

La serie "Borradores Semanales de Economía" es una publicación de la Subgerencia de Estudios Económicos del Banco de la República. Los Trabajos son de carácter provisional, las opiniones y posibles errores son responsabilidad exclusiva de los autores y sus contenidos no comprometen al Banco de la República ni a su Junta Directiva.

## **PRONOSTICOS CONDICIONADOS PARA MODELOS VAR**

**Por:  
Luis Fernando Melo V.**

**1996**

**No. 62**

Para comentarios favor dirigirse al autor:  
Fax: 2865936 - Teléfono 3347193

# **PRONOSTICOS CONDICIONADOS PARA MODELOS VAR**

**Luis Fernando Melo V.\***

**Santafé de Bogotá, Octubre de 1996**

---

\* Las opiniones expresadas son responsabilidad exclusiva del autor y no necesariamente reflejan las del Banco de la República. Se agradecen los comentarios de Martha Misas, Fabio Nieto, Norberto Rodríguez, Hugo Oliveros y Alvaro Riascos.

## I. Introducción

La incorporación y evaluación de metas en modelos econométricos juegan un papel importante en la toma de decisiones y el planeamiento de políticas económicas. En este documento se desarrolla una metodología para construir pronósticos condicionados sobre modelos que admiten una representación VAR<sup>1</sup>. Adicionalmente, se encuentra una aproximación al sesgo generado en los pronósticos de las series en la escala original al aplicar la transformación logaritmo en modelos VAR.

La metodología de pronósticos condicionados, presentada en la sección II, se basa en el procedimiento desarrollado para modelos de función de transferencia<sup>2</sup> sugerido por Guerrero (1991). Las restricciones son introducidas en forma de combinaciones lineales de los valores futuros. Los pronósticos condicionados se construyen incorporando las restricciones a pronósticos obtenidos de un modelo multivariado VAR. Este método además de generar los pronósticos condicionados con intervalos de confianza, facilita la evaluación de la verosimilitud del sendero sugerido con respecto a la evolución histórica de las variables. Por último se presenta una aplicación de esta metodología con datos simulados bajo la estructura de un modelo VAR.

## II. Metodología

### A. Modelos VAR

Sea  $Z_t = (z_{1t}, z_{2t}, \dots, z_{kt})'$  un vector compuesto por  $k$  variables observado durante  $t = 1, 2, \dots, N$ , el cual sigue un modelo VAR de orden  $p$ :

$$(1) \quad Z_t = v + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t$$

---

<sup>1</sup> El caso univariado con modelos ARIMA es considerado en Melo y Oliveros (1991).

<sup>2</sup> Una amplia discusión sobre modelos de función de transferencia es encontrada en Box y Jenkins (1970) y Haugh y Box (1977) entre otros.

donde los  $\phi_i$  son las matrices ( $K \times K$ ) de coeficientes,  $v$  es el vector  $k$ -dimensional de interceptos del modelo y  $a_t = (a_{1t}, a_{2t}, \dots, a_{kt})'$  es un vector ruido blanco con distribución  $N_k(0, \Sigma)$ , donde  $\Sigma$  es la matriz de varianzas covarianzas asociada al vector  $a_t$ .

El modelo especificado en (1) también puede ser representado como:

$$(2) \quad \Phi(B)Z_t = v + a_t$$

donde  $\Phi(B) = (I_k - \phi_1 B - \dots - \phi_p B^p)$  y  $B$  es el operador de rezago.

Si este proceso VAR( $p$ ) es estacionario puede ser expresado de la forma (Judge et al, 1988):

$$(3) \quad Z_t = \mu + \Psi(B)a_t$$

donde  $\Psi(B) = (I_k + \psi_1 B + \psi_2 B^2 + \dots)$ ,  $\psi_i$  son matrices de dimensión  $k \times k$  y  $\mu = E(Z_t)$  de dimensión  $k \times 1$ .

Es de recordar (e.g. Judge et al, 1988) que el pronóstico con error cuadrático medio mínimo (ECMM) de  $Z_{N+h}$ , para  $h=1,2,\dots$ , viene dado por la esperanza condicional:

$$E_0(Z_{N+h}) = E(Z_{N+h} | Z_0)$$

donde  $Z_0 = (Z'_1, Z'_2, \dots, Z'_N)'$  es un vector de dimensión  $kN \times 1$  y contiene la información histórica, observada hasta el tiempo  $N$ .

El error de pronóstico para este modelo esta dado por:<sup>3</sup>

$$(4) \quad e_N(h) = Z_{N+h} - E_0(Z_{N+h}) = \sum_{j=0}^{h-1} \psi_j a_{N+h-j}$$

---

<sup>3</sup> Para modelos no estacionarios aunque la ecuación (3), dada por la el teorema de descomposición de Wold (1938), no se cumple, la ecuación (4) si es valida. Este resultado puede ser consultado entre otros en Wei (1993), para modelos ARIMA en las pág. 87-88 y para modelos VAR en la pág. 364 o en Lütkepohl (1987).

Si la longitud del horizonte de pronóstico es denotada por  $H$  y los vectores  $Z$  y  $a$  para todo este periodo son escritos como  $Z_F = (Z'_{N+1}, Z'_{N+2}, \dots, Z'_{N+H})'$  y  $a_F = (a'_{N+1}, a'_{N+2}, \dots, a'_{N+H})'$ , respectivamente, ambos de dimensión  $kH \times 1$ , entonces:

$$(5) \quad Z_F - E_0(Z_F) = \Psi a_F$$

donde la matriz  $\Psi$  es de dimensión  $kH \times kH$  y está dada por:

$$(6) \quad \Psi = \begin{bmatrix} \psi_0 & 0 & 0 & \dots & 0 \\ \psi_1 & \psi_0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_{H-1} & \psi_{H-2} & \psi_{H-3} & \dots & \psi_0 \end{bmatrix}, \psi_0 = I_k$$

además como  $a_t \sim N_k(0, \Sigma)$  es un vector ruido blanco:

$$(7) \quad a_F \sim N_{kH}(0, I_H \otimes \Sigma)$$

Este resultado se analiza en mayor detalle en el anexo.

Basados en (5) y (7) podemos deducir la matriz de covarianzas de los errores de pronóstico del modelo VAR:

$$(8) \quad Cov(Z_F - E_0(Z_F)) = \Psi(I_H \otimes \Sigma)\Psi'$$

## B. Pronósticos VAR condicionados

Las restricciones o metas sobre los valores futuros de  $\{Z_t\}$ , es decir sobre  $Z_F$ , se asumen como combinaciones lineales de estos valores; así, las  $m$  restricciones sobre el horizonte de pronóstico  $H$ , pueden ser expresadas como:

$$(9) \quad Y = CZ_F$$

donde  $C$  es una matriz de dimensión  $m \times kH$  con rango  $m \leq kH$ , la cual indica las variables y el horizonte de tiempo sobre las cuales se aplican las restricciones. En el vector  $Y$  se especifican los valores de las restricciones.

Si se tienen  $m$  restricciones sobre el horizonte de pronóstico  $H$  utilizando  $k$  variables, el vector  $Y$  será, por lo tanto, de orden  $m \times 1$ , la matriz  $C$  de orden  $m \times kH$  y el vector  $Z_F$  de orden  $kH \times 1$ .

Por ejemplo, si se tienen dos restricciones ( $y_1, y_2$ ), una sobre la primera variable en el primer periodo de tiempo fuera de muestra ( $N+1$ ) y otra sobre la  $k$ -ésima variable en el último periodo del horizonte de pronóstico ( $N+H$ ), entonces:

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{2 \times 1}, \quad C = \begin{bmatrix} \underbrace{1 \quad 2 \quad \dots \quad k}_{1} & \underbrace{0 \quad 0 \quad \dots \quad 0}_{2} & \dots & \underbrace{0 \quad 0 \quad \dots \quad 0}_{H} & \dots & \underbrace{0 \quad 0 \quad \dots \quad 0}_{H} \\ \underbrace{0 \quad 0 \quad \dots \quad 0}_{1} & \underbrace{0 \quad 0 \quad \dots \quad 0}_{2} & \dots & \dots & \dots & \underbrace{0 \quad 0 \quad \dots \quad 1}_{H} \end{bmatrix}_{2 \times kH}$$

### 1. Obtención del Pronóstico condicionado y su matriz de covarianzas.

Esta sección generaliza al caso Multivariado la metodología sugerida por Guerrero (1989), que inicialmente se realizó para pronósticos condicionados con un modelo univariado ARIMA.

Incorporando la ecuación (5) en la información adicional (9) se tiene:

$$(10) \quad Y = CE_0(Z_F) + C\Psi a_F$$

para obtener los pronósticos condicionados,  $\tilde{Z}_F$ , que incorporen las condición (9) debe existir una matriz  $A$  de dimensión  $H \times m$  tal que  $\tilde{Z}_F = AY$ , es decir, utilizando (10):

$$(11) \quad \tilde{Z}_F = AC(E_0(Z_F) + \Psi a_F)$$

por insesgamiento del error de pronóstico:

$$(12) \quad E_0(\tilde{Z}_F - Z_F) = 0$$

tomando valor esperado en (11):

$$(13) \quad E_0(\tilde{Z}_F) = AC E_0(Z_F)$$

reemplazando (13) en (12) se tiene:

$$(14) \quad (AC - I)E_0(Z_F) = 0$$

reemplazando (5) en (14):

$$(15) \quad (AC - I)(Z_F - \Psi a_F) = 0$$

simplificando (15) y utilizando (9) y la relación  $\tilde{Z}_F = AY$ :

$$(16) \quad \tilde{Z}_F - Z_F = (AC - I)\Psi a_F$$

y la matriz de covarianzas de (16) es:

$$(17) \quad \text{cov}(\tilde{Z}_F - Z_F) = (AC - I)\Psi (I \otimes \Sigma) \Psi'(C'A' - I)$$

Si la estimación de  $\tilde{Z}_F$  es realizada minimizando el error cuadrático medio, la matriz **A** debe minimizar la varianza generalizada, definida como:

$$(18) \quad \text{var}_F = \text{tr} \left\{ \text{cov}_0(\tilde{Z}_F - Z_F) \right\}$$

de tal modo que **A** se encuentra resolviendo  $\partial \text{var}_F / \partial A = 0$ , de lo cual se obtiene:

$$(19) \quad \hat{A} = \Psi(I \otimes \Sigma)\Psi'C' [C\Psi(I \otimes \Sigma)\Psi'C']^{-1}$$

reemplazando la ecuación (14) en (11) y estimando **A** se obtiene:

$$(20) \quad \hat{Z}_F = E_0(Z_F) + \hat{A}C\Psi a_F$$

donde  $\hat{Z}_F$  corresponde a la estimación del vector  $\tilde{Z}_F$ .

Reemplazando la ecuación (10) en (20) se obtiene la siguiente expresión final para los pronósticos condicionados<sup>4</sup>:

$$(21) \quad \hat{Z}_F = E_0(Z_F) + \hat{A}[Y - CE_0(Z_F)]$$

si definimos el error del pronóstico condicionado como  $e_F = Z_F - \tilde{Z}_F$ , de la ecuación (16) y utilizando el supuesto de distribución de  $a_F$  dado en (7), se tiene:

$$(22) \quad E_0(e_F | Y) = 0 \quad C\hat{O}V_0(e_F | Y) = (I - \hat{A}C)\Psi(I \otimes \Sigma)\Psi'$$

---

<sup>4</sup> Los resultados obtenidos sobre la estimación de los pronósticos condicionados están sujetos al supuesto que los parámetros involucrados en esta metodología son conocidos, sin embargo, si los parámetros son desconocidos y por tanto estimados, los resultados de este documento siguen siendo válidos asintóticamente. Esta propiedad se analiza en mayor detalle en el anexo.



Así, los intervalos de confianza del 100(1- $\alpha$ )% para el vector de pronósticos condicionados están dados por:

$$(23) \quad \hat{Z}_F \pm \delta_{1-\alpha/2} [\hat{VAR}_0(e_F | Y)]^{1/2}$$

donde  $\delta_{1-\alpha/2}$  es el percentil 100(1- $\alpha$ ) de la distribución normal,  $VAR_0(e_F | Y)$  es la diagonal de la matriz  $COV_0(e_F | Y)$  y el exponente  $1/2$  indica que a cada elemento del vector de varianzas se le es extraída la raíz cuadrada.

Los resultados finales especificados en las ecuaciones (19) , (21) y (22) son similares a los derivados por Guerrero (1991) para pronósticos condicionales con modelos de función de transferencia.

### C. Prueba de hipótesis de compatibilidad

Para decidir si las metas especificadas en el vector Y pueden ser consideradas consistentes con los datos históricos bajo el modelo VAR especificado , se plantea la siguiente prueba de hipótesis:

$$(24) \quad H_0: Y = CZ_F$$

Asumiendo la representación VAR dada en (1) para las series en estudio, de la ecuación(5) se tiene:

$$(25) \quad Z_F = E_0(Z_F) + \Psi a_F$$

donde  $a_F \sim N_{kH}(0, I_H \otimes \Sigma)$ .

incorporando la restricción  $Y = CZ_F$  en (25) :

$$(26) \quad Y = C[\Psi a_F + E_0(Z_F)]$$

es decir:

$$(27) \quad e_Y = Y - CE_0(Z_F) = C\Psi a_F$$

donde:

$$(28) \quad e_y \sim N_m(0, C\Psi(I_H \otimes \Sigma)\Psi'C')$$

para evaluar la hipótesis de compatibilidad se puede construir la siguiente estadística K:

$$(29) \quad K = e_y' [C\Psi(I \otimes \Sigma)\Psi'C']^{-1} e_y$$

como la estadística K es una forma cuadrática relacionada a la distribución normal dada en (28) se tiene que bajo la hipótesis nula:

$$(30) \quad K \sim \chi_{(m)}^2$$

Si la estadística K no es lo suficientemente alta, la hipótesis nula no podrá ser rechazada y por lo tanto se puede derivar intervalos de confianza para los pronósticos restringidos, estos son especificados en la ecuación (23).

#### **D. Corrección del sesgo por transformación logaritmo en modelos VAR.**

Dadas las características de algunas series macroeconómicas colombianas, una de las transformaciones,  $T(.)^5$ , necesarias para obtener estacionariedad en varianza es la llevada a cabo mediante la función logaritmo. Al realizar pronósticos utilizando modelos VAR o VEC donde algunas de las series del modelo han sido transformadas bajo dicha función, los pronósticos en niveles de las series son obtenidos simplemente aplicando la función inversa, es decir, la función exponencial. Sin embargo, al igual que en el caso de series de tiempo univariado, modelos ARIMA, es ampliamente conocido que si se realiza una transformación de potencia, los pronósticos en niveles de las series deben ser corregidos por el sesgo producido por esta transformación. En esta sección, se realiza una aproximación para corregir este sesgo en el caso multivariado, la cual

---

<sup>5</sup> En esta sección el tipo de transformaciones que se analizan son sólo aquellas referentes a transformaciones de potencia del tipo Box-Cox (1964) que generalmente se utilizan para estabilizar la varianza.

corresponde a la generalización planteada por Guerrero (1991) para el caso univariado.

Si se supone que una serie multivariada de tiempo ha sido transformada por medio de una función  $T(\cdot)$ , por ejemplo logaritmo, es de interés obtener pronósticos de la serie original  $\{\bar{Z}_t\}$  a partir de los pronósticos del modelo  $\{T(\bar{Z}_t)\}$ .

Notacionalmente,  $\bar{\mu}_j \equiv E_t[T(\bar{Z}_{t+j})]$  se define como el pronóstico de la serie transformada, donde el subíndice  $t$  en el valor esperado indica que éste se realiza condicional a la información conocida hasta el tiempo  $t$ , en tanto que los pronósticos de la serie original son  $E_t[\bar{Z}_{t+j}]$ ,  $j = 1, 2, \dots$ .

### 1. Aproximación de $\bar{\mu}_j$

Expandiendo a  $T(\bar{Z}_t)$  mediante series multivariadas de Taylor alrededor de  $E_t(\bar{Z}_{t+j})$  se tiene:

$$(31) \quad T(\bar{Z}_{t+j}) \cong T[E_t(\bar{Z}_{t+j})] + \sum_{l=1} \left[ \frac{\partial T}{\partial Z_l} [E_t(\bar{Z}_{t+j})] \right] [Z_{l,t+j} - E_t(Z_{l,t+j})] + \frac{1}{2} \sum_{i,l} \left[ \frac{\partial^2 T}{\partial Z_i \partial Z_l} [E_t(\bar{Z}_{t+j})] \right] [Z_{i,t+j} - E_t(Z_{i,t+j})][Z_{l,t+j} - E_t(Z_{l,t+j})]$$

donde  $Z_{l,t+j}$  representa el componente  $l$  del vector  $\bar{Z}_{t+j}$ .

Tomando valor esperado condicional<sup>6</sup> a (31) se obtiene:

$$(32) \quad \bar{\mu}_j \cong T[E_t(\bar{Z}_{t+j})] + \frac{1}{2} \sum_{i,l} \left[ \frac{\partial^2 T}{\partial Z_i \partial Z_l} [E_t(\bar{Z}_{t+j})] \right] [\text{cov}_t(Z_{i,t+j}, Z_{l,t+j})]$$

### 2. Aproximación de la varianza de $T(\bar{Z}_t)$

Tomando la aproximación dada por los dos primeros términos de (31) se tiene:

$$(33) \quad T(\bar{Z}_{t+j}) \cong T[E_t(\bar{Z}_{t+j})] + \sum_{l=1} \left[ \frac{\partial T}{\partial Z_l} [E_t(\bar{Z}_{t+j})] \right] [Z_{l,t+j} - E_t(Z_{l,t+j})]$$

<sup>6</sup> Bajo esta notación el subíndice  $t$  del valor esperado en (2) indica que éste se encuentra condicionado a la información disponible hasta el tiempo  $t$ .

tomando varianzas condicionales a ambos lados de la ecuación (33):

$$(34) \quad \vec{V}_t[T(\bar{Z}_{t+j})] \cong E_t \left\{ \left( \sum_{i=1}^n \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right] [Z_{i,t+j} - E_t(Z_{i,t+j})] \right) \left( \sum_{i=1}^n \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right]^T [Z_{i,t+j} - E_t(Z_{i,t+j})] \right) \right\}$$

donde  $\vec{V}[\cdot]$  corresponde a la matriz de varianzas covarianzas y el superíndice  $T$  indica que el vector es transpuesto.

Simplificando la anterior expresión resulta:

$$(35) \quad \vec{V}_t[T(\bar{Z}_{t+j})] \cong E_t \left\{ \sum_{i,j} \left[ \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right] \left[ \frac{\partial T}{\partial Z_j} [E_t(\bar{Z}_{t+j})] \right]^T [Z_{i,t+j} - E_t(Z_{i,t+j})] [Z_{j,t+j} - E_t(Z_{j,t+j})] \right] \right\}$$

utilizando la definición de covarianza en (35):

$$(36) \quad \vec{V}_t[T(\bar{Z}_{t+j})] \cong \sum_{i,j} \left[ \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right] \left[ \frac{\partial T}{\partial Z_j} [E_t(\bar{Z}_{t+j})] \right]^T [Cov_t(Z_{i,t+j}, Z_{j,t+j})] \right]$$

tomando sólo la diagonal de las matrices a ambos lados de la ecuación (36):

$$(37) \quad diag(\vec{V}_t[T(\bar{Z}_{t+j})]) \cong \sum_{i,j} \left[ \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right] * \left[ \frac{\partial T}{\partial Z_j} [E_t(\bar{Z}_{t+j})] \right] [Cov_t(Z_{i,t+j}, Z_{j,t+j})] \right]$$

donde el producto \* indica multiplicación elemento por elemento.

### 3. Obtención del resultado final para la transformación logarítmica.

Debido a que las transformaciones usualmente utilizadas, entre estas el logaritmo, son transformaciones que dependen solo de la misma variable<sup>7</sup>.

Las expresiones (32) y (37) pueden ser simplificadas respectivamente a:

$$(38) \quad \bar{\mu}_j \cong T[E_t(\bar{Z}_{t+j})] + \frac{1}{2} \sum_i \left[ \frac{\partial^2 T}{\partial Z_i^2} [E_t(\bar{Z}_{t+j})] [V_t(Z_{i,t+j})] \right]$$

$$(39) \quad diag(\vec{V}_t[T(\bar{Z}_{t+j})]) \cong \sum_i \left[ \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right] * \left[ \frac{\partial T}{\partial Z_i} [E_t(\bar{Z}_{t+j})] \right] [V_t(Z_{i,t+j})] \right]$$

donde  $V[\cdot]$  corresponde a la varianza.

<sup>7</sup> Por ejemplo, para  $\{Z_{i,t}\}$ , la  $i$ -ésima variable del vector  $\{\bar{Z}_t\}$ ,  $T(Z_{i,t})$  sólo depende de  $Z_i$  y no de las demás variables del vector.

Si se trabaja con la transformación logaritmo<sup>8</sup> en la ecuación (38) con k variables se tiene:

$$(40) \quad \begin{bmatrix} \mu_{1,j} \\ \vdots \\ \mu_{k,j} \end{bmatrix} \cong \begin{bmatrix} \ln E_t(Z_{1,t+j}) \\ \vdots \\ \ln E_t(Z_{k,t+j}) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \frac{-1}{E_t(Z_{1,t+j})^2} V_t(Z_{1,t+j}) \\ \vdots \\ \frac{-1}{E_t(Z_{k,t+j})^2} V_t(Z_{k,t+j}) \end{bmatrix}$$

y en la ecuación (39):

$$(41) \quad \begin{bmatrix} V_t[\ln(Z_{1,t+j})] \\ \vdots \\ V_t[\ln(Z_{k,t+j})] \end{bmatrix} \cong \begin{bmatrix} \frac{1}{E_t(Z_{1,t+j})^2} V_t(Z_{1,t+j}) \\ \vdots \\ \frac{1}{E_t(Z_{k,t+j})^2} V_t(Z_{k,t+j}) \end{bmatrix}$$

finalmente, reemplazando los términos  $V_t(Z_{i,t+j})$ ,  $i=1,\dots,k$  de (41) en (40) y despejando  $E_t[Z_{i,t+j}]$  se obtiene:

$$(42) \quad \begin{bmatrix} E_t(Z_{1,t+j}) \\ \vdots \\ E_t(Z_{k,t+j}) \end{bmatrix} \cong \begin{bmatrix} \exp\left(\mu_{1,j} + \frac{1}{2} V_t[\ln(Z_{1,t+j})]\right) \\ \vdots \\ \exp\left(\mu_{k,j} + \frac{1}{2} V_t[\ln(Z_{k,t+j})]\right) \end{bmatrix}$$

es decir, bajo esta metodología, para modelos VAR, la corrección del sesgo causado por la transformación logaritmo en los pronósticos de la variable  $Z_i$  en el tiempo  $t+j$ , implica que el antilogaritmo del pronóstico, para el logaritmo de esta variable, debe ser multiplicado por  $\exp\left(\frac{1}{2} V_t[\ln(Z_{i,t+j})]\right)$ . De lo contrario los pronósticos de las series en niveles son subestimados.

### III Aplicaciones

#### A. Modelo de Corrección de Errores

La metodología anterior es directamente aplicable a modelos VAR donde todas las series involucradas son estacionarias. Si estas series no son  $I(0)$  pero se

---

<sup>8</sup> Es decir, para k variables: 
$$T \begin{bmatrix} Z_{1,t} \\ \vdots \\ Z_{k,t} \end{bmatrix} = \begin{bmatrix} \ln(Z_{1,t}) \\ \vdots \\ \ln(Z_{k,t}) \end{bmatrix}$$

encuentran cointegradas, Lütkepohl (1991) muestra como el modelo VAR(p) descrito en la ecuación (1) puede ser representado a través del siguiente modelo de corrección de errores:

$$(43) \quad \Delta Z_t = v + F_1 \Delta Z_{t-1} + \dots + F_{p-1} \Delta Z_{t-p+1} - \Pi Z_{t-1} + a_t$$

donde:

$$(44) \quad F_i = -(\phi_{i+1} + \dots + \phi_p), \quad i = 1, \dots, p-1$$

$$(45) \quad \Pi = I - \phi_1 - \dots - \phi_p$$

donde la matriz  $\Pi$  de orden  $k \times k$  contiene la información sobre las relaciones de largo plazo entre las variables, y su rango indica el número de vectores de cointegración. Las matrices  $\phi_i$  son definidas en la ecuación (1).

Por lo tanto, si las variables especificadas en el modelo VAR se encuentran cointegradas podemos estimar la ecuación (43)<sup>9</sup>, y por medio de las recurrencias (44) y (45) podemos obtener las estimaciones de los parámetros y de la matriz de covarianzas del vector de errores del modelo inicial especificado en la ecuación (1), y de esta forma aplicar la metodología para pronósticos condicionados, sugerida en las secciones anteriores.

## B. Aplicación de la Metodología.

### 1. Simulación.

Como ejercicio empírico, esta metodología se aplicó a datos simulados del siguiente modelo VAR(2) sobre dos variables diferenciadas,  $\Delta X_t$  y  $\Delta Y_t$ :

$$(46) \quad \begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} \Delta X_{t-1} \\ \Delta Y_{t-1} \end{bmatrix} + \begin{bmatrix} -0.3 & -0.2 \\ 0.09 & -0.5 \end{bmatrix} \begin{bmatrix} \Delta X_{t-2} \\ \Delta Y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \quad t = 1, \dots, 154$$

donde  $\bar{\varepsilon}_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$  se distribuye normal bivariada con valor esperado  $\bar{0}$  y matriz de

varianzas-covarianzas  $\Sigma = \begin{bmatrix} 1.0 & 0.2 \\ 0.2 & 0.4 \end{bmatrix}$ .

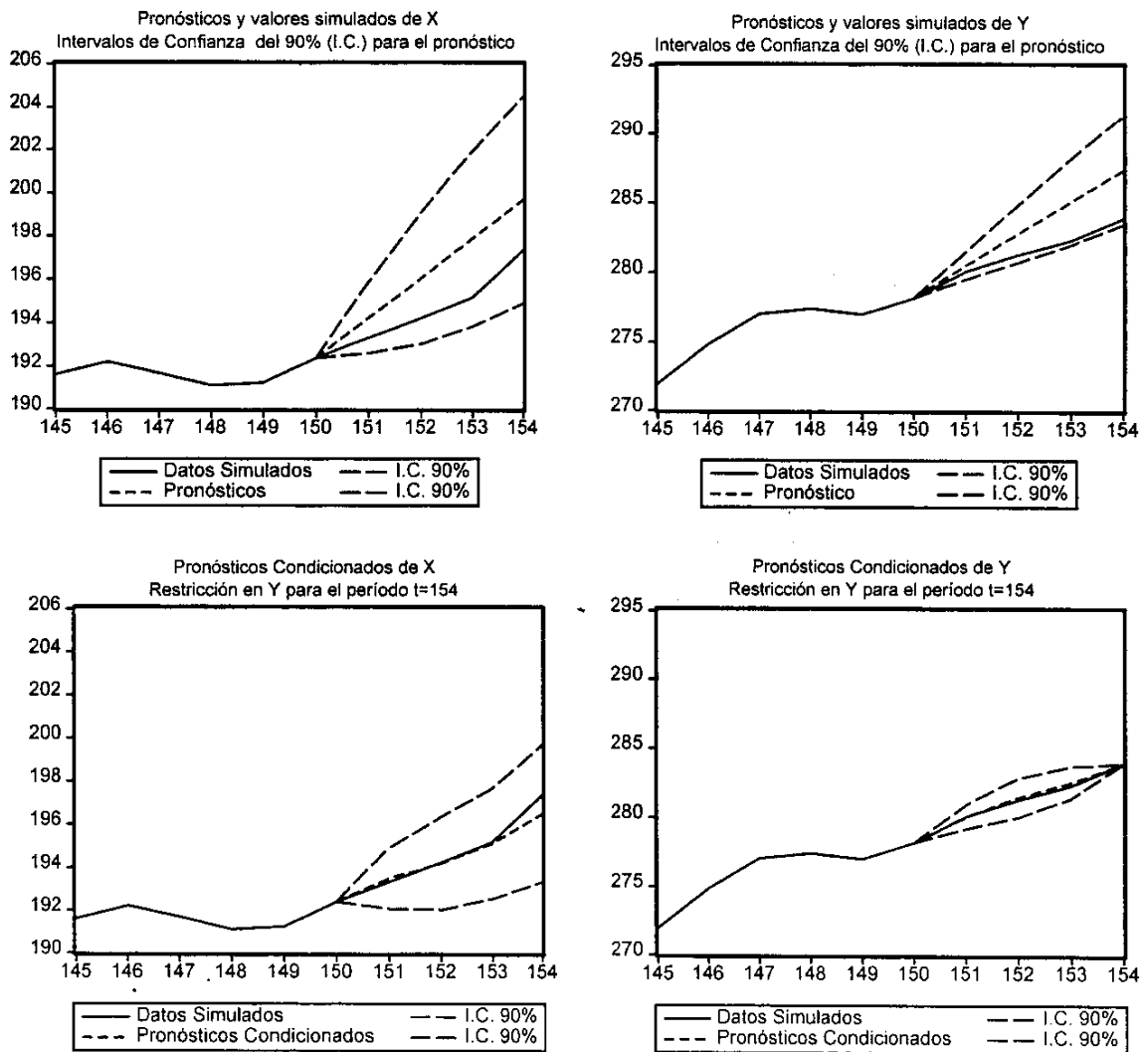
<sup>9</sup> Véase Lütkepohl (1991).

Una vez simulados los 154 datos para las variables X y Y. Se estimó un modelo VAR(2) utilizando las primeras 150 observaciones, las restantes cuatro observaciones son utilizadas para posteriores comparaciones con los pronósticos del modelo. Los resultados de la estimación son los siguientes:

$$(47) \begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = \begin{bmatrix} 0.83 \\ (0.19) \\ 1.02 \\ (0.12) \end{bmatrix} + \begin{bmatrix} 0.49 & 0.33 \\ (0.08) & (0.12) \\ 0.32 & 0.61 \\ (0.05) & (0.08) \end{bmatrix} \begin{bmatrix} \Delta X_{t-1} \\ \Delta Y_{t-1} \end{bmatrix} + \begin{bmatrix} -0.29 & -0.22 \\ (0.09) & (0.11) \\ 0.12 & -0.47 \\ (0.06) & (0.07) \end{bmatrix} \begin{bmatrix} \Delta X_{t-2} \\ \Delta Y_{t-2} \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_{1,t} \\ \hat{\varepsilon}_{2,t} \end{bmatrix} \quad t = 1, \dots, 150$$

los errores estándar de los parámetros son indicados entre paréntesis.

FIGURA 1



## 2. Pronósticos Condicionados.

Utilizando la metodología descrita en la sección II y diferentes conjuntos de condiciones o metas se estimaron pronósticos condicionados sobre el modelo VAR(2) anteriormente descrito, tomando como horizonte de pronóstico a  $t=151, \dots, 154$ .

### a) Metas sobre el valor simulado para $Y_{154}$ .

Aunque en la Figura 1 se observa que los intervalos de confianza de los pronósticos del modelo VAR(2), estimado en (47), contienen los valores simulados, la trayectoria no es exactamente la misma. En este ejercicio se impone una sola restricción, sobre la segunda variable, de tal forma que el valor del pronóstico de la variable  $Y$  en  $t=154$  coincide con el valor verdadero o simulado. Los resultados, mostrados en las dos gráficas inferiores de la Figura 1, muestran que tan solo incorporando esta restricción se reproduce la trayectoria total para las dos series de forma muy aproximada.

### b) Metas incluyendo choques en ambas variables en $t=151$ .

El objetivo de este ejercicio es incorporar escenarios más realistas con respecto a las restricciones de los pronósticos fuera de muestra. En el ejercicio anterior la restricción utilizada fue extraída usando mecanismos de simulación exactamente iguales para el período de estimación ( $t=1, \dots, 150$ ) y para el período de pronósticos fuera de muestra ( $t=151, \dots, 154$ ). En este ejercicio se comparan los pronósticos condicionados con los pronósticos del modelo cuando se ha realizado un choque en ambas variables en el primer período fuera de muestra<sup>10</sup>.

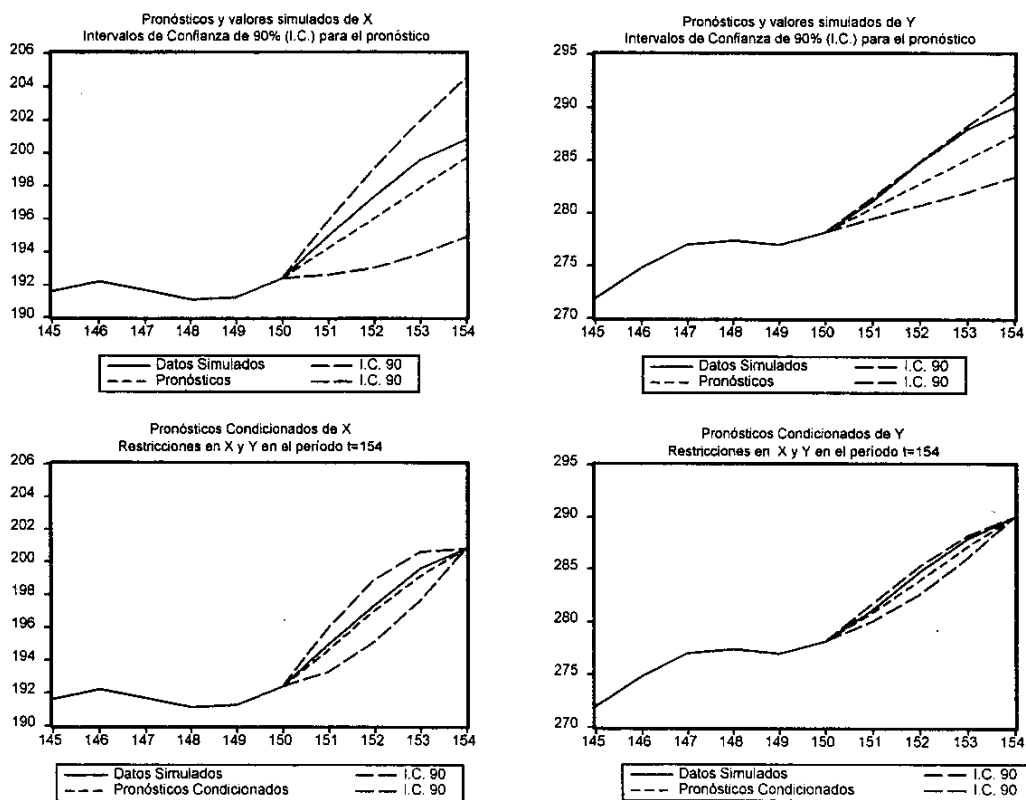
Para incluir los efectos sobre las simulaciones de las series en el período fuera de muestra, tal como se mencionó anteriormente, se realizó la siguiente modificación a las simulaciones originales; los valores de las series  $X_t$  y  $Y_t$  para el período  $t=151$  se tomaron como el límite superior del intervalo de pronóstico del 90% de las simulaciones iniciales, lo cual podría ser interpretado como el



resultado de un choque positivo en ambas variables<sup>11</sup> o como uno de los posibles resultados determinados aleatoriamente. Los valores simulados posteriores a este periodo,  $t=152, \dots, 154$ , son simplemente determinados por la estructura del modelo.

Dado que los valores de las variables para la muestra utilizada en la estimación ( $t=1, \dots, 150$ ) no fueron modificados, los pronósticos del modelo son iguales a los obtenidos en el ejercicio anterior. Sin embargo, como se observa en las dos primeras gráficas de la figura 2, aunque los valores simulados con los dos “choques positivos” se encuentran dentro del intervalo de confianza de los pronósticos, estos se encuentran cercanos al límite superior de los intervalos de confianza.

**FIGURA 2**  
Simulaciones incluyendo choques en X y Y en el periodo  $t=151$



<sup>10</sup> Es decir, los pronósticos verdaderos o simulados contienen una pequeña desviación con respecto a los pronósticos basados sólo en los datos utilizados en la estimación del modelo.

<sup>11</sup> Lo cual no es muy improbable, dada la estructura del modelo, ya que estos valores están contenido dentro del intervalo de confianza del 90% para los pronósticos.

En este ejercicio se impusieron dos restricciones sobre los pronósticos de las series X y Y, las cuales especifican que estos sean iguales a los nuevos valores simulados en el período  $t=154$ . El resultado de este ejercicio se presenta en las dos gráficas inferiores de la Figura 2, donde se observa que estas nuevas trayectorias de los pronósticos condicionados, imponiendo las dos restricciones, son muy similares a las obtenidas a partir de los datos simulados.

### 3) Prueba de Compatibilidad

Debido a que las restricciones impuestas sobre los pronósticos, presentadas en los dos ejercicios anteriores, están diseñadas a partir de valores simulados del mismo modelo, es de esperarse que los test de compatibilidad, explicado en la sección II.C, indiquen que los pronósticos del modelo y las restricciones no son significativamente diferentes<sup>12</sup>. Este tipo de resultados es obtenido en las pruebas de compatibilidad para los ejercicios realizados en las secciones 2a y 2b:

**Tabla 1**  
**Pruebas de Compatibilidad**

<b>Restricción sobre los pronósticos</b>	<b>Estadística</b>	<b>Grados de Libertad</b>	<b>P-Value</b>
Para la serie $Y_t$ en el periodo $t=154$ (sección 2a)	2.1434	1	0.143
Para las series $X_t$ y $Y_t$ en el periodo $t=154$ (sección 2b)	1.6124	2	0.447

Como segundo ejercicio se evaluó la compatibilidad entre los pronósticos del modelo estimado en (47) con respecto a restricciones en las cuales se especificaba, para las series X y Y, valores menores al límite inferior del intervalo

<sup>12</sup> Es importante notar, que la prueba de compatibilidad en lugar de comparar directamente los pronósticos del modelo con las restricciones o metas, se compara una la combinación lineal de los pronósticos, definida por los valores de las filas de la matriz C (especificada en la sección II.B), con las restricciones.

de pronóstico del 90% en el período  $t=154$ <sup>13</sup>. Los resultados de este ejercicio se presentan en la tabla 2 donde se indica una compatibilidad baja, resultados esperados puesto que estas restricciones no se encuentran incluidas en los intervalos de pronóstico del 90% para el modelo.

**Tabla 2**  
**Pruebas de Compatibilidad**

Restricciones impuestas en $X_{154}$	Restricciones impuestas en $Y_{154}$	Estadística	Grados de Libertad	P-Value
192.0	-----	6.980	1	0.00824
-----	289.0	12.227	1	0.00047
192.0	289.0	12.228	2	0.00221

#### IV Conclusiones

La metodología de pronósticos condicionados en modelos multivariados de series de tiempo, específicamente sobre modelos VAR, se presenta como óptima. Lo anterior en la medida que estima los pronósticos de un modelo VAR restringidos a un conjunto específico de metas sobre las variables involucradas, permitiendo así evaluar si las metas impuestas ex-ante son y continúan siendo compatibles con la evolución histórica de las variables. Adicionalmente, si la hipótesis de compatibilidad no es rechazada, se pueden construir intervalos de confianza sobre los pronósticos condicionados.

Es de señalar, que esta metodología puede ser igualmente aplicada utilizando modelos VAR o modelos de corrección de errores (VEC) si las variables analizadas están cointegradas.

En este documento se demuestra que los pronósticos de las series en niveles para un modelo VAR, donde algunas de las series han sido transformadas

<sup>13</sup> Los intervalos de confianza del 90% para los pronósticos de las series X y Y en el período  $t=154$  son (192.7, 201.9) y (279.6, 288.1) respectivamente.

bajo logaritmo, deben ser corregidos por el sesgo producido por esta transformación<sup>14</sup>, de lo contrario estos pronósticos son subestimados<sup>15</sup>.

---

<sup>14</sup> Una aproximación al valor con el cual se corrige este sesgo es encontrado en la sección II.d del documento.

<sup>15</sup> Es decir, cuando algunas de las series de un modelo VAR han sido transformadas bajo logaritmo y los pronósticos de las series en niveles son obtenidos aplicando la función inversa, en este caso la función exponencial, estos pronósticos finales son menores comparados con los que teóricamente son correctos.

## Referencias

Box, G.E.P. and Cox, D.R. "An Analysis of Transformations", *Journal of the Royal Statistical Society*, B-26, 1964.

Box, G.E.P. y Jenkins, G.M. *Time Series Analysis, forecasting and control*, San Francisco: Holden Day, 1970.

Cromwell, J. B. et al. *Multivariate tests for time series models*. Sage publications, 1994.

Doornik, J. A. & Hansen, H. "An omnibus test for univariate and multivariate normality". Nuffield College, Oxford, 1994.

Godfrey, L. G. *Misspecification tests in econometrics. The Lagrange Multiplier principle and other applications*. Cambridge University Press, 1988.

Guerrero, V.M. "Optimal Conditional ARIMA Forecasts", *Journal of Forecasting* V 8 pp 215-229 , 1989.

Guerrero, V.M. "Pronósticos restringidos con modelos de transferencia", Instituto Autónomo de México, ITAM, mimeo, 1991.

Guerrero, V.M. "Time Series Supported by Power Transformations", Instituto Autónomo de México, ITAM, mimeo, 1991

Haugh, L.D. y Box, G.E.P. "Identification of dynamic regression (distributed lag) models connecting two time series", *Journal of the American Statistical Association* 72, pp. 121-130, 1977.

Johansen, S. y Juselius, K. . "Maximum Likelihood Estimation and Inference on Cointegration - with Applications to the Demand for Money". *Oxford Bulletin of Economics and Statistics*, 52, pp. 169-210, 1990.

Judge, G. G. et al. *Introduction to the theory and practice of econometrics*. John Wiley & Sons, 1988.

Lütkepohl, H. *Forecasting Aggregated Vector ARMA Processes*. Springer-Verlag. Heidelberg, 1987.

Lütkepohl, H. *Introduction to Multiple Time Series Analysis*. Springer-Verlag. Heidelberg, 1991.

Melo, L.F. y Oliveros H. "Pronósticos condicionados: método y aplicación al caso de la inflación 1991", *Ensayos sobre Política Económica* v. 20 pp. 87-105, B.R., 1991.

Serfling, R.J. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, 1980.

Shenton, L. R. & Bowman, K. O. (1977). "A bivariate model for the distribution of  $\sqrt{b_1}$  and  $b_2$ ". *Journal of the American Statistical Association*, 72, pp. 206-211, 1977.

Spanos, Aristis. *Statistical foundations of econometric modeling*. Cambridge University Press, 1986.

Wei, William. *Time Series Analysis: Univariate and Multivariate Methods*. Addison-Wesley, 1993.

Wold, Herman. *A study in the Analysis of Stationary Time Series*. Uppsala, Sweden: Almqvist and Wiksell, 1938.

## ANEXO

**1. Si  $a_t \sim N_k(0, \Sigma)$  vector ruido blanco, entonces  $a_F \sim N_{kH}(0, I_H \otimes \Sigma)$**

Demostración:

Sea  $a_t = (a_{1t}, a_{2t}, \dots, a_{kt})'$  donde  $a_t \sim N_k(0, \Sigma)$  es un vector ruido blanco *iid* para  $t = 1, \dots, H$  y  $a_F = (a'_1, a'_2, \dots, a'_H)'$  entonces:

$$\begin{aligned}
 f_{a_F}(a) &= \prod_{i=1}^H f_{a_i}(a_i) \\
 &= \prod_{i=1}^H \left[ (2\pi)^{-k/2} (\det \Sigma)^{-1/2} \exp(-1/2 a'_i \Sigma^{-1} a_i) \right] \\
 &= (2\pi)^{-kH/2} (\det \Sigma)^{-H/2} \exp \left[ -1/2 \sum_{i=1}^H a'_i \Sigma^{-1} a_i \right] \\
 &= (2\pi)^{-kH/2} (\det \Sigma)^{-H/2} \exp \left\{ -1/2 [a'_1, a'_2, \dots, a'_H] (I^{-1} \otimes \Sigma^{-1}) [a'_1, a'_2, \dots, a'_H]' \right\} \\
 &= (2\pi)^{-kH/2} [\det(I)^k \det(\Sigma)^H]^{-1/2} \exp[-1/2 A'(I \otimes \Sigma)^{-1} A] \\
 &= (2\pi)^{-kH/2} [\det(I \otimes \Sigma)]^{-1/2} \exp[-1/2 A'(I \otimes \Sigma)^{-1} A]^{16}
 \end{aligned}$$

por lo tanto  $a_F \sim N_{kH}(0, I_H \otimes \Sigma)$ .

## 2. Distribución de los pronósticos condicionados cuando los parámetros del modelo VAR

**son desconocidos.**

Mediante la metodología utilizada en la sección II se obtuvo la siguiente expresión final para la estimación de los pronósticos condicionados:

$$(48) \quad \hat{Z}_F = E_0(Z_F) + \hat{A}[Y - CE_0(Z_F)]$$

donde :

$$(49) \quad \hat{A} = \Psi(I \otimes \Sigma)\Psi' C' [C\Psi(I \otimes \Sigma)\Psi' C']^{-1}$$

---

<sup>16</sup>  $\det(I_H)^k \det(\Sigma)^H = \det(I_H \otimes \Sigma)$  véase, entre otros, Spanos (1986).

Estos resultados están sujetos al supuesto de que los parámetros del modelo VAR, contenidos en las matrices  $\Psi$  y  $\Sigma$ , son conocidos. Sin embargo, si estos parámetros son desconocidos y por lo tanto deben ser estimados, las ecuaciones (48) y (49) deben ser reformuladas como sigue:

$$(50) \quad \hat{Z}_F = E_0(Z_F) + \hat{A}[Y - CE_0(Z_F)]$$

donde:

$$(51) \quad \hat{A} = \hat{\Psi}(I \otimes \hat{\Sigma})\hat{\Psi}'C'[C\hat{\Psi}(I \otimes \hat{\Sigma})\hat{\Psi}'C']^{-1}$$

El error de pronóstico condicionado definido como  $\hat{e}_F \equiv Z_F - \hat{Z}_F$  utilizando la ecuación (16) es igual a:

$$(52) \quad \hat{e}_F = \hat{B}a_F$$

donde:

$$(53) \quad \hat{B} \equiv (\hat{A}C - I)\hat{\Psi}$$

y  $a_F$  es un vector ruido blanco con distribución normal multivariada.

Como los estimadores de los parámetros del modelo VAR son consistentes, es decir, la sucesión de estimadores<sup>17</sup> converge en probabilidad a los valores poblacionales y como esta propiedad se preserva para funciones continuas<sup>18</sup> (teorema 1.7 Serfling, 1980) entonces se tiene<sup>19</sup>:

$$(54) \quad \{\hat{A}\} \xrightarrow{p} \hat{A}, \quad \{\hat{B}\} \xrightarrow{p} B, \quad \{\hat{Z}_F\} \xrightarrow{p} \hat{Z}_F, \quad \{Z_F - \hat{Z}_F\} \xrightarrow{p} \{Z_F - \hat{Z}_F\}$$

<sup>17</sup> El índice de la sucesión es el tamaño de la muestra.

<sup>18</sup>  $\hat{A}$  y  $\hat{B}$  son funciones continuas de las matrices de los estimadores de los parámetros del modelo VAR.

<sup>19</sup> En las expresiones a continuación se utiliza la siguiente notación:

- $\{\hat{\Theta}\} \xrightarrow{p} \Theta$  indica que la sucesión  $\{\hat{\Theta}_n\}$  converge en probabilidad a  $\Theta$ ; i.e.  $\lim_{n \rightarrow \infty} \Pr(|\hat{\Theta}_n - \Theta| < \varepsilon) = 1$ , esto implica a *grosso* modo que incrementando el tamaño de la muestra indefinidamente se puede estar casi seguro de encontrar a  $\hat{\Theta}_n$  tan cercano como se desee de  $\Theta$ .
- $\{X\} \xrightarrow{d} Y$  indica que la sucesión  $\{X_n\}$  converge en distribución a  $Y$ ; i.e.  $\lim_{n \rightarrow \infty} F_n(x) = F(y)$ ; es decir, a medida que se aumenta el tamaño de la muestra, la distribución de probabilidad de  $X_n$  tiende a la distribución de probabilidad de  $Y$  para cada punto de continuidad  $x$  de  $F(x)$ .



así, en particular:

$$(55) \quad \{Z_F - \hat{Z}_F\} \xrightarrow{P} B a_F$$

donde  $B \equiv (\hat{A}C - I)\Psi$ .

Convergencia en probabilidad implica convergencia en distribución, entonces:

$$(56) \quad \{Z_F - \hat{Z}_F\} \xrightarrow{d} B a_F$$

en las ecuaciones indicadas en (22) se mostró que:

$$(57) \quad B a_F \sim N_{kH}(\bar{0}, (I - \hat{A}C)\Psi(I \otimes \Sigma)\Psi')$$

es decir, la distribución asintótica de  $\{Z_F - \hat{Z}_F\}$  es normal multivariada, con los parámetros especificados en (57), que es la misma que se había encontrado en el documento asumiendo parámetros conocidos. Por lo tanto, los resultados de este documento siguen siendo válidos asintóticamente aún cuando los parámetros del modelo VAR son desconocidos y por tanto estimados.