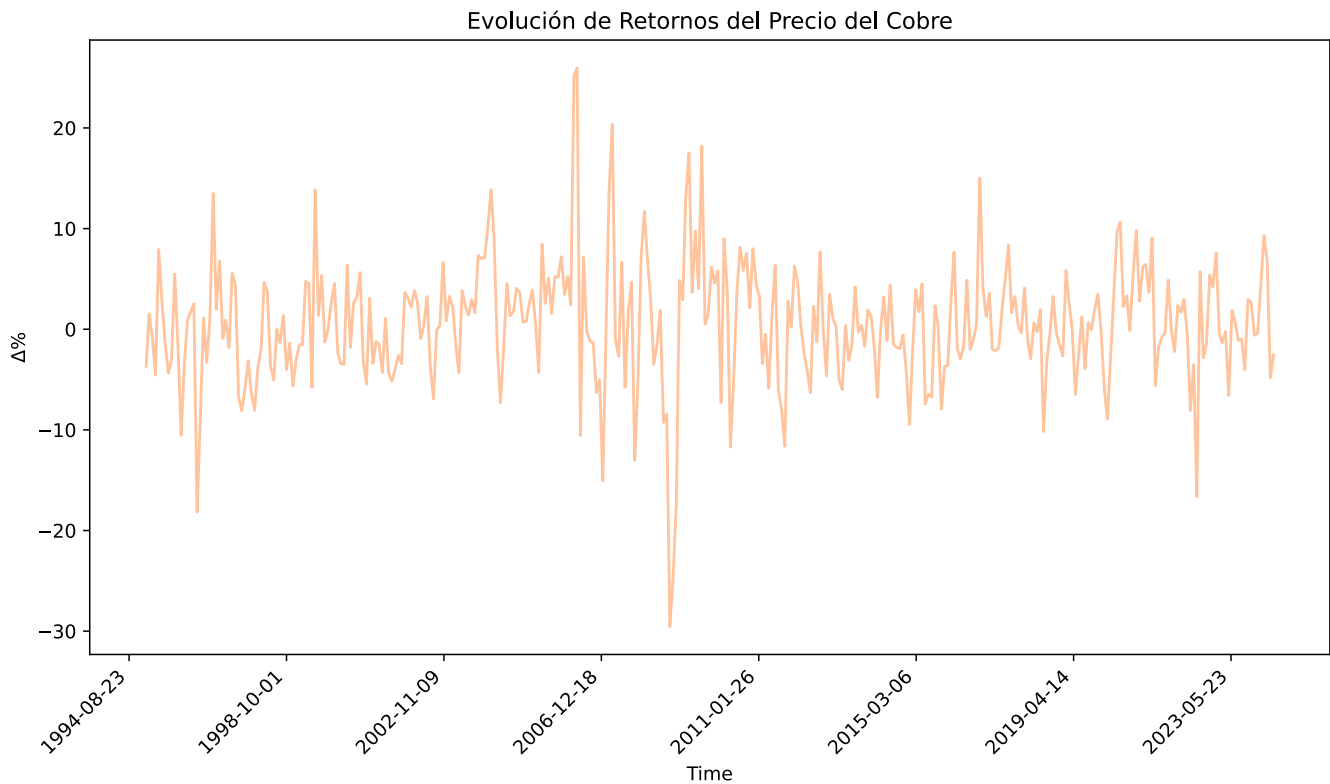
$H_0$ : La serie es de raíz unitaria, i.e., no es estacionaria.

$H_1$  : La serie es estacionaria.

#### Augmented Dickey-Fuller Test for Precio del Cobre Internacional

ADF Statistic	-1.283245
p-value	0.636803

Dado que el valor-p se encuentra por sobre el nivel de tolerancia estandar  $\alpha = 5\%$ , no podemos la hipótesis nula  $H_0$  y establecemos que existe evidencia estadística de que la serie de de la evolución mensual del precio del cobre internacional es no estacionaria y requiere transformación. Ante esto, se ha aplicado primera diferencia logarítmica:



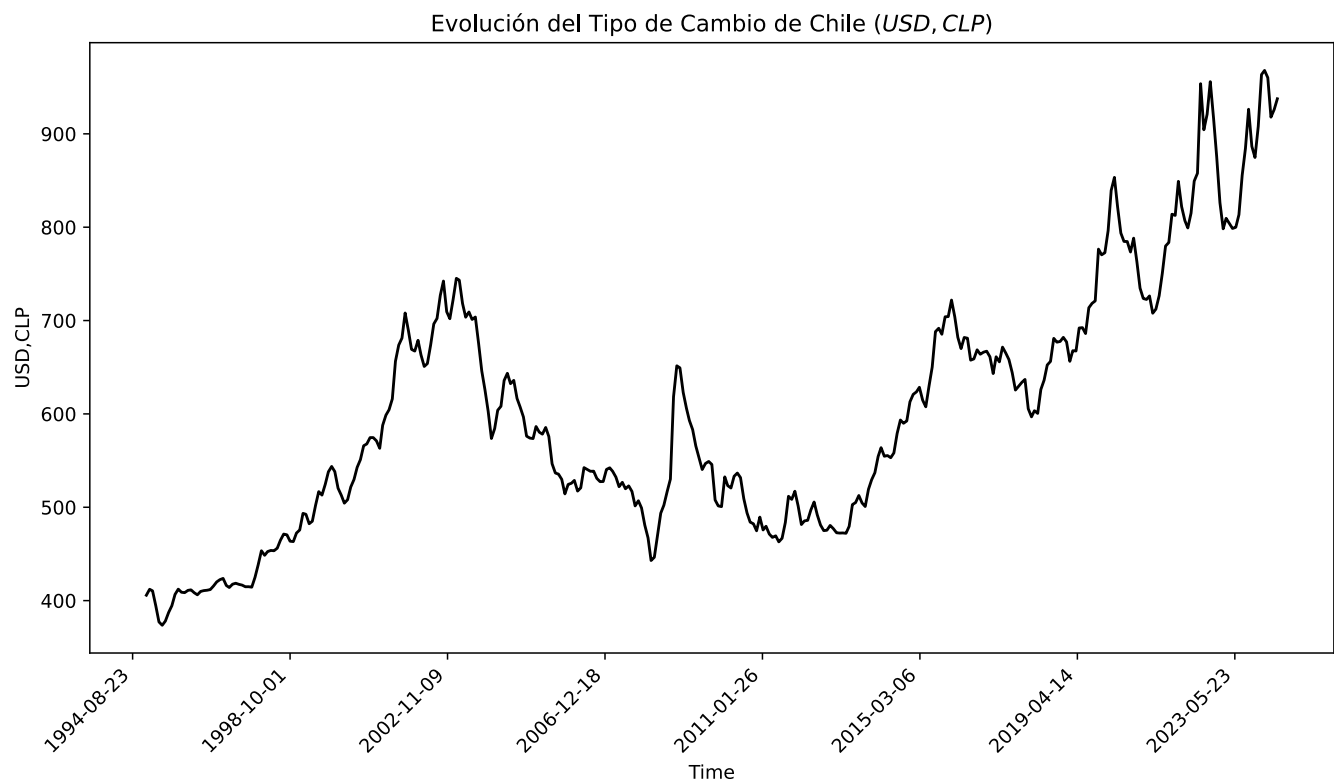
Luego de realizar la transformación, se testea estacionariedad:

Augmented Dickey-Fuller Test for Retornos del Precio del Cobre

ADF Statistic	-12.713640
p-value	0.000000

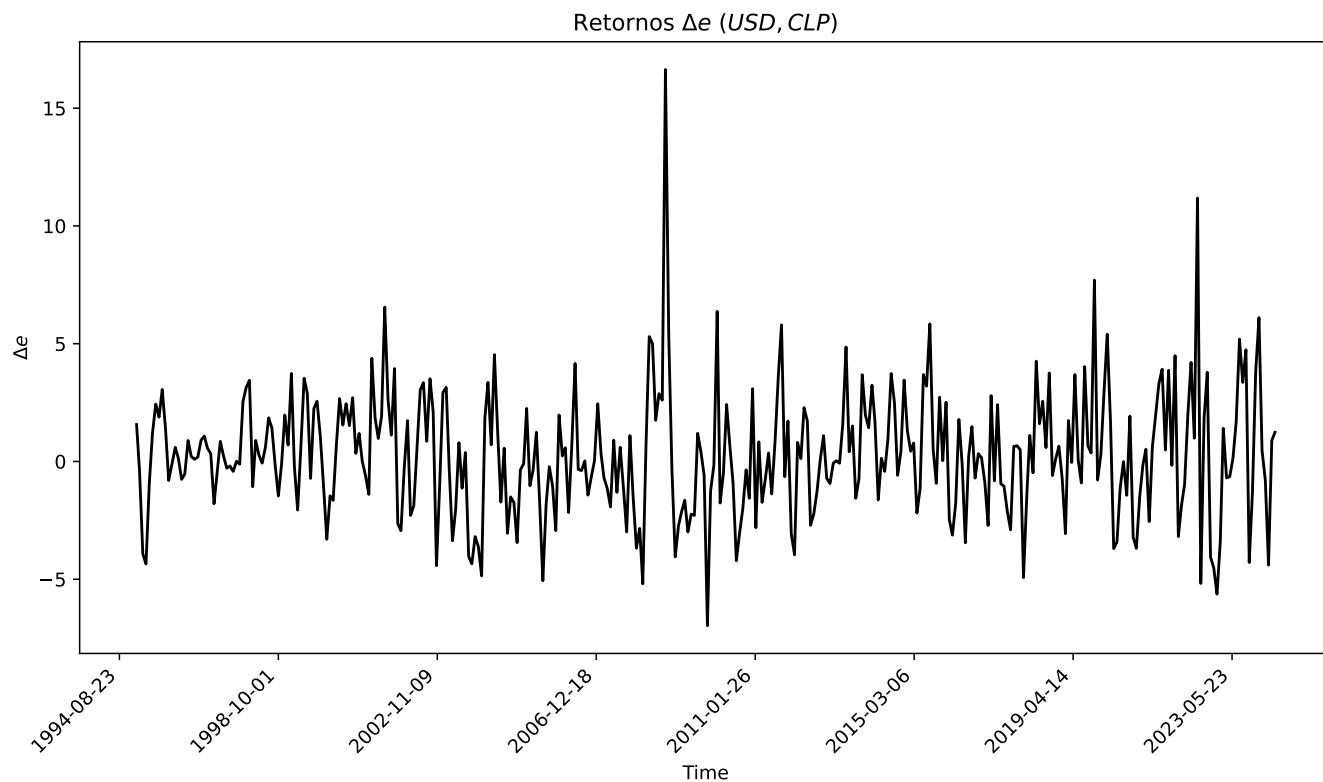
Así, podemos establecer que la serie de retornos del cobre ya es estacionaria:  $p\text{-value} < \alpha$ .

A continuación, se visualizan los datos de Tipo de Cambio de Chile descargados sin manipular desde el sitio web del *Banco Central de Chile*.



Replicando la aplicación del *Augmented Dickey-Fuller test*, encontramos evidencia de no estacionariedad dado que  $p\text{-value} > \alpha$ .

Ante esto, se ha aplicado primera diferencia logarítmica:



Luego de realizar la transformación, se testea estacionariedad al retorno  $USD, CLP$ :

Augmented Dickey-Fuller Test for Retornos Tipo de Cambio USD,CLP	
ADF Statistic	-13.941785
p-value	0.000000

Así, podemos establecer que la última serie ya es estacionaria: p-value  $\alpha$ .

## Pregunta 2.b

Para determinar el largo del  $VAR(p)$  se ha ejecutado el comando de Stata `varsoc copper_returns spot_returns, maxlag(12)`, obteniendo:

Lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-1868.8				193.371	10.9404	10.9493	10.9628
1	-1834.24	69.13*	4	0.000	161.72*	10.7616*	10.7884*	10.8289*
2	-1833.3	1.8727	4	0.759	164.644	10.7795	10.8242	10.8917
3	-1832.1	2.3976	4	0.663	167.365	10.7959	10.8585	10.9529
4	-1829.4	5.4142	4	0.247	168.637	10.8035	10.8839	11.0053
5	-1827.53	3.7301	4	0.444	170.759	10.816	10.9142	11.0626
6	-1823.78	7.5012	4	0.112	171.014	10.8174	10.9336	11.109
7	-1820.75	6.0572	4	0.195	171.995	10.8231	10.9571	11.1595
8	-1817.16	7.1914	4	0.126	172.411	10.8255	10.9773	11.2067
9	-1813.92	6.479	4	0.166	173.19	10.8299	10.9997	11.256
10	-1812.02	3.7936	4	0.435	175.348	10.8422	11.0298	11.3132
11	-1808.83	6.3757	4	0.173	176.2	10.847	11.0524	11.3628
12	-1807.13	3.3977	4	0.494	178.608	10.8604	11.0838	11.4211

Cuadro 1: VAR Lag Order Selection Criteria.

Podemos ver que el largo indicado es de 1 según todos los criterios.

Luego, se estimo el  $VAR(1)$  usando `var copper_returns spot_returns, lags(1)`, obteniendo:

Cuadro 2: VAR Model Summary

Equation	Parms	RMSE	R-sq	Chi2	P>Chi2
Copper Returns	3	5.64124	0.1478	61.198	0.0000
Spot Returns	3	2.53042	0.0843	32.48246	0.0000

Se aprecia que el VAR es conjuntamente significativo por el valor-p asociado a los tests  $\chi^2$ .

Cuadro 3: VAR Model Coefficients

Equation	Variable	Coef. del Lag	Std. Err.	z	P >  z	[95 % Conf. Interval]
Copper Returns	Copper Returns	0.3085451	0.0568722	5.43	0.000	[0.1971, 0.4200]
	Spot Returns	-0.2804384	0.1314471	-2.13	0.033	[-0.5381, -0.0228]
	<i>Constant</i>	0.4371309	0.3042162	1.44	0.151	[-0.1591, 1.0334]
Spot Returns	Copper Returns	-0.0204854	0.0255104	-0.80	0.422	[-0.0705, 0.0295]
	Spot Returns	0.2634805	0.0589616	4.47	0.000	[0.1479, 0.3790]
	<i>Constant</i>	0.2072001	0.1364584	1.52	0.129	[-0.0603, 0.4747]

Dado que las variables están en puntos porcentuales, los lags tienen magnitudes de afectación de un mes a otro de alrededor de 0.2 puntos porcentuales, hasta 0.43 puntos, lo que es bastante relevante en términos anualizados.

Luego, se testeó *Granger Causality* usando el comando `vargranger`, obteniendo:

Cuadro 4: Granger Causality Wald Tests

Equation	Excluded	Chi2	P > Chi2
Copper Returns	Spot Returns	4.5517	0.033
Copper Returns	ALL	4.5517	0.033
Spot Returns	Copper Returns	0.6448	0.422
Spot Returns	ALL	0.6448	0.422

La causalidad de Granger se puede expresar analíticamente de la siguiente forma<sup>4</sup>:

$$\begin{aligned} \text{Si } \theta(L) \neq 0 \text{ en } y_t = \phi(L)y_{t-1} + \theta(L)z_{t-1} + \varepsilon_t^y \\ \implies z_t \text{ causa-a-la-Granger a } y_t \end{aligned}$$

Excluyendo la variable de los retornos de tipo de cambio en la ecuación del precio del cobre se encontró un valor-p asociado al test  $\chi^2$  de  $0,033 < \alpha$ , indicando Granger causalidad de cambios en el dólar al precio del cobre. Esto, probablemente porque movimientos en el dólar podrían tener un efecto en el precio del cobre percibido por los mercados internacionales, lo que en última instancia, a través de la demanda, podría alterar el precio del cobre. Esto, en consideración de que Chile representa una fracción relevante en el mercado mundial. Es decir, cambios en el dólar se traducen en cambios en el precio del cobre en el próximo mes en puntos porcentuales **muy** significativos.

Sin embargo, excluyendo la variable del cobre en los retornos de tipo de cambio se encontró un valor-p asociado al test  $\chi^2$  de  $0,42 > \alpha$ , indicando no existencia de Granger causalidad en los cambios del tipo de cambio en función del precio del cobre rezagado. Es decir, movimientos en el cobre no afectan el dólar el próximo mes. El dólar del próximo periodo se determinaría por otros factores.

En resumen, dólar causa a la Granger el precio del cobre, pero el cobre no causa a la Granger el dólar, en contra de lo encontrado por Schmitt-Grohé & Uribe, 2017.

## Pregunta 2.c

La causalidad de Sims se puede expresar analíticamente de la siguiente forma<sup>5</sup>:

$$\begin{aligned} \text{Si } \theta(L) \neq 0 \text{ en } y_t = c + \sum_{i=0}^{\infty} \phi_i(L)z_{t-i} + \sum_{i=1}^{\infty} \theta_i(L)z_{t+i} + \varepsilon_t^y \\ \implies z_t \text{ causa-a-la-Sims a } y_t \end{aligned}$$

<sup>4</sup>Para este tarea se testeó con un solo rezago.

<sup>5</sup>Para este tarea se testeó con un solo rezago y adelanto. Para una inferencia robusta ante la presencia de autocorrelación se testeó con el comando de Stata `newey`.

Es decir, si controlando por el valor contemporáneo y los rezagos, los valores futuros<sup>6</sup> de la variable contraria son relevantes para determinarla.

La causalidad de Geweke, Meese & Dent, en cambio, se puede expresar de la siguiente forma:

$$\text{Si } \theta(L) \neq 0 \text{ en } y_t = c - \sum_{i=1}^{\infty} \lambda_i(L) y_{t-i} + \sum_{i=0}^{\infty} \phi_i(L) z_{t-i} + \sum_{i=1}^{\infty} \theta_i(L) z_{t+i} + \varepsilon_t^y \\ \implies z_t \text{ causa-a-la-GMD a } y_t$$

Es decir, controlando por sus auto-rezagos, valores contemporáneos de la otra variable y rezagos de la otra variable, si los valores futuros<sup>7</sup> de la variable contraria son relevantes para determinarla.

Para testear à-la-Sims se hizo una transformación inteligente para sobreponerse a la no existencia de un comando directo de estos tests. Se crearon variables para los adelantos, y se estimaron las regresiones correspondientes para cada ecuación con inferencia robusta, sin usar sistemas dinámicos:

```
//
// Non-Granger Causality Tests
//

gen copper_returns_f1 = F.copper_returns    // Forward one period
gen spot_returns_f1 = F.spot_returns       // Forward one period

// Sims causality

newey copper_returns 1.spot_returns spot_returns spot_returns_f1, lag(0)
newey spot_returns 1.copper_returns copper_returns copper_returns_f1, lag(0)

// Geweke, Meese & Dent causality

newey copper_returns 1.copper_returns 1.spot_returns spot_returns spot_returns_f1, lag(0)
newey spot_returns 1.spot_returns 1.copper_returns copper_returns copper_returns_f1, lag(0)
```

---

<sup>6</sup>como *proxy* de variables *forward-looking*.

<sup>7</sup>como *proxy forward-looking*.



Los resultados obtenidos (estimación) para el Test de Causalidad de Sims del tipo de cambio sobre el cobre son:

Cuadro 5: Newey–West Regression Results: Copper Returns Equation

Equation	Variable	Coefficient	Std. Err.	$t$	$P >  t $	[95 % Conf. Interval]
Copper Returns	Lag Spot Returns	-0.3345199	0.1412552	-2.37	0.018	[-0.6123, -0.0567]
	Spot Returns	-1.039384	0.1266897	-8.20	0.000	[-1.2886, -0.7902]
	Forward Spot Returns	-0.0971899	0.0954092	-1.02	0.309	[-0.2848, 0.0905]
	<i>Constant</i>	0.9209708	0.2783971	3.31	0.001	[0.3734, 1.4685]

Se observa que el coeficiente del adelanto del tipo de cambio no es significativo, por lo tanto, el tipo de cambio no causa à-la-Sims al precio del cobre. Este es un cambio respecto a los resultados en (2.b).

Los resultados obtenidos (estimación) para el Test de Causalidad de Sims del cobre sobre el tipo de cambio:

Cuadro 6: Newey–West Regression Results: Spot Returns Equation

Equation	Variable	Coefficient	Std. Err.	$t$	$P >  t $	[95 % Conf. Interval]
Spot Returns	Lag Copper Returns	0.0026779	0.0221484	0.12	0.904	[-0.0409, 0.0462]
	Copper Returns	-0.2020786	0.0352453	-5.73	0.000	[-0.2714, -0.1328]
	Forward Copper Returns	-0.0454774	0.0251424	-1.81	0.071	[-0.0949, 0.0040]
	<i>Constant</i>	0.3939107	0.1242432	3.17	0.002	[0.1495, 0.6383]

Se observa que el coeficiente del adelanto del cobre no es significativo, por lo tanto, el cobre no causa à-la-Sims al tipo de cambio, en línea con (2.b).

Los resultados obtenidos (estimación) para el Test de Causalidad de Geweke, Meese & Dent del tipo de cambio sobre el cobre son:

Cuadro 7: Newey–West Regression Results: Copper Returns Equation (Extended Model)

Equation	Variable	Coefficient	Std. Err.	$t$	$P >  t $	[95 % Conf. Interval]
Copper Returns	Lag Copper Returns	0.2870655	0.0800671	3.59	0.000	[0.1296, 0.4445]
	Lag Spot Returns	-0.0076976	0.1300291	-0.06	0.953	[-0.2634, 0.2480]
	Spot Returns	-1.013766	0.1226446	-8.27	0.000	[-1.2550, -0.7725]
	Forward Spot Returns	-0.0989826	0.0820144	-1.21	0.228	[-0.2603, 0.0623]
	<i>Constant</i>	0.6743118	0.2633354	2.56	0.011	[0.1564, 1.1922]

Se observa que el coeficiente del adelanto del tipo de cambio no es significativo, por lo tanto, el tipo de cambio no causa à-la-Geweke, Meese & Dent al precio del cobre. Este es un cambio respecto a los resultados en (2.b).

Los resultados obtenidos (estimación) para el Test de Causalidad de Geweke, Meese & Dent del cobre sobre el tipo de cambio son:

Cuadro 8: Newey–West Regression Results: Spot Returns Equation (Extended Model)

<b>Equation</b>	<b>Variable</b>	<b>Coefficient</b>	<b>Std. Err.</b>	<b><i>t</i></b>	<b><math>P &gt;  t </math></b>	<b>[95 % Conf. Interval]</b>
Spot Returns	Lag Spot Returns	0.1962705	0.0537631	3.65	0.000	[0.0905, 0.3020]
	Lag Copper Returns	0.0423053	0.0222736	1.90	0.058	[-0.0015, 0.0861]
	Copper Returns	-0.195749	0.0351473	-5.57	0.000	[-0.2649, -0.1266]
	Forward Copper Returns	-0.0385617	0.0241627	-1.60	0.111	[-0.0861, 0.0090]
	<i>Constant</i>	0.3135154	0.1227802	2.55	0.011	[0.0720, 0.5550]

Se observa que el coeficiente del adelanto del cobre no es significativo, por lo tanto, el cobre no causa à-la-Geweke, Meese & Dent al tipo de cambio, en línea con (2.b).

## Referencias

Banco Central de Chile (2024), Base de Datos Estadísticos del Banco Central de Chile, Economía Internacional, Materias primas, Precios (data set). Database link: Precio del cobre refinado BML (dólares/libra).

Banco Central de Chile (2024), Base de Datos Estadísticos del Banco Central de Chile, Tipos de Cambio, Tipo de cambio nominal distintas monedas (pesos por unidad de moneda extranjera (data set). Database link: Estados Unidos: Dólar de Estados Unidos (Observado).

Geweke, J., Meese, R., & Dent, W. (1983). “Comparing alternative tests of causality in temporal systems: Analytic results and experimental evidence”. *Journal of Econometrics*, 21(2), 161–194. [https://doi.org/10.1016/0304-4076\(83\)90012-X](https://doi.org/10.1016/0304-4076(83)90012-X)

Granger, C. W. J. (1969). “Investigating Causal Relations by Econometric Models and Cross-spectral Methods”. *Econometrica*, 37(3), 424–438. <https://doi.org/10.2307/1912791>.

Sims, C. (1972). “Money, Income, and Causality”. *The American Economic Review*, vol. 62, no. 4, 1972, pp. 540–52. JSTOR, <http://www.jstor.org/stable/1806097>

Soto, R. (2024), “Econometría de Series de Tiempo (Parte 4, 2022)”[Slides Teoría Econométrica III]. Pontificia Universidad Católica de Chile.

Uribe, M. & Schmitt-Grohé, S. (2017), “Open Economy Macroeconomics”. Princeton University Press.

Vicondoa, A. (2023), “Sección 7: VAR Models”[Slides Macroeconometría Aplicada]. Pontificia Universidad Católica de Chile.

Vicondoa, A. (2023), “SVAR Models”[Slides Macroeconometría Aplicada]. Pontificia Universidad Católica de Chile.