

Pronósticos con restricciones para series de tiempo

Jesús Alberto Albarracín
Harney Palacios Bejarano



Maestría en Matemáticas Aplicadas
Departamento de Ciencias Básicas
Universidad EAFIT
2011

Pronósticos con restricciones para series de tiempo

Jesús Alberto Albarracín
Harney Palacios Bejarano

Trabajo de grado presentado como requisito parcial para optar el título de
Maestría en Matemáticas Aplicadas

Director
Ermilson Velásquez Ceballos
PhD. Ciencias Matemáticas



Maestría en Matemáticas Aplicadas
Departamento de Ciencias Básicas
2011

Resumen

Uno de los aspectos centrales y que actualmente adquiere gran relevancia es la Modelación de series de tiempo con fines de predicción. Los modelos ARIMA por su misma estructura no son muy potentes en este sentido.

En este trabajo se presentan las consideraciones teóricas de los modelos Autorregresivos y de Medias Móviles Integrados (ARIMA), los modelos de Vectores Autorregresivos (VAR) y se desarrolla el procedimiento empleado por Víctor Guerrero (2002) que permite obtener predicciones condicionadas a determinadas metas o restricciones.

Los elementos formales se ilustran con un ejemplo numérico y finalmente se realiza un caso de aplicación para la economía colombiana utilizando la variable PIB.

CONTENIDO

INTRODUCCIÓN	5
1. PRELIMINARES	7
1.1. PROCESO ESTOCÁSTICO	7
1.1.1. Procesos estacionarios	8
1.1.2. Procesos lineales	9
1.2. MODELOS PARA SERIES DE TIEMPO UNIVARIADAS	9
1.2.1. Modelos Autorregresivos (AR)	9
1.2.2. Modelos de Medias Móviles	13
1.2.3. Modelos mixtos: Autorregresivos-Medias móviles (ARMA)	16
1.3. PROCESOS NO ESTACIONARIOS	18
1.3.1. Modelos ARIMA	18
1.4. Pronósticos óptimos de Series de tiempo	20
1.4.1. Pronóstico de series estacionarias	20
2. PRONÓSTICOS CON RESTRICCIONES	24
2.1. RESTRICCIONES CIERTAS, COMPATIBLES CON LA HISTORIA	24
2.2. RESTRICCIONES INCIERTAS, COMPATIBLES O NO CON LA HISTORIA	30
2.3. RESTRICCIONES CIERTAS INCOMPATIBLES CON LA HISTORIA	33
3. PRONÓSTICOS RESTRINGIDOS PARA SERIES MÚLTIPLES	42
3.1. MODELOS VAR Y VECM	42
3.1.1. Aspectos metodológicos	42
3.2. PRONÓSTICOS CON RESTRICCIONES PARA MODELOS VAR	47
4. APLICACIONES	52
4.1. SEGUIMIENTO AL PIB DE COLOMBIA	52
CONCLUSIONES	61
APÉNDICE	62
BIBLIOGRAFÍA	63

Introducción

La modelación de Series de tiempo económicas es un aspecto fundamental en la teoría econométrica; los aportes de Box y Jenkins permitieron la construcción de modelos que consideran la historia de la serie (modelos autorregresivos), las innovaciones (modelos de medias móviles) o combinaciones de ambos (modelos ARMA). Si las series presentan raíces unitarias, éstas se diferencian con lo cual se utilizan los modelos Integrados ARIMA.

Después de determinado el modelo apropiado siguiendo la metodología Box-Jenkins, se pueden generar pronósticos que permitan tomar decisiones; teniendo en cuenta la naturaleza de la serie, suelen presentarse situaciones externas que modifican la trayectoria de la serie y que pueden darse en un periodo de tiempo corto o que pueden aparecer de forma abrupta para luego lentamente desaparecer; por ejemplo la creación de un nuevo impuesto puede alterar el IPC (índice de precios al consumidor). Para estas circunstancias suele agregarse al componente estocástico modelado, otro llamado efecto de intervención.

Puede ocurrir que para determinadas series como la inflación o el Producto Interno Bruto (PIB) la autoridad económica del país, que en el caso de Colombia es el Banco de la República, defina metas que son establecidas para un horizonte que normalmente es de un año. El gobierno, puede entonces emprender acciones que posibiliten el cumplimiento de la política establecida como el aumento o disminución de las tasas de interés, el ingreso al mercado de divisas o el estímulo a la creación de empresas mediante la implementación de determinada ayuda. Es más, Lucas[10] en su crítica afirma que las predicciones basadas en la información histórica podrían ser inválidas cuando el gobierno tiene metas precisamente por la capacidad del gobierno para que con sus actuaciones tratar de llegar a las metas. El conocimiento por parte del analista económico de estas, puede servirle para mejorar los pronósticos de la variable en estudio. La meta, es entonces una información adicional o si se quiere decirlo una restricción porque se establece un límite en el crecimiento, el promedio, la suma, o el valor que se ha definido para la serie en un tiempo determinado.

Antes de considerar la información adicional, se deben tener en cuenta los pronósticos del modelo; aquellos que poseen el mínimo error cuadrático promedio, se determinan en forma matemática por la esperanza condicional de la serie en los tiempos $t + 1, t + 2, \dots, t + h$, dado los valores de la serie hasta el t . Estos pronósticos serán conocidos en el documento como irrestrictos.

Para incorporar las restricciones se determina una matriz para la cual pueda demostrarse que el nuevo pronóstico minimiza el error; en el documento se hace el desarrollo matemático ampliado al procedimiento seguido por Guerrero[6] y que antes habían establecido Trabetsi y Pankratz. Se obtiene una expresión que combina los pronósticos irrestrictos con las metas de forma que se generan unos pronósticos que se llamarán restringidos. Con estos pronósticos se suministra la representación de una trayectoria compatible con la historia

de la variable de interés y que permita al término de un horizonte cumplir con el objetivo planteado.

Con esta herramienta matemática puede también establecerse si la información histórica es compatible con las metas propuestas; esto se hace mediante la definición de un estadístico que utilice la restricción o restricciones bajo los supuestos de normalidad de los residuos generados.

Pero la metodología de pronósticos con restricciones también puede aplicarse al caso de la series múltiples; si se tiene un sistema de series que se pueden representar en un modelo de vectores autorregresivos (VAR), tal como lo explica Lutketpool[11], o si hay cointegración, con lo cual se podría permitir la implementación de un modelo de corrección de errores (VEC), el cual fue explicado por Engle y Granger (1987), se pueden generar pronósticos irrestrictos. De manera similar, puede encontrarse una representación para incorporar las restricciones y generar los pronósticos respectivos, lo cual es mostrado en este documento.

Luis Fernando Melo[12] describe los procedimientos que permiten la obtención de los pronósticos restringidos para un modelo VAR y propone un método para corregir el sesgo cuando a las variables de la especificación se les hace una transformación logarítmica para estacionarizarlas en varianza. Por otro lado, Peña y Guerrero[8] plantea cómo usar las restricciones para actualizar los pronósticos, la estimación de datos faltantes y pone de manifiesto la relación que hay entre la expresión matemática de la restricción y la que permite desagregar series temporales.

Para aplicar la teoría demostrada, se tiene en cuenta el Producto interno Bruto de Colombia, para el cual se estableció que al finalizar 2010 crecería en un 3,0%; se propone un modelo ARIMA para esta serie para el cual se generan los pronósticos en un horizonte de un año; teniendo en cuenta la meta establecida, se introduce la restricción y se establece una trayectoria para monitorear el cumplimiento de esta meta. Por último se hace un comparativo de los errores de los pronósticos irrestrictos con los que tienen restricciones.

1. Preliminares

1.1. Proceso estocástico

Definición 1.1. *Un proceso estocástico se define como una colección indexada de variables aleatorias $\{Y_t\}$, donde el subíndice t toma valores de un conjunto T dado. Con frecuencia T se toma como el conjunto de enteros no negativos que pueden equivaler a unidades de tiempo como meses, días o trimestres.*

Cuando se fija un valor en el tiempo, el proceso estocástico pasa a ser una variable aleatoria que tendrá su correspondiente distribución de probabilidad. Así, para $t = t_i$ la distribución de probabilidad será expresada por $F[Y(t_i)]$. Si en lugar de un valor fijamos dos valores del tiempo, se obtendría una variable bidimensional con una distribución bivalente. Así para $t = t_i$ y $t = t_j$ la distribución de probabilidad sería $F[Y(t_i), Y(t_j)]$. En general, para un conjunto finito de valores del tiempo, se obtendría una función de distribución conjunta. Así para t_1, t_2, \dots, t_n la función de distribución conjunta sería

$$F[Y(t_1), Y(t_2), \dots, Y(t_n)] \quad (1)$$

Se dice que un proceso estocástico está perfectamente caracterizado cuando se pueden determinar las funciones de distribución conjunta para cada conjunto finito de variables del proceso, es decir, para cada valor finito de n en (1). Sin embargo, la determinación de las características del proceso a partir de las funciones de distribución no es un procedimiento directo, por lo que se acostumbra a utilizar los momentos.

En un proceso estocástico al que se designará, para simplificar la notación, por Y_t la media o momento de primer orden se define de la siguiente forma

$$\mu_t = E(Y_t) \quad (2)$$

Como momentos de segundo orden respecto a la media es preciso considerar, además de la varianza, las covarianzas entre variables referidas a distintos momentos de tiempo o autocovarianzas a las que se designarán por

$$\gamma_{t,s} = Cov(Y_t, Y_s) = E(Y_t - \mu_t)(Y_s - \mu_s) \quad (3)$$

Cuando $s = t$ se tiene definida la varianza:

$$\gamma_{t,s} = \text{var}(Y_t) = E(Y_t - \mu_t)^2 \quad (4)$$

Como forma alternativa de caracterización de un proceso estocástico se utilizan los coeficientes de autocorrelación

$$\rho_{t,s} = \frac{Cov(Y_t, Y_s)}{\sqrt{\text{var}(Y_t)\text{var}(Y_s)}} \quad (5)$$

Las autocorrelaciones conjuntamente con las varianzas proporcionan idéntica información que las autocovarianzas. Sin embargo, es preferible utilizar las autocorrelaciones ya que éstas proporcionan unas medidas relativas a diferencia de lo que ocurre con las autocovarianzas que vienen afectadas por la escala que se utiliza.

1.1.1. Procesos estacionarios

Se dice que un proceso estocástico es estacionario en sentido estricto cuando al realizar un mismo desplazamiento en el tiempo de todas las variables de cualquier distribución conjunta finita, resulta que esta distribución no varía.

Considerando la función de distribución conjunta

$$F[Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}] \quad (6)$$

si se adopta el supuesto de que todos los elementos de la anterior distribución se desplazan m períodos, la nueva función de distribución conjunta será

$$F[Y_{t_1+m}, Y_{t_2+m}, \dots, Y_{t_k+m}] \quad (7)$$

Si el proceso es estacionario en sentido estricto se deberá verificar que

$$F[Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}] = F[Y_{t_1+m}, Y_{t_2+m}, \dots, Y_{t_k+m}] \quad (8)$$

e igualmente, se deberá obtener un resultado análogo para cualquier otra distribución conjunta que tenga carácter finito.

También en este caso es más complejo el análisis de la estacionariedad a partir de las funciones de distribución que si se efectúa a partir de los momentos. En contrapartida, el concepto de estacionariedad será más limitado. Se dice que un proceso es estacionario de primer orden, o en media, si se verifica que

$$E(Y_t) = \mu, \quad \forall t \quad (9)$$

es decir, en un proceso estacionario en media, la esperanza matemática permanece constante a lo largo del tiempo.

Se dice que un proceso estocástico es estacionario de segundo orden (o en sentido amplio) cuando se verifican las dos condiciones siguientes:

i) la varianza es finita y permanece constante a lo largo del tiempo, es decir

$$E(Y_t - \mu)^2 = \sigma^2 < \infty, \quad \forall t \quad (10)$$

ii) La autocovarianza entre dos períodos distintos de tiempo únicamente viene afectada por el lapso de tiempo transcurrido entre esos dos períodos. Así:

$$E(Y_{t+k} - \mu)(Y_t - \mu) = \gamma_k, \quad \forall t \quad (11)$$

En un proceso estacionario las autocorrelaciones quedan definidas de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (12)$$

1.1.2. Procesos lineales

Los procesos lineales se caracterizan porque se pueden representar como una combinación lineal de variables aleatorias. Los más comunes son: procesos puramente aleatorios, procesos autorregresivos, procesos de medias móviles y los obtenidos como combinación de estos dos últimos.

El proceso puramente aleatorio es el más simple de todos. Podría expresarse de la siguiente forma

$$Y_t = a_t \quad (13)$$

donde a_t goza de las siguientes propiedades:

$$E(a_t) = 0 \quad \forall t \quad (14a)$$

$$E(a_t^2) = \sigma_a^2 \quad \forall t \quad (14b)$$

$$E(a_t a_{t'}) = 0 \quad t \neq t' \quad (14c)$$

En conclusión, a_t se caracteriza por que su media es igual a cero y su varianza es constante a lo largo del tiempo y no existe relación entre valores referidos a momentos distintos del tiempo. Este es un proceso que se conoce como *Ruido blanco*.

1.2. Modelos para series de tiempo univariadas

En el trabajo a desarrollar es necesario considerar procesos: autorregresivos (AR), de medias móviles (MA), mixtos (ARMA) y modelos ARIMA.

1.2.1. Modelos autorregresivos (AR)

Un modelo *autorregresivo* de orden p , o abreviadamente un modelo $AR(p)$ se define de la siguiente forma:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad (15)$$

donde a_t es una variable aleatoria de *ruido blanco*. Si se utiliza el operador de retardos:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (16)$$

el modelo puede expresarse de forma compacta

$$\phi(L)Y_t = a_t \quad (17)$$

Si $p = 1$ se obtiene un proceso $AR(1)$

Modelo AR(1) Un modelo $AR(1)$ viene definido por

$$Y_t = \phi_1 Y_{t-1} + a_t \quad (18)$$

o, utilizando el operador de retardos, por

$$(1 - \phi_1 L)Y_t = a_t \quad (19)$$

El modelo definido en (18) se caracteriza porque $E(a_t Y_{t-\tau}) = 0 \quad \forall \tau > 0$.

Además, para que el modelo definido en (18) sea estacionario, la raíz del polinomio característico

$$1 - \phi_1 L = 0 \quad (20)$$

deben caer fuera del círculo unidad. Es decir

$$|L| = \left| \frac{1}{\phi_1} \right| > 1 \quad \text{o mejor} \quad (21a)$$

$$|\phi_1| < 1 \quad (21b)$$

Ahora bien, si el proceso se considera iniciado en $-\infty$, entonces para $|\phi_1| < 1$,¹ y cualquiera que sea el valor inicial, se verificará que $E(Y_t) = 0$, pero si en el modelo se introduce una constante, se tendrá que

$$Y_t = \delta + \phi_1 Y_{t-1} + a_t$$

Si el proceso es estacionario, se tendrá que

$$\mu = \frac{\delta}{1 - \phi_1}$$

Para determinar las autocovarianzas y las autocorrelaciones del modelo $AR(1)$ se sigue el siguiente procedimiento:

Sea Y_t como en (18), entonces multiplicando por $Y_{t-\tau}$ y tomando esperanzas se obtiene

$$Y_t Y_{t-\tau} = \phi_1 Y_{t-1} Y_{t-\tau} + a_t Y_{t-\tau} \quad (22a)$$

$$E(Y_t Y_{t-\tau}) = \phi_1 E(Y_{t-1} Y_{t-\tau}) + E(a_t Y_{t-\tau}) \quad (22b)$$

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + E(a_t Y_{t-\tau}) \quad (22c)$$

Para $\tau = 0$, resulta:

$$\gamma_0 = \phi_1 \gamma_{-1} + E(a_t Y_t) \quad (23a)$$

$$\gamma_0 = \phi_1 \gamma_{-1} + E(a_t [\phi_1 Y_{t-1} + a_t]) \quad (23b)$$

$$\gamma_0 = \phi_1 \gamma_{-1} + \phi_1 E(a_t Y_{t-1}) + E(a_t^2) \quad (23c)$$

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_a^2 \quad (23d)$$

¹Condición necesaria y suficiente para que un modelo $AR(1)$ sea estacionario.

Para $\tau = 1$, se obtiene:

$$\gamma_1 = \phi_1\gamma_0 + E(a_t Y_{t-1}) \quad (24a)$$

$$\gamma_1 = \phi_1\gamma_0 \quad (24b)$$

Reemplazando (24b) en (23d)

$$\begin{aligned} \gamma_0 &= \phi_1(\phi_1\gamma_0) + \sigma_a^2 \\ \gamma_0 &= \phi_1^2\gamma_0 + \sigma_a^2 \quad \text{y así} \\ \gamma_0 &= \frac{\sigma_a^2}{1 - \phi_1^2} \end{aligned}$$

Para $\tau > 0$ resulta que (22c) se transforma en

$$\gamma_\tau = \phi_1\gamma_{\tau-1}$$

luego, las autocorrelaciones serían

$$\begin{aligned} \rho_\tau &= \frac{\gamma_\tau}{\gamma_0} = \frac{\phi_1\gamma_{\tau-1}}{\gamma_0} \\ &= \phi_1\rho_{\tau-1} \end{aligned}$$

En un proceso $AR(1)$ estacionario, se puede despejar Y_t , utilizando el inverso del operador polinomial de retardos. Así,

$$Y_t = \phi_1^{-1}(L)a_t = \frac{1}{1 - \phi_1 L}a_t$$

Si se considera $|\phi_1 L| < 1$, la fracción del segundo miembro puede contemplarse como la suma de los términos de una progresión geométrica infinita de razón $\phi_1 L$, que es convergente. Es decir

$$Y_t = \sum_{j=0}^{\infty} (\phi_1 L)^j a_t = \sum_{j=0}^{\infty} \phi_1^j a_{t-j}$$

el cual es un modelo que se conoce como *proceso de medias móviles con infinitos retardos*, que será analizado más adelante.

Modelo AR(2) Un modelo $AR(2)$ viene definido por

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t \quad (25)$$

o, utilizando el operador de retardos,

$$(1 - \phi_1 L - \phi_2 L^2)Y_t = a_t \quad (26)$$

Para que el proceso anterior sea estacionario se requiere que las raíces de la ecuación

$$1 - \phi_1 L - \phi_2 L^2 = 0$$

caigan fuera del círculo unidad.

Si se multiplican ambos miembros de (25) por $Y_{t-\tau}$ y se toman esperanzas se tiene que

$$E(Y_t Y_{t-\tau}) = \phi_1 E(Y_{t-1} Y_{t-\tau}) + \phi_2 E(Y_{t-2} Y_{t-\tau}) + E(\epsilon_t Y_{t-\tau})$$

O mejor,

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + E(a_t Y_{t-\tau})$$

Para $\tau = 0$, resulta:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2$$

Para $\tau = 1$:

$$\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1$$

Para $\tau = 2$:

$$\gamma_2 = \phi_1 \gamma_1 + \phi_2 \gamma_0$$

Resolviendo algebraicamente las expresiones anteriores, se llega a que

$$\begin{aligned} \gamma_1 &= \frac{\phi_1 \gamma_0}{1 - \phi_2} \\ \gamma_2 &= \left(\frac{\phi_1^2 \gamma_0}{1 - \phi_2} \right) + \phi_2 \gamma_0 \end{aligned}$$

donde

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi_1^2 - \phi_2^2}$$

Se tiene que para $\tau > 0$, γ_τ se transforma en $\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2}$, de esta forma, las autocorrelaciones de este modelo quedan definidas como

$$\begin{aligned} \rho_\tau &= \frac{\gamma_\tau}{\gamma_0} \\ &= \frac{\phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2}}{\gamma_0}, \quad \text{por lo que} \\ &= \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2}, \quad \tau > 0 \end{aligned}$$

Modelo AR(p) Un modelo $AR(p)$ viene definido por

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t \quad (27)$$

o, alternativamente por

$$\phi(L)Y_t = a_t$$

done $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$. Para que el proceso sea estacionario, se requiere que las raices de la ecuación polinomial

$$1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p = 0$$

estén fuera del círculo unidad.

Multiplicando ambos miembros de (27) por $Y_{t-\tau}$ y tomando esperanzas se tiene que

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \cdots + \phi_p \gamma_{\tau-p} + E(a_t Y_{t-\tau})$$

Para $\tau = 0$, se obtiene

$$\gamma_0 = \phi_1 \gamma_1 + \cdots + \phi_p \gamma_p + \sigma_a^2$$

Para valores de $\tau > 0$ el resultado obtenido es el siguiente

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \cdots + \phi_p \gamma_{\tau-p}; \quad \tau > 0$$

Dividiendo ambos miembros de la ecuación anterior por γ_0 se obtiene la ecuación de orden p para las autocorrelaciones,

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \cdots + \phi_p \rho_{\tau-p}$$

Si se despeja $\phi(L)$ en $\phi(L)Y_t = a_t$ obtenemos

$$Y_t = \phi^{1-}(L)a_t = \frac{1}{\phi(L)}a_t$$

De esta forma se ha pasado de un modelo $AR(p)$ a un modelo $MA(\infty)$.

1.2.2. Modelos de medias móviles (MA)

Un modelo de *medias móviles* de orden q , o abreviadamente un $MA(q)$, se define como:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \quad (28)$$

Utilizando el operador polinomial de retardos,

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q$$

el modelo puede expresarse en forma compacta como

$$Y_t = \theta(L)a_t$$

A continuación se examinan las propiedades de un $MA(1)$ y de un $MA(2)$, para generalizarlas posteriormente a un $MA(q)$.

Modelo MA(1) Un modelo MA(1) viene definido por

$$Y_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 L)a_t$$

Si se multiplican ambos miembros de la ecuación anterior por $Y_{t-\tau}$, y se toman esperanzas se tiene que

$$\begin{aligned} E(Y_t Y_{t-\tau}) &= E(a_t - \theta_1 a_{t-1})(a_{t-\tau} - \theta_1 a_{t-\tau-1}) \\ \gamma_\tau &= E[a_t a_{t-\tau} - \theta_1 a_{t-\tau-1} a_t - \theta_1 a_{t-1} a_{t-\tau} + \theta_1^2 a_{t-1} a_{t-\tau-1}] \\ &= E[a_t a_{t-\tau}] - \theta_1 E[a_{t-\tau-1} a_t] - \theta_1 E[a_{t-1} a_{t-\tau}] + \theta_1^2 E[a_{t-1} a_{t-\tau-1}] \end{aligned}$$

Para $\tau = 0$, teniendo en cuenta que $E(a_t a_{t'}) = 0$ para $t \neq t'$, se obtiene

$$\gamma_0 = (1 + \theta_1^2) \sigma_a^2$$

Para $\tau = 1$

$$\gamma_1 = -\theta_1 \sigma_a^2$$

Para valores de $\tau > 1$, se deduce que

$$\gamma_\tau = 0, \quad \tau > 1.$$

De esta manera, dividiendo γ_τ entre γ_0 , se obtienen los coeficientes de autocorrelación

$$\begin{aligned} \rho_1 &= \frac{-\theta_1}{1 + \theta_1^2} \\ \rho_\tau &= 0, \quad \tau > 1 \end{aligned}$$

Dado que la varianza o autocovarianza de orden 0 (γ_0) es constante, se concluye que un modelo MA(1) es siempre estacionario independientemente del valor que tenga θ_1 .

Si expresamos el modelo así:

$$a_t = Y_t + \theta_1 a_{t-1}$$

Y si $|\theta_1| < 1$ puede hacerse

$$\begin{aligned} a_t &= \frac{1}{1 - \theta_1 L} Y_t = \sum_{j=0}^{\infty} (\theta_1 L)^j Y_t \\ &= Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots \end{aligned}$$

De esta forma, se ha pasado de un MA(1) a un AR(∞). A la condición que ha permitido pasar de un modelo a otro, es decir, que $|\theta_1| < 1$ se le denomina condición de invertibilidad. Como puede verse, la condición de invertibilidad de un modelo MA(1) es equivalente en sentido formal a la condición de estacionariedad de un modelo AR(1). Ahora bien, conviene no olvidar que un modelo MA(1) es siempre estacionario y que la condición de invertibilidad se establece para poder pasar a un modelo AR(∞).

Modelo MA(2) Un modelo MA(2) viene definido por

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = (1 - \theta_1 L - \theta_2 L^2) a_t$$

Si se multiplican ambos miembros de la ecuación anterior por $Y_{t-\tau}$, y se toman esperanzas se tiene que

$$E(Y_t Y_{t-\tau}) = E(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t-\tau} - \theta_1 a_{t-\tau-1} - \theta_2 a_{t-\tau-2})$$

Para distintos valores de τ se obtienen los siguientes resultados o autocovarianzas Si se multiplican ambos miembros de la ecuación anterior por $Y_{t-\tau}$, y se toman esperanzas se tiene que

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \theta_2^2) \sigma_a^2 \\ \gamma_1 &= (\theta_1 \theta_2 - \theta_1) \sigma_a^2 \\ \gamma_2 &= -\theta_2 \sigma_a^2 \\ \gamma_\tau &= 0; \quad \tau > 2\end{aligned}$$

A partir de las expresiones anteriores se obtienen fácilmente los coeficientes de autocorrelación

$$\begin{aligned}\rho_1 &= \frac{\theta_1 \theta_2 - \theta_1}{1 + \theta_1^2 + \theta_2^2} \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} \\ \rho_\tau &= 0; \quad \tau > 2.\end{aligned}$$

Para que un proceso MA(2) sea invertible se requiere que las raíces del polinomio característico

$$1 - \theta_1 L - \theta_2 L^2 = 0$$

caigan fuera del círculo unidad.

Modelo MA(q) Si en (28) se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas, se obtienen los siguientes resultados

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2 \\ \gamma_\tau &= \begin{cases} (-\theta_\tau + \theta_1 \theta_{\tau+1} + \dots + \theta_{q-\tau} \theta_q) \sigma_a^2 & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases}\end{aligned}$$

Por tanto, en las autocovarianzas existe un corte brusco en q , ya que después de ese retardo su valor es igual a 0. El mismo fenómeno se presenta con los coeficientes de autocorrelación que vendrán dados por

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1 \theta_{\tau+1} + \dots + \theta_{q-\tau} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases}$$

Para que un modelo MA(q) sea invertible se requieren que las raíces de la ecuación polinomial

$$1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q = 0$$

estén fuera del círculo unidad.

1.2.3. Modelos mixtos autorregresivos - medias móviles (ARMA)

Un modelo *autorregresivo y de medias móviles* ARMA(p, q) viene definido de la siguiente forma:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (29)$$

Utilizando los operadores polinomiales de retardo, el modelo queda expresado en forma compacta del siguiente modo:

$$\phi(L)Y_t = \theta(L)a_t$$

Para que el modelo sea estacionario se requiere que las raíces de la ecuación polinomial

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 0$$

caigan fuera del círculo unidad.

Bajo condiciones de estacionariedad, el modelo ARMA(p, q) se puede expresar como un MA(∞), pudiendo representarse de la siguiente forma:

$$Y_t = \frac{\theta(L)}{\phi(L)} a_t = \psi(L)a_t$$

Luego, los coeficientes del operador polinomial $\psi(L)$ cumplen la siguiente igualdad

$$\phi(L)\psi(L) \equiv \theta(L)$$

o, equivalentemente

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 + \psi_1 L + \psi_2 L^2 + \dots) = (1 - \theta_1 L - \dots - \theta_q L^q)$$

A partir de la identidad anterior, se pueden deducir un conjunto de ecuaciones que nos permiten obtener los ψ_i , en función de los coeficientes autorregresivos y de medias móviles, siendo un derivado principal que los valores de ψ_τ para $\tau > q$ se puede deducir la siguiente ecuación

$$\psi_\tau = \phi_1 \psi_{\tau-1} + \dots + \phi_p \psi_{\tau-p} \quad \tau > q$$

Por otra parte, para que el modelo (29) sea invertible, se requiere que las raíces de la ecuación polinomial

$$\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q = 0$$

se sitúen fuera del círculo unidad.

Bajo condiciones de invertibilidad, el modelo ARMA(p, q) se puede expresar mediante un AR(∞):

$$a_t = \frac{\phi(L)}{\theta(L)} Y_t = \pi(L) Y_t$$

donde el operador $\pi(L)$ se define como

$$\pi(L) = 1 - \pi_1 L - \pi_2 L^2 - \dots$$

los coeficientes del operador polinomial $\pi(L)$ deben cumplir la siguiente igualdad

$$\pi(L)\theta(L) = \phi(L)$$

Resolviendo la equivalencia anterior, se encuentra que para $\tau > p$ los coeficientes π_τ siguen la siguiente ecuación

$$\pi_\tau = \theta_1 \pi_{\tau-1} + \dots + \theta_q \pi_{\tau-q} \quad \tau > p$$

Modelo ARMA(1, 1) Un modelo ARMA(1, 1) viene dado por la expresión

$$Y_t = \phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1}$$

Multiplicando ambos miembros por $Y_{t-\tau}$ y tomando esperanzas tenemos que

$$\gamma_\tau = E(Y_t Y_{t-\tau}) = \phi_1 \gamma_{\tau-1} + E(a_t Y_{t-\tau}) - \theta_1 E(a_{t-1} Y_{t-\tau})$$

Teniendo en cuenta que $E(a_t Y_t) = \sigma_a^2$ y $E(a_{t-1} Y_t) = E[a_{t-1}(\phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1})] = (\phi_1 - \theta_1) \sigma_a^2$.

Para $\tau = 0$, se obtiene la expresión

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_a^2 - \theta_1 (\phi_1 - \theta_1) \sigma_a^2$$

Para $\tau = 1$, fácilmente se comprueba que

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_a^2$$

Para $\tau > 1$ resulta que

$$\gamma_\tau = \phi_1 \gamma_{\tau-1}$$

sustituyendo el valor de γ_1 en la ecuación para γ_0 y operando, se obtiene

$$\gamma_0 = \frac{1 - 2\theta_1\phi_1 + \theta_1^2}{1 - \phi_1^2} \sigma_a^2$$

por lo tanto,

$$\gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_a^2$$

De esta manera, se obtienen los siguientes coeficientes de autocorrelación

$$\begin{aligned} \rho_1 &= \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - 2\theta_1\phi_1 + \theta_1^2} \\ \rho_\tau &= \phi_1\rho_{\tau-1}; \quad \tau > 1 \end{aligned}$$

Modelo ARMA(p, q) Um modelo ARMA(p, q) viene definido como en (29). si en esta expresión se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas, se obtiene

$$\gamma_\tau - \phi_1\gamma_{\tau-1} - \cdots - \phi_p\gamma_{\tau-p} = E(a_t Y_{t-\tau}) - \theta_1 E(a_t Y_{t-\tau}) - \cdots - \theta_q E(a_{t-q} Y_{t-\tau})$$

Si se tiene en cuenta que $E(a_t Y_{t'}) = 0$ para $t' < t$, resulta que

$$\gamma_\tau - \phi_1\gamma_{\tau-1} - \cdots - \phi_p\gamma_{\tau-p} = 0, \quad \text{para } \tau > q$$

Para las autocorrelaciones se obtiene

$$\rho_\tau - \phi_1\rho_{\tau-1} - \cdots - \phi_p\rho_{\tau-p} = 0, \quad \text{para } \tau > q$$

En los modelos ARMA(p, q) es conveniente factorizar la parte autorregresiva y la parte de medias móviles, con objeto de analizar si existen raíces comunes. Si existiera alguna raíz idéntica en los dos miembros, entonces el modelo estaría sobreparametrizado innecesariamente, ya que un modelo con las mismas propiedades sería un ARMA($p - 1, q - 1$).

1.3. Procesos no estacionarios

Dado que la gran mayoría de las series económicas se caracterizan por que son generadas por procesos no estacionarios, se hace necesario incorporar herramientas que hagan posible la transformación de una serie no estacionaria en estacionaria

En la práctica, lo más común es que las series que se analizan sean no estacionarias, ya sea porque exhiben algún tipo de tendencia, porque su varianza no sea constante o porque estén influenciadas por algún factor de tipo semideterminista como puede ser la estacionalidad. Si el problema es que se aprecia una tendencia en el comportamiento de la serie, es bastante posible que dicha tendencia sea de carácter polinomial adaptativo y

que por tanto, pueda eliminarse mediante la aplicación del operador diferencia, lo cual da origen a los modelos ARIMA. Por otro lado, si la no estacionariedad se debe también a que la varianza no es constante, quizá la causa sea que en cada punto de observación t , la variable Y_t tiene varianza σ_t^2 la cual es función de su media μ_t ; de ocurrir esto, la transformación adecuada puede consistir en tomar logaritmos. Los estadísticos Box y Cox (1964) definieron una transformación instantánea. Esta transformación instantánea se define por

$$Y_t^\lambda = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln(Y_t) & \lambda = 0 \end{cases}$$

1.3.1. Modelos ARIMA

Los modelos *autorregresivos e integrados de medias móviles* (ARIMA) pueden ser vistos como una generalización de los modelos ARMA. Yaglon(1955) sugirió la posibilidad de que un cierto tipo de no estacionariedad mostrado por algunas series de tiempo, podía representarse mediante la simple toma sucesiva de diferencias de la serie original. esto, permite gran flexibilidad de representación a los modelos ARMA, puesto que en realidad lo que se hace al aplicar el operador diferencia ∇^d es eliminar una posible tendencia polinomial de orden d , presente en la serie que se analice. Es decir, si el proceso original Y_t presenta no estacionariedad causada por una tendencia polinomial no determinista es posible construir el proceso estacionario $\{W_t\}$, en donde

$$W_t = \nabla^d Y_t \quad \text{para toda } t \quad (30)$$

Para esta nueva serie es posible obtener un modelo ARMA: $\phi(L)W_t = \theta(L)a_t$, equivalente a considerar el modelo ARIMA

$$\phi(L)\nabla^d Y_t = \theta(L)a_t \quad d \geq 1 \quad (31)$$

El término "integrado" se refiere a que Y_t se obtiene de la relación (30) por inversión del operador ∇^d , dando como resultado una suma infinita (o una integración) de términos W_t . Si $W_t = \nabla Y_t$, aplicando el desarrollo en serie de McLaurin, se tendría

$$Y_t = \nabla^{-1}W_t = W_t + W_{t-1} + W_{t-2} + \dots$$

Un modelo ARIMA(p, d, q) indica que consta de un polinomio autorregresivo de orden p , de una diferencia de orden d y de un polinomio de medias móviles de orden q . De esta manera, el modelo (31) se escribe como

$$W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad \text{con } W_t = \nabla^d Y_t \quad (32)$$

Para que el modelo sea estacionario e invertible, se requiere que las raíces del polinomio $\phi(L) = 0$ y las raíces de $\theta(L) = 0$ se encuentren fuera del círculo unitario, pero si consideramos el operador autorregresivo generalizado

$$\varphi(L) = \phi(L)\nabla^d \quad (33)$$

la condición es que d de las raíces de $\varphi(L) = 0$ sean unitarias, mientras que las restantes deben estar fuera del círculo unitario. Con el uso de la ecuación (33), la expresión (32) se convierte en

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_{p+d} Y_{t-p-d} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

y la representación en términos de suma ponderada de choques aleatorios viene dada por

$$Y_t = \varphi^{-1}(L)\theta(L)a_t = \psi(L)a_t \quad (34)$$

en donde, debido a que

$$(1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_{p+d} L^{p+d})(1 - \psi_1 L - \psi_2 L^2 - \cdots) = (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q)$$

se deduce que las ponderaciones ψ deben satisfacer la ecuación en diferencia siguiente

$$\varphi(L)\psi_j = \phi(L)\nabla^d \psi_j = 0 \quad \text{para } j > \text{máx}\{p + d - 1, q\}$$

Asimismo, si el proceso definido en (34) es invertible, se tiene

$$\pi(L)Y_t = \theta^{-1}(L)\varphi(L)Y_t = a_t$$

con las ponderaciones π que satisfacen

$$\theta(L)\pi_j = 0 \quad \text{para } j > \text{máx}\{p + d, q\} \quad (35)$$

1.4. Pronósticos óptimos de series de tiempo

La teoría de pronósticos se presenta exclusivamente para el caso de series no estacionales, esto se debe a que una exposición que cubra el caso general del modelo multiplicativo para series estacionales no aporta mayor contribución al entendimiento de la teoría. Sin embargo, los resultados obtenidos para series no estacionales, son aplicables de manera prácticamente directa al caso de series estacionales cuyo modelo ARIMA sea del tipo multiplicativo.

1.4.1. Pronóstico de series estacionarias

Sea $\{Z_t\}$ una serie de tiempo, T una transformación y supongamos $W_t = \nabla^d T(Z_t)$ es estacionaria con media cero. Supóngase además que W_t admite la representación $W_t = \psi(L)a_t$, para la cual existe un modelo ARMA equivalente, que se va a utilizar en la obtención de pronósticos de la serie. En particular si, a partir del origen t , se desea pronosticar la observación W_{t+h} , un pronóstico cualquiera de esta observación, que se obtenga como combinación lineal de los valores de la serie $\{W_t\}$, y en consecuencia de los errores $\{a_t\}$, será denotado por $\tilde{W}_t(h)$, mientras que el **pronóstico óptimo** se escribirá como $\hat{W}_t(h)$.

Definición 1.2. *Se denomina error medio cuadrático mínimo a*

$$E_t[W_{t+h} - \hat{W}_t(h)]^2 = \min_{\tilde{W}_t(h)} E_t[W_{t+h} - \tilde{W}_t(h)]^2 \quad (36)$$

en la cual E_t denota la esperanza condicional, dada toda la información hasta el momento t , o sea

$$E_t[W_{t+h} - \tilde{W}_t(h)]^2 = E\{[W_{t+h} - \tilde{W}_t(h)]^2 | Z_t, Z_{t-1}, \dots\} \quad (37)$$

Ya que $\tilde{W}_t(h)$, y por consiguiente también $\hat{W}_t(h)$, debe ser de la forma

$$\tilde{W}_t(h) = C_h a_t + C_{h+1} a_{t-1} + \dots = \sum_{j=h}^{\infty} C_j a_{t+h-j} \quad (38)$$

el problema de obtener $\tilde{W}_t(h)$ se traduce en la determinación de los valores C_h, C_{h+1}, \dots , de tal manera que se cumpla la relación 36. Con este fin, la observación W_{t+h} se escribirá como

$$W_{t+h} = - \sum_{j=0}^{\infty} \psi_j a_{t+h-j} \quad (39a)$$

$$= - \sum_{j=0}^{h-1} \psi_j a_{t+h-j} - \sum_{j=h}^{\infty} \psi_j a_{t+h-j}, \quad \text{con } \psi_0 \equiv -1 \quad (39b)$$

en donde la primera suma corresponde a información desconocida al tiempo t (ya que abarca observaciones desde $t+1$ hasta $t+h$), mientras que la segunda suma consta de información conocida al tiempo t (desde $-\infty$ hasta t). De (39) y (38) se obtiene

$$W_{t+h} - \tilde{W}_t(h) = - \sum_{j=0}^{h-1} \psi_j a_{t+h-j} - \sum_{j=h}^{\infty} (\psi_j + C_j) a_{t+h-j} \quad (40)$$

y, debido a que $E(a_{t+i} a_{t+j}) = 0$ para $i \neq j$, se sigue que

$$E_t[W_{t+h} - \tilde{W}_t(h)]^2 = \sum_{j=0}^{h-1} \psi_j^2 \sigma_a^2 + \sum_{j=h}^{\infty} (\psi_j + C_j)^2 \sigma_a^2 \quad (41)$$

el mínimo de esta expresión se obtiene haciendo $C_j = -\psi_j$, para $j = h, h + 1, \dots$, y esta selección de valores C_j conduce a tener

$$E_t[W_{t+h} - \hat{W}_t(h)]^2 = \sum_{j=0}^{h-1} \psi_j^2 \sigma_a^2 \quad (42)$$

con $\hat{W}_t(h)$ dado, según (38), por

$$\hat{W}_t(h) = -\psi_h a_t - \psi_{h+1} a_{t-1} - \dots = -\sum_{j=h}^{\infty} \psi_j a_{t+h-j} \quad (43)$$

Ahora bien, por definición de la esperanza condicional, se tiene

$$E_t(a_{t+h-j}) = \begin{cases} a_{t+h-j}, & \text{si } j \geq h \\ 0, & \text{si } j < h \end{cases} \quad (44)$$

de donde se sigue que

$$E_t(W_{t+h}) = -E_t\left(\sum_{j=0}^{h-1} \psi_j a_{t+h-j}\right) - E_t\left(\sum_{j=h}^{\infty} \psi_j a_{t+h-j}\right) \quad (45a)$$

$$= -\sum_{j=h}^{\infty} \psi_j a_{t+h-j} \quad (45b)$$

y, en consecuencia, al comparar esta última expresión con (43), se advierte que $E_t(W_{t+h})$ proporciona el pronóstico con error cuadrático medio mínimo, es decir

$$\tilde{W}_t(h) = E_t(W_{t+h}) \quad (46)$$

el error de pronóstico (46) con origen en t , viene dado por

$$e_t(h) = W_{t+h} - \hat{W}_t(h) \quad (47a)$$

$$= -\sum_{j=0}^{h-1} \psi_j a_{t+h-j} \quad (47b)$$

por lo cual es inmediato que

$$E_t[e_t(h)] = 0 \quad \text{y} \quad \text{Var}_t[e_t(h)] = \sum_{j=0}^{h-1} \psi_j^2 \sigma_a^2 \quad (48)$$

con $\text{Var}_t[e_t(h)] = \text{Var}[e_t(h)|Z_t, Z_{t-1}, \dots]$; aquí se aprecia en particular que los pronósticos $\hat{W}_t(h)$ son **insesgados**. Además, se tiene

$$\text{Var}_t[e_t(h)] - \text{Var}_t[e_t(h-1)] = \psi_{h-1}^2 \sigma_a^2 \geq 0 \quad \text{para } h \geq 1 \quad (49)$$

de donde se deduce que al emplear los pronósticos óptimos, mientras mas alejado se desee el pronóstico (mayor sea h) mayor será la varianza (menor la precisión) del mismo.

Si se toma $h = 1$ en (47) se obtiene

$$e_t(1) = W_{t+1} - \hat{W}_t(1) = a_{t+1} \quad (50)$$

lo cual es válido para $t = d, d+1, \dots, N$, y que por tanto muestra que los errores del pronóstico un período hacia adelante no están correlacionados. En general, los errores de pronóstico $e_d(h), e_{d+1}(h), \dots, e_N(h)$, se encuentran correlacionadas para $h \geq 2$ ya que, como lo muestra la expresión (47), la serie $\{e_t(h)\}$ con h fijo, sigue un proceso MA($h-1$). Por otro lado, si el origen permanece fijo, los errores de pronóstico para diversos períodos hacia adelante, $e_t(1), e_t(2), \dots, e_t(H)$ con $H \geq 2$, también están correlacionados.

Ya que se desea obtener pronósticos de la serie $\{W_t\}$, considérese conocido el modelo $\phi(L)W_t = \theta(L)a_t$, entonces

$$\tilde{W}_t(h) = E_t(W_{t+h}) \quad (51a)$$

$$= E_t(\phi_1 W_{t+h-1} + \dots + \phi_p W_{t+h-p} + a_{t+h} - \theta_1 a_{t+h-1} - \dots - \theta_q a_{t+h-q}) \quad (51b)$$

$$= \phi_1 E_t(W_{t+h-1}) + \dots + \phi_p E_t(W_{t+h-p}) + E_t(a_{t+h}) - \theta_1 E_t(a_{t+h-1}) - \dots - \theta_q E_t(a_{t+h-q}) \quad (51c)$$

donde

$$E_t(W_{t+h-j}) = \begin{cases} W_{t+h-j}, & \text{si } j \geq h \\ \tilde{W}_t(h-j), & \text{si } j < h \end{cases} \quad (52)$$

y, empleando (44) y (50), también se obtiene

$$E_t(a_{t+h-j}) = \begin{cases} W_{t+h-j} - \hat{W}_{t+h-j-1}(1), & \text{si } j \geq h \\ 0, & \text{si } j < h \end{cases} \quad (53)$$

esta expresión es importante y necesaria ya que, en contraste con (44), en donde aparecen los errores aleatorios a_{t+h-j} no observables, aquí aparecen los valores de la serie y sus pronósticos, que sí son observables.

Ejemplo 1.1. *Considerese el proceso W_t definido por*

$$(1 - 0,6L)W_t = (1 + 0,2L)a_t$$

entonces

$$W_t = 0,6W_{t-1} + a_t + 0,2a_{t-1}$$

de tal forma que los pronósticos $\hat{W}_t(h)$ se obtienen como

$$\begin{aligned}
 \hat{W}_t(1) &= E_t(W_{t+1}) \\
 &= 0,6E_t(W_t) + E_t(a_{t+1}) + 0,2E_t(a_t) \\
 &= 0,6W_t + 0,2[W_t - \hat{W}_{t-1}(1)] \\
 \hat{W}_t(2) &= 0,6E_t(W_{t+1}) + E_t(a_{t+2}) + 0,2E_t(a_{t+1}) \\
 &= 0,6\hat{W}_t(1)
 \end{aligned}$$

y, en general

$$\hat{W}_t(h) = 0,6\hat{W}_t(h-1) \quad \text{para } h \geq 2$$

es decir, los pronósticos se pueden obtener de manera recursiva.

Se debe tener en cuenta que $\hat{W}_t(1)$ involucra al pronóstico $\hat{W}_{t-1}(1)$, el cual a su vez utiliza $\hat{W}_{t-2}(1)$ ya que

$$\hat{W}_{t-1}(1) = 0,6W_{t-1} + 0,2[W_{t-1} - \hat{W}_{t-2}(1)]$$

procediendo recursivamente se tiene

$$\hat{W}_1(1) = 0,6W_1 + 0,2[W_1 - \hat{W}_0(1)]$$

en donde, ya que no existe información para calcular $\hat{W}_0(1)$, se supone $\hat{W}_0(1) = W_1$, es decir, suponemos $a_1 = 0$; así pues, para obtener los pronósticos $\hat{W}_t(h)$ con $h \geq 2$, hay necesidad de calcular los valores

$$\begin{aligned}
 \hat{W}_1(1) &= 0,6W_1 \\
 \hat{W}_2(1) &= 0,8W_2 - 0,2\hat{W}_1(1) \\
 &\vdots \\
 \hat{W}_t(1) &= 0,8W_t - 0,2\hat{W}_{t-1}(1)
 \end{aligned}$$

2. Pronósticos con restricciones

Los modelos ARIMA pueden mejorarse mediante diversos métodos: modelos de la familia ARCH, modelos VAR, análisis de los efectos de intervención. En esta sección se hace uso de uno de estos modelos para clarificar la metodología de pronósticos con restricciones.

Se trata entonces de mejorar la calidad de los pronósticos con modelos ARIMA mediante la inclusión de información adicional referente al comportamiento futuro de la serie; dicha información está dada en forma de restricciones lineales y para su inclusión en el pronóstico se siguen procedimientos formales de la estadística.

Se debe tener en cuenta que no toda información adicional se debe anexar al pronóstico, ya que conviene verificar la compatibilidad de ésta información con los datos históricos de la variable. Por lo anterior se examinan las siguientes posibilidades:

2.1. Restricciones ciertas, compatibles con la historia

Para seguir la metodología propuesta por Guerrero[?] sobre la incorporación de la restricción lineal en los pronósticos del modelo ARIMA, consideremos:

$\{a_t\}$ una serie ruido blanco y $\{Z_t\}$ una serie de tiempo con $t = 1, 2, \dots, T$ donde $Z_t \sim \text{ARIMA}(p, d, q)$

$$\phi(L)\nabla^d Z_t = \theta(L)a_t$$

con

$$\begin{aligned}\phi(L) &= 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \\ \theta(L) &= 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q\end{aligned}$$

La serie puede escribirse así

$$\nabla^d Z_t = \frac{\theta(L)}{\phi(L)} a_t = \Psi(L) a_t \quad \text{con} \quad \Psi(L) = \frac{\theta(L)}{\phi(L)}$$

donde $\Psi(L)$ es un polinomio con la siguiente representación

$$\begin{aligned}\Psi(L) &= 1 - \psi_1 L - \psi_2 L^2 - \psi_3 L^3 - \dots \\ \Psi(L) &= \psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 - \dots\end{aligned}$$

Consideremos los vectores

$$\begin{aligned}Z_0 &= (Z_1, Z_2, \dots, Z_T)' \\ Z_F &= (Z_{T+1}, Z_{T+2}, \dots, Z_{T+H})' \\ a_F &= (a_{T+1}, a_{T+2}, \dots, a_{T+H})'\end{aligned}$$

Donde $\{Z_0\}$ recoge la información histórica de la variable aleatoria y $\{Z_F\}$ recoge los valores futuros de la serie Z_{T+H} para $H \geq 1$ períodos por delante y $\{a_F\}$ recoge la información futura de la variable ruido blanco y consideremos la matriz

$$\Psi = \begin{pmatrix} -\psi_0 & 0 & 0 & \cdots & 0 \\ -\psi_1 & -\psi_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\psi_{H-1} & -\psi_{H-2} & -\psi_{H-3} & \cdots & -\psi_0 \end{pmatrix}$$

Llamada *matriz de ponderaciones*, que recoge los valores de ψ para el primer pronóstico, hasta $H - 1$.

Definición 2.1. *El pronóstico óptimo de la serie Z_t con error cuadrático medio (ECM) mínimo, para un horizonte de h períodos, es la esperanza condicional de la variable dada la información histórica contenida en el vector Z_0 .*

$$\hat{Z}_t(h) = E(Z_{t+h} | Z_0)$$

Definición 2.2. *El error de pronóstico óptimo (con ECM mínimo) se define como la diferencia entre los valores futuros y la esperanza condicional de la variable, dado el vector Z_0 .*

Proposición 2.1. *El error de pronóstico óptimo (con ECM mínimo) admite la siguiente representación:*

$$Z_{T+h} - E(Z_{T+h} | Z_0) = - \sum_{k=0}^{h-1} \psi_j a_{T+h-k} \quad \text{para } h = 1, 2, \dots \quad (54)$$

Demostración. Supongamos que la serie $\{Z_t\}$ es no estacionaria, entonces puede considerarse generada por la serie estacionaria $\{W_t\}$ aplicando diferencias y una transformación T , de la forma $W_t = \nabla^d T(Z_t)$. Luego, admite la representación

$$\delta(L)Z_t = W_t \quad \text{con} \quad \delta(L) = 1 - \delta_1 L - \cdots - \delta_d L^d \quad \text{y} \quad d \geq 1$$

dond el vector de valores iniciales $Z_* = (Z_1, \dots, Z_d)'$ está fijo, de tal forma que, dados $\{W_t\}$ y Z_* , los valores subsecuentes de la serie se generan recursivamente a partir de

$$Z_t = \delta_1 Z_{t-1} + \cdots + \delta_d Z_{t-d} + W_t \quad \text{para } t > d$$

Por tanto, para $t > d$ se tiene la representación

$$Z_t = \underline{A}'_t Z_* + \sum_{i=0}^{t-d-1} \zeta_i W_{t-i}$$

con ζ_0, ζ_1, \dots los coeficientes del polinomio $\zeta(L) = \delta^{-1}(L)$, dados por

$$\zeta_0 = 1 \quad \text{y} \quad \zeta_i = \sum_{k=1}^{\min(d,i)} \delta_k W_{i-k} \quad \text{para} \quad i \geq 1$$

Por ejemplo, si $\delta(L) = 1-L$, entonces $d = 1, \delta_1 = 1$ y $\zeta_i = 1$ para $i \geq 0$, o si $\delta(L) = (1-L)^2$, entonces $d = 1, \delta_1 = 2, \delta_2 = -1$ y $\zeta_i = i + 1$ para $i \geq 0$. Además $\underline{A}'_t = (A'_{1t}, \dots, A'_{dt})$ está definido por $A'_{dt} = \zeta_{t-d}$ y

$$A'_{it} = \zeta_{t-i} - \sum_{k=0}^{d-i} \zeta_{t-k-i} \delta_k \quad \text{para} \quad i = 1, \dots, d-1.$$

Así se obtiene la representación

$$Z_t = \underline{A}'_t Z_* + \sum_{i=0}^{t-d-1} \zeta_i \psi^*(L) a_{t-i}$$

donde $\psi^*(L)$ es el polinomio de retraso asociado con la representación MA pura de la serie $\{W_t\}$, o sea $W_t = \psi^*(L) a_t$ con $\sum_{j=0}^{\infty} |\psi_j^*| < \infty$. De esta manera se tiene que

$$Z_{T+h} = \underline{A}'_{T+h} Z_* - \sum_{i=0}^{T+h-d-1} \zeta_i \sum_{j=0}^{\infty} \psi_j^* a_{T+h-j-i}$$

y

$$\begin{aligned} E(Z_{T+h} | Z_0) &= \underline{A}'_{T+h} Z_* - \sum_{i=0}^{T+h-d-1} \zeta_i E \left(\sum_{j=0}^{\infty} \psi_j^* a_{T+h-j-i} \middle| Z_0 \right) \\ &= \underline{A}'_{T+h} Z_* - \sum_{i=0}^{T+h-d-1} \zeta_i \sum_{j=h-i}^{\infty} \psi_j^* a_{T+h-j-i} \end{aligned}$$

Luego

$$Z_{T+h} - E(Z_{T+h}|Z_0) = - \sum_{i=0}^{T+h-d-1} \zeta_i \sum_{j=0}^{\infty} \psi_j^* a_{T+h-j-i} + \sum_{i=0}^{T+h-d-1} \zeta_i \sum_{j=h-i}^{\infty} \psi_j^* a_{T+h-j-i} \quad (55a)$$

$$= - \sum_{i=0}^{T+h-d-1} \zeta_i \sum_{j=0}^{h-i-1} \psi_j^* a_{T+h-j-i} \quad (55b)$$

$$= - \sum_{i=0}^{T+h-d-1} \zeta_i \sum_{k=i}^{h-1} \psi_{k-i}^* a_{T+h-k} \quad (55c)$$

$$= - \sum_{i=0}^{h-1} \zeta_i \sum_{k=i}^{h-1} \psi_{k-i}^* a_{T+h-k} \quad (55d)$$

$$= - \sum_{i=0}^{h-1} \sum_{k=i}^{h-1} \zeta_i \psi_{k-i}^* a_{T+h-k} \quad (55e)$$

$$= - \sum_{k=0}^{h-1} \psi_k a_{T+h-k} \quad (55f)$$

donde la cuarta igualdad se cumple debido a que

$$\sum_{h=i}^{h-1} \psi_{k-i} a_{T+h-k} = 0 \quad \text{si } i > h - 1.$$

□

Si consideramos $h = 1, \dots, H$ la ecuacion (54) se puede escribir en forma matricial como

$$Z_F - E(Z_F|Z_0) = \Psi a_F \quad (56)$$

El vector a_F es tal que

$$\begin{aligned} E(a_F|Z_0) &= 0 \quad \text{y} \\ \Sigma_{a_F} &= E[(a_F - E(a_F))(a_F - E(a_F))'|Z_0] \\ &= E(a_F a_F'|Z_0) \\ &= \sigma_a^2 I \end{aligned}$$

donde σ_a^2 es la varianza de $\{a_t\}$.

Al usar (56), se puede considerar que la matriz de varianza-covarianza del error de pronóstico anterior es:

$$\begin{aligned} Cov(Z_F - E(Z_F|Z_0)) &= Cov(\Psi a_F) \\ &= E[\Psi a_F a_F' \Psi'] \\ &= \sigma_a^2 \Psi \Psi' \end{aligned}$$

Sea m el número de restricciones y Y un vector columna que contiene la información adicional e impone las restricciones sobre los valores futuros de la serie. Asumiendo que hay compatibilidad entre los valores históricos de la serie y la restricción, pueden relacionarse con la especificación de una matriz C de dimensión $m \times H$ con $m < H$, de rango completo y cuyos valores son constantes, de la siguiente manera:

$$Y = CZ_F \quad (57)$$

Si por ejemplo $m = 1$, se tendría:

$$Y = (C_1, C_2, \dots, C_H) \begin{pmatrix} Z_{T+1} \\ Z_{T+2} \\ \vdots \\ Z_{T+H} \end{pmatrix} = C_1 Z_{T+1} + C_2 Z_{T+2} + \dots + C_H Z_{T+H}$$

Proposición 2.2. *El pronóstico óptimo (lineal y con ECM mínimo) de la serie que incorpora la restricción establecida en Y , está dado por*

$$\hat{Z}_F = E(Z_F|Z_0) + \hat{A}[Y - CE(Z_F|Z_0)] \quad (58)$$

con

$$\hat{A} = \Psi\Psi' C'(C\Psi\Psi')^{-1} \quad (59)$$

donde, la matriz de varianza covarianza para el error de pronóstico es

$$\Sigma_{\hat{Z}_F - Z_F} = \text{Cov}(\hat{Z}_F - Z_F) = \sigma_a^2 \Psi\Psi'(I - \hat{A}C)' \quad (60)$$

$$= \text{Cov}[E(\hat{Z}_F|Z_0) - Z_F] - \sigma_a^2 \Psi\Psi' C' \hat{A}' \quad (61)$$

Demostración. Para demostrar los resultados (58) y (59) consideremos un pronóstico lineal general que incluya tanto a Y como a $E(Z_F|Z_0)$ de la forma

$$\tilde{Z}_F = AY + A_*E(Z_F|Z_0) \quad (62)$$

donde las matrices A y A_* son de constantes desconocidas y \tilde{Z}_F es un pronóstico cualquiera en el sentido de que no satisface la condición de optimalidad del error cuadrático medio mínimo. Reemplazando la ecuación (56) en la ecuación (57) resulta que

$$Y = CZ_F = C[E(Z_F|Z_0) + \Psi a_F] \quad (63)$$

Reemplazando (63) en (62) encontramos

$$\tilde{Z}_F = AC[E(Z_F|Z_0) + \Psi a_F] + A_*E(Z_F|Z_0) \quad (64)$$

$$= (AC + A_*)E(Z_F|Z_0) + AC\Psi a_F \quad (65)$$

Como es necesario que el pronóstico sea insesgado (dado Z_0 y Y), se debe tener:

$$\begin{aligned}
E(\tilde{Z}_F|Z_0, Y) &= E(Z_F|Z_0, Y), \quad \text{es decir,} \\
0 &= E(\tilde{Z}_F - Z_F|Z_0, Y) \\
&= E[(AC + A_*)E(Z_F|Z_0) + AC\Psi a_F - E(Z_F|Z_0) - \Psi a_F|Z_0, Y] \\
&= (AC + A_* - I)E(Z_F|Z_0) + (AC - I)\Psi E(a_F|Z_0, Y) \\
&= (AC + A_* - I)E(Z_F|Z_0)
\end{aligned}$$

Luego,

$$(AC + A_*)E(Z_F|Z_0) = E(Z_F|Z_0) \quad (66)$$

De otro lado,

$$\tilde{Z}_F - Z_F = (AC + A_*)E(Z_F|Z_0) + AC\Psi a_F - E(Z_F|Z_0) - \Psi a_F \quad (67)$$

Reemplazando (66) en (67) tenemos

$$\tilde{Z}_F - Z_F = (AC - I)\Psi a_F \quad (68)$$

Sea $\zeta = \tilde{Z}_F - Z_F$, la covarianza de este vector aleatorio es:

$$\begin{aligned}
\Sigma_\zeta &= E[(\zeta - E(\zeta))(\zeta - E(\zeta))'] \quad \text{como } E(\zeta) = 0, \quad \text{tenemos} \\
&= E[(AC - I)\Psi a_F a_F' \Psi' (AC - I)'] \\
&= [(AC - I)\Psi] E(a_F a_F') [(AC - I)\Psi]' \\
&= (AC - I)\Psi (\sigma_a^2 I) \Psi' (A' C' - I) \\
&= \sigma_a^2 (AC\Psi - \Psi)(\Psi' C' A' - \Psi')
\end{aligned}$$

resultando,

$$Cov(\tilde{Z}_F - Z_F) = \sigma_a^2 (AC\Psi\Psi' C' A' - AC\Psi\Psi' - \Psi\Psi' C' A' + \Psi\Psi') \quad (69)$$

Ahora, para que \tilde{Z}_F tenga ECM mínimo, se debe elegir A de tal manera que se obtenga el mínimo de la varianza generalizada, definida como la traza de $Cov(\tilde{Z}_F - Z_F)$, es decir:

$$\text{var}(\zeta) = \text{tr}[Cov(\zeta)] \quad (70)$$

Se requiere resolver el siguiente problema de optimización:

$$\min_A [\text{Var}(\zeta)]$$

la siguiente ecuación es la condición de primer orden para optimizar y permite elegir el valor apropiado de A .

$$\left. \frac{\partial \text{var}(\zeta)}{\partial A} \right|_{\hat{A}} = 0 \quad (71)$$

Entonces,

$$\frac{\partial \text{var}(\zeta)}{\partial A} = \sigma_a^2(2AC\Psi\Psi'C' - 2\Psi\Psi'C') = 0 \quad (72)$$

Sea \hat{A} el valor que optimiza, de acuerdo con (71). Entonces,

$$\hat{A}C\Psi\Psi'C' = \Psi\Psi'C'$$

la matriz $C\Psi\Psi'C'$ es no singular, por lo que se obtiene

$$\hat{A} = \Psi\Psi'C'(C\Psi\Psi'C')^{-1} \quad (73)$$

Reemplazando (66) en (64) resulta

$$\hat{Z}_F = E(Z_F|Z_0) + \hat{A}C\Psi a_F \quad (74)$$

Nuevamente al utilizar el hecho de que $Y - CE(Z_F|Z_0) = C\Psi a_F$, se demuestra la expresión (58). Para obtener la expresión que representa la Covarianza de $\widehat{Z}_F - Z_F$, sustituimos \hat{A} por A en: $\sigma_a^2(AC\Psi - \Psi)(\Psi'C'A' - \Psi')$, resultando:

$$\begin{aligned} \text{Cov}(\widehat{Z}_F - Z_F) &= (\hat{A}C - I)\Psi\sigma_a^2 I\Psi'(C'\hat{A}' - I) \\ &= \sigma_a^2[\Psi\Psi'C'(C\Psi\Psi'C')^{-1}C - I]\Psi\Psi'[C'(C\Psi\Psi'C')^{-1}C\Psi\Psi' - I] \\ &= \sigma_a^2[\Psi\Psi'C'(C\Psi\Psi'C')^{-1}C\Psi\Psi' - \Psi\Psi'] [C'(C\Psi\Psi'C')^{-1}C\Psi\Psi' - I] \\ &= \sigma_a^2[\Psi\Psi' - \Psi\Psi'C'(C\Psi\Psi'C')^{-1}C\Psi\Psi'] \\ &= \sigma_a^2\Psi\Psi'[I - C'(C\Psi\Psi'C')^{-1}C\Psi\Psi'] \\ &= \sigma_a^2\Psi\Psi'[I - C'\hat{A}'] \end{aligned}$$

Como $\text{Cov}(\Psi a_F|Z_0) = \sigma_a^2\Psi\Psi'$ y $E(\widetilde{Z}_F - Z_0|Z_F) - Z_F = \Psi a_F$, se tiene que:

$$\text{Cov}[E(Z_F|Z_0) - Z_F] = \sigma_a^2\Psi\Psi',$$

□

Dado que la matriz $\sigma_a^2\Psi\Psi'C'\hat{A}'$ es semidefinida positiva, el error de pronóstico restringido tiene una covarianza igual o menor que la del pronóstico irrestricto.

Por último, se hace necesario validar la compatibilidad entre la información adicional (Y) y la historia de la serie (Z_0); como se sabe, La información adicional se puede representar como combinación lineal de los valores futuros de la serie. Ahora, si realmente hay compatibilidad, esta igualdad debe mantenerse aun si se tiene en cuenta la historia de la variable; para representar esta relación consideraremos la esperanza condicional, con se contempla la siguiente hipótesis

$$E(Y|Z_0) = CE(Z_F|Z_0) \quad (75)$$

Asumiendo que $a_F \sim N_H(0, \sigma_a^2 I)$. Bajo el supuesto de veracidad de la hipótesis nula, se tendría que el error entre lo observado de Y y lo esperado $E(Y|Z_0)$, vendría a ser

$$e_Y = Y - E(Y|Z_0) = Y - CE(Z_F|Z_0) = C\Psi a_F \sim N_m(0, \sigma_a^2 C\Psi\Psi' C^t) \quad (76)$$

Para construir un estadístico que permita la validación de la anterior hipótesis, consideraremos la siguiente proposición

Proposición 2.3. *Sea Z un vector de V.A con distribución normal, de media cero y matriz de varianza covarianza Σ , la cual es no singular y de tamaño $m \times m$. Entonces*

$$Z'(\Sigma)^{-1}Z \sim \chi_m^2 \quad (77)$$

Aplicando el resultado dado por (77) resulta,

$$\frac{1}{\sigma_a^2} e_Y' (C\Psi\Psi' C^t)^{-1} e_Y \sim \chi_m^2$$

Como se requiere estimar σ_a^2 se puede utilizar la media cuadrática residual, y por tanto

$$\hat{\sigma}_a^2 = \frac{a_0' a_0}{T - k} \quad (78)$$

Donde $a_0 = (a_1, a_2, \dots, a_T)'$ y k es el número de parámetros en el modelo ARIMA; se puede hacer una aproximación de un estadístico para un tamaño de muestra grande

$$K_1 = \frac{e_Y^t (C\Psi\Psi^t C^t)^{-1} e_Y}{\sigma_a^2} \sim \chi_m^2 \quad (79)$$

2.2. Restricciones inciertas, compatibles o no con la historia

Consideremos nuevamente la restricción Y , pero con alguna incertidumbre en la información adicional. Esta será almacenada en un vector aleatorio u que suponemos distribuye normal e independiente de los errores del modelo ARIMA.

$$Y - CZ_F = u \quad \text{con} \quad u \sim N_m(\mathbf{0}, U) \quad (80)$$

Con C como se definió en (57). Se observa que:

$$i) \text{Cov}[Z_0, u'] = 0$$

Demostración. u debe estar incorrelacionado tanto con la historia (Z_0) como con los errores de pronóstico (a_F), luego,

$$\begin{aligned} \text{Cov}(Z_0, u') &= E[Z_0 - E(Z_0)][u' - E(u')] = E[Z_0 u'] \\ &= Z_0 E[u'] \\ &= 0 \end{aligned}$$

puesto que u tiene distribución normal con media 0. □

$$ii) \text{Cov}[a_F, u'] = 0$$

Demostración. Tenemos que $\text{Cov}(a_F, u') = E[(a_F - E(a_F))(u - E(u))'] = E(a_F u')$.

$$E(a_F u') = E(a_F)E(u') = 0$$

□

Haciendo uso de (80), resulta que

$$\begin{aligned} Y - E(Y|Z_0) &= (CZ_F + u) - E[(CZ_F + u)|Z_0] \\ &= (CZ_F + u) - CE(Z_F|Z_0) + E(u|Z_0) \\ &= (CZ_F + u) - CE(Z_F|Z_0) \end{aligned}$$

Pero, por (56) tenemos que $Z_F - E(Z_F|Z_0) = \Psi a_F$. y por tanto $C[Z_F - E(Z_F|Z_0)] + u = C\Psi a_F + u$, luego

$$Y - E(Y|Z_0) = C\Psi a_F + u \quad (81)$$

Por consiguiente, debido a (81) y (56), se puede escribir

$$E\{[Z_F - E(Z_F|Z_0)]|Y - E(Y|Z_0)\} = E(\Psi a_F | C\Psi a_F + u) \quad (82)$$

Bajo el supuesto de normalidad para a_F el esperado debe ser lineal, por lo que:

$$E(\Psi a_F | C\Psi a_F + u) = A_U(C\Psi a_F + u) \quad (83)$$

donde A_U es de tamaño $H \times m$, formada por constantes conocidas y que debe satisfacer la siguiente condición de ortogonalidad, para que la suma de cuadrados de la diferencia entre los vectores Ψa_F y $A_U(C\Psi a_F + u)$ sea mínima

$$E\{(C\Psi a_F + u)[\Psi a_F - \hat{A}_U(C\Psi a_F + u)]\} = 0 \quad (84)$$

operando resulta,

$$\begin{aligned} 0 &= E[C\Psi a_F a_F' \Psi' + u a_F' \Psi' - (C\Psi a_F a_F' \Psi^t C^t + C\Psi a_F u' + u a_F' \Psi' C' + u u') \hat{A}_U'] \\ &= \sigma_a^2 C\Psi \Psi' - (\sigma_a^2 C\Psi \Psi' C' + U) \hat{A}_U' \end{aligned}$$

luego

$$\hat{A}_U = \Psi \Psi' C' (C\Psi \Psi' C' + U/\sigma_a^2)^{-1} \quad (85)$$

Esta matriz minimiza el error, por lo que al sustituir \hat{A}_U en (83) resulta

$$E\{[Z_F - E(Z_F|Z_0)]|Y - E(Y|Z_0)\} = \hat{A}_U(C\Psi a_F + u) = \hat{A}_U[Y - E(Y|Z_0)] \quad (86)$$

Y por tanto

$$\hat{Z}_{F,U} = E(Z_F|Z_0, Y) = E(Z_F|Z_0) + \hat{A}_U[Y - E(Y|Z_0)] \quad (87)$$

Para calcular el error de pronóstico, encontramos

$$\begin{aligned} \hat{Z}_{F,U} - Z_F &= E(Z_F|Z_0) + \hat{A}_U[Y - E(Y|Z_0)] - [E(Z_F|Z_0) + \Psi a_F] \\ &= \hat{A}_U[Y - E(Y|Z_0)] - \Psi a_F \\ &= \hat{A}_U[C\Psi a_F + u] - \Psi a_F \end{aligned}$$

Y por tanto el error es

$$\hat{Z}_{F,U} - Z_F = (\hat{A}_U C - I)\Psi a_F + \hat{A}_U u \quad (88)$$

y así, haciendo $\zeta_{F,U} = \hat{Z}_{F,U} - Z_F$, resulta:

$$\begin{aligned} \text{Cov}(\zeta_{F,U}) &= E[(\zeta_{F,U} - E(\zeta_{F,U}))(\zeta_{F,U} - E(\zeta_{F,U}))'] \\ &= E\left\{[(\hat{A}_U C - I)\Psi a_F + \hat{A}_U U][(\hat{A}_U C - I)\Psi a_F + \hat{A}_U U]'\right\} \\ &= \sigma_a^2(\hat{A}_U C - I)\Psi\Psi'(C'\hat{A}_U' - I) + \hat{A}_U U \hat{A}_U' \\ &= \sigma_a^2(\hat{A}_U C\Psi\Psi'C'\hat{A}_U' - 2\Psi\Psi'C'\hat{A}_U' + \Psi\Psi^t) + \hat{A}_U U \hat{A}_U' \\ &= \sigma_a^2\Psi\Psi'(I - 2C'\hat{A}_U') + \sigma_a^2\hat{A}_U(C\Psi\Psi'C' + U/\sigma_a^2)\hat{A}_U' \\ &= \sigma_a^2\Psi\Psi'(I - 2C'(C\Psi\Psi'C' + U/\sigma_a^2)^{-1}C\Psi\Psi') \\ &\quad + \sigma_a^2\Psi\Psi'C'(C\Psi\Psi'C' + U/\sigma_a^2)^{-1}C\Psi\Psi' \end{aligned}$$

De este modo

$$\text{Cov}(\hat{Z}_{F,U} - Z_F) = \sigma_a^2\Psi\Psi'(I - \hat{A}_U C) \quad (89)$$

expresion que es similar a (60), es decir

$$\text{Cov}(\hat{Z}_{F,U} - Z_F) = \text{Cov}[E(\hat{Z}_F|Z_0) - Z_F] - \sigma_a^2\Psi\Psi'C'\hat{A}_U' \quad (90)$$

Se puede concluir que $\hat{Z}_{F,U}$ es mas preciso que el pronóstico irrestricto $E(Z_F|Z_0)$; dado que si $U \rightarrow 0$, entonces, $\hat{Z}_{F,U} \rightarrow \hat{Z}_F$ y $\text{Cov}(\hat{Z}_{F,U} - Z_F) \rightarrow \text{Cov}(\hat{Z}_F - Z_F)$, mientras que si $U \rightarrow \infty$, resulta que $\hat{Z}_{F,U} \rightarrow E(Z_F|Z_0)$ y $\text{Cov}(\hat{Z}_{F,U} - Z_F) \rightarrow \text{Cov}[E(Z_F|Z_0) - Z_F]$, por lo que $\hat{Z}_{F,U}$ oscila entre \hat{Z}_F y $E(Z_F|Z_0)$.

Por último se establece un estadístico que contraste la hipótesis nula de que la informacion adicional impuesta por las restricción incierta (80) es compatible con la historia de la serie, o sea

$$e_{Y,U} = Y - CE(Z_F|Z_0) = C\Psi a_F + u \sim N_m(0, \sigma_a^2(C\Psi\Psi^t C^t + U)) \quad (91)$$

si U y σ_a^2 se suponen conocidas, se obtiene el estadístico

$$K_U = e_{Y,U}^t (C\Psi\Psi^t C^t + U/\sigma_a^2)^{-1} e_{Y,U}/\sigma_a^2 \sim \chi_m^2 \quad (92)$$

cuya contraparte muestral produce el estadístico calculado

$$K_{U, \text{calc}} = [Y - CE(Z_F|Z_0)]^t (C\Psi\Psi^t C^t + \hat{U}/\hat{\sigma}_a^2)^{-1} [Y - CE(Z_F|Z_0)]\hat{\sigma}_a^2 \quad (93)$$

el cual, cuando la muestra es grande puede compararse con una distribución χ_m^2 para determinar la significación estadística del valor calculado. En la práctica $\hat{\sigma}_a^2$ se obtiene al estimar el modelo ARIMA mientras que \hat{U} es una estimación que se debe suministrar.

2.3. Restricciones ciertas incompatibles con la historia

Hasta aquí se han analizado los casos de series con restricciones ciertas o que pueden tener cierto grado de incertidumbre. Hay otro caso que corresponde a cuando las restricciones son ciertas pero no son compatibles con el modelo ARIMA que se ha construido. Puede suceder que un fenómeno futuro no esté relacionado con la historia de la serie, por lo que habría una gran posibilidad de que se cambie la estructura de la serie: su componente tendencial, su componente estocástico o bien los valores del parámetros.

Analizando el caso del cambio en la estructura determinista del modelo, se puede establecer una función de intervención en algún momento τ que corresponde al momento en que ocurra el fenómeno que altere dicha estructura. La función de intervención que se plantea es:

$$(1 - \delta L)\varepsilon_{\tau,t} = \omega S_{\tau,t} \quad (94)$$

donde

$$S_{\tau,t} = \begin{cases} 1, & \text{si } t \geq \tau \\ 0, & \text{si } t < \tau \end{cases}$$

Con esta función de intervención sucede que si $\delta = 0$, la función corresponde a un cambio de nivel de magnitud ω , una disminución del tipo exponencial cuando $\delta < 1$, un aumento rectilíneo cuando $\delta = 1$ y un cambio explosivo si $\delta > 1$.

Es posible que también resulte un cambio en la estructura estocástica del modelo, en respuesta al efecto de la variable exógena; con fines ilustrativos consideremos que un proceso ruido blanco $\{v_t\}$ se suma al modelo ARIMA construido inicialmente. El nuevo modelo que incluye los efectos determinísticos y estocásticos es:

$$Z_{t,D,V} = Z_t + \left(\frac{\omega}{1 - \delta L} + v_t \right) S_{\tau,t} \quad \text{con } t = 1, \dots, N, \dots, N + H. \quad (95)$$

Se definen los siguientes vectores que contienen los efectos determinísticos y estocásticos de la serie: $D_F = (D_{N+1}, \dots, D_{N+H})'$ y $V_F = (V_{N+1}, \dots, V_{N+H})'$ El vector de predicciones es por tanto:

$$Z_{F,D,V} = Z_F + D_F + V_F \quad (96)$$

Si se toma la esperanza condicional dada la historia Z_0 resultará que

$$E(Z_{F,D,V}|Z_0) = E(Z_F|Z_0) + D_F \quad (97)$$

dado que D_F tiene una naturaleza no aleatoria y el origen del vector V_F es un ruido blanco, es decir su esperanza es 0. Además, esta esperanza corresponde al pronóstico con Error cuadrático medio mínimo, tal como se vio anteriormente.

Realizando el cálculo de la matriz de varianza-covarianza del error de pronóstico dada la información histórica, se tiene:

$$\text{Cov}\{[Z_{F,D,V} - E(Z_{F,D,V}|Z_0)]|Z_0\} = \text{cov}[Z_F + D_F + V_F - E(Z_F|Z_0) - D_F|Z_0] \quad (98)$$

Asumiendo la independencia entre $Z_F - E(Z_F|Z_0)$ y V_F ,

$$\text{Cov}\{[Z_{F,D,V} - E(Z_{F,D,V}|Z_0)]|Z_0\} = \text{Cov}[Z_F - E(Z_F|Z_0)|Z_0] + \text{Cov}(V_F|Z_0) \quad (99)$$

$$= \text{Cov}[\Psi a_F|Z_0] + \text{Cov}(V_F|Z_0) \quad (100)$$

$$= \sigma_a^2 \Psi \Psi' + \sigma_v^2 I \quad (101)$$

Teniendo en cuenta que tanto ω , δ y σ_v^2 son desconocidos, se debe utilizar la información adicional acerca de los valores futuros de la serie para estimarlos. Esta información adicional está contenida en el vector:

$$Y = CZ_{F,D,V} \quad (102)$$

Consideremos un pronóstico lineal de la forma

$$\tilde{Z}_{F,D,V} = AY + A_* E(Z_{F,D,V}|Z_0) \quad (103a)$$

$$= A[C(E[Z_F|Z_0] + \Psi a_F + D_F + V_F)] + A_* E(Z_F + D_F + V_F|Z_0) \quad (103b)$$

$$= (AC + A_*)E(Z_F|Z_0) + AC(\Psi a_F + D_F + V_F) + A_* D_F \quad (103c)$$

Para asegurar el insesgamiento hacemos, establecemos que la diferencia entre las esperanzas de \tilde{Z}_F y Z_F

$$\begin{aligned} \tilde{0} &= E(\tilde{Z}_{F,D,V} - Z_{F,D,V}|Z_0, Y) \\ &= E[(AC + A_*)E(Z_F|Z_0) + AC(\Psi a_F + D_F + V_F) \\ &\quad + A_* D_F - E(Z_F|Z_0) - \Psi a_F - D_F - V_F|Z_0, Y] \\ &= E[(AC + A_* - I)E(Z_F|Z_0) + (AC - I)\Psi a_F \\ &\quad + (AC - I)D_F + (AC - I)V_F + A_* D_F|Z_0, Y] \\ &= (AC + A_* - I)E(Z_F|Z_0) + (AC - I)D_F + A_* D_F \end{aligned}$$

Luego, obtenemos

$$E(Z_F|Z_0) = (AC + A_*)E(Z_F|Z_0) + (AC - I)D_F + A_*D_F \quad (104)$$

$$(AC + A_*)E(Z_F|Z_0) = E(Z_F|Z_0) - (AC - I)D_F - A_*D_F \quad (105)$$

El error de pronóstico será

$$\begin{aligned} \tilde{Z}_{F,D,V} - Z_{F,D,V} &= (AC + A_*)E(Z_F|Z_0) + AC(\Psi a_F + D_F + V_F) \\ &\quad + A_*D_F - E(Z_F|Z_0) - \Psi a_F - D_F - V_F \end{aligned}$$

Reemplazando (105) en la expresión anterior, resulta

$$\begin{aligned} \tilde{Z}_{F,D,V} - Z_{F,D,V} &= AC(\Psi a_F + D_F + V_F) + A_*D_F - \Psi a_F \\ &\quad - (AC - I)D_F - A_*D_F - D_F - V_F \\ &= AC\Psi a_F + ACV_F - \Psi a_F - V_F \\ &= (AC - I)\Psi a_F + (AC - I)V_F \end{aligned}$$

De esta manera,

$$\tilde{Z}_{F,D,V} - Z_{F,D,V} = (AC - I)(\Psi a_F + V_F) \quad (106)$$

entonces,

$$\begin{aligned} \text{Cov}(\tilde{Z}_{F,D,V} - Z_{F,D,V}) &= (AC - I)E[(\Psi a_F + V_F)(\Psi a_F + V_F)'](AC - I)' \\ &= (AC - I)E[\Psi a_F a_F' \Psi' + \Psi a_F V_F' + V_F a_F' \Psi' + V_F V_F'] (AC - I)' \\ &= (AC - I)(\Psi \sigma_a^2 \Psi' + \sigma_v^2 I)(AC - I)' \\ &= (AC \Psi \sigma_a^2 \Psi' + AC \sigma_v^2 - \Psi \sigma_a^2 \Psi' - \sigma_v^2 I)(C^t A' - I) \\ &= AC \Psi \sigma_a^2 \Psi' C' A' - AC \Psi \sigma_a^2 \Psi' + AC \sigma_v^2 C^t A' - AC \sigma_v^2 \\ &\quad - \Psi \sigma_a^2 \Psi' C' A' + \Psi \sigma_a^2 \Psi' - \sigma_v^2 I C' A' + \sigma_v^2 I \end{aligned}$$

Debe encontrarse A de tal manera que se minimice la varianza, es decir la traza de la Covarianza. Se plantea nuevamente un problema de optimización, que se resuelve derivando la expresión anterior:

$$\begin{aligned} \frac{\partial \text{Var}}{\partial A} &= 2AC\Psi\Psi'C'\sigma_a^2 - \Psi\Psi'C'\sigma_a^2 + 2ACC'\sigma_v^2 - C'\sigma_v^2 - \Psi\Psi'C'\sigma_a^2 - C'\sigma_v^2 \\ &= 2AC\Psi\Psi'C'\sigma_a^2 - 2\Psi\Psi'C'\sigma_a^2 + 2ACC'\sigma_v^2 - 2C'\sigma_v^2 \end{aligned}$$

De la condición de primer orden para la optimización se tiene que

$$\hat{A} = (\Psi\Psi'C'\sigma_a^2 + C'\sigma_v^2)(C\Psi\Psi'C'\sigma_a^2 + CC'\sigma_v^2)^{-1} \quad (107)$$

Reemplazando (105) en (103c) se obtiene el pronóstico restringido

$$\begin{aligned}
\hat{Z}_{F,D,V} &= E(Z_F|Z_0) - (AC - I)D_F + AC(\Psi a_F + D_F + V_F) \\
&= E(Z_F|Z_0) + D_F + \hat{A}C\Psi a_F + \hat{A}CV_F \\
&= E(Z_{F,D,V}|Z_0) + \hat{A}C(\Psi a_F + V_F) \\
&= E(Z_{F,D,V}|Z_0) + \hat{A}C[Z_{F,D,V} - E(Z_{F,D,V}|Z_0)] \\
&= E(Z_{F,D,V}|Z_0) + \hat{A}[CZ_{F,D,V} - CE(Z_{F,D,V}|Z_0)]
\end{aligned}$$

Por lo que

$$\hat{Z}_{F,D,V} = E(Z_{F,D,V}|Z_0) + \hat{A}[Y - CE(Z_{F,D,V}|Z_0)] \quad (108)$$

La solución del pronóstico restringido que se expresa en la ecuaciones anteriores, carece de utilidad práctica ya que se desconocen los parámetros del cambio estructural. Esta dificultad se supera cuando se especifica la solución a los dos casos de cambio estructural tomándolos por separado, lo cual limita el número de parámetros por estimar. Por lo cual se usarán las expresiones

$$Z_{F,D} = Z_F + D_F \quad y \quad Z_{F,V} = Z_F + V_F$$

Para el caso de $Z_{F,D} = Z_F + D_F$ y razonando como antes, se propone un pronóstico de la forma

$$\tilde{Z}_{F,D} = AY + A_*E(Z_F, D|Z_0)$$

donde $Y = CZ_{F,D} = C[E(Z_F|Z_0) + \Psi a_F + D_F]$, de esta manera, sustituyendo en la ecuación anterior resulta:

$$\tilde{Z}_{F,D} = AC[E(Z_F|Z_0) + \Psi a_F + D_F] + A_*E(Z_F, D|Z_0) \quad (109a)$$

$$= (AC + A_*)E(Z_F|Z_0) + AC(\Psi a_F + D_F) + A_*D_F \quad (109b)$$

Para probar el Insesgamiento, hacemos

$$\begin{aligned}
E(\tilde{Z}_{F,D} - Z_{F,D}|Z_0, Y) &= E[(AC + A_*)E(Z_F|Z_0) + AC(\Psi a_F + D_F) \\
&\quad + A_*D_F - E(Z_F|Z_0) - \Psi a_F - D_F|Z_0, Y]
\end{aligned}$$

o mejor,

$$\begin{aligned}
E(\tilde{Z}_{F,D} - Z_{F,D}|Z_0, Y) &= E[(AC + A_* - I)E(Z_F | Z_0) + (AC - I)\Psi a_F + (AC - I)D_F | Z_0, Y] \\
&= E[(AC + A_* - I)E(Z_F | Z_0) + (AC - I)(\Psi a_F + D_F) | Z_0, Y] \\
&= (AC + A_* - I)E(Z_F | Z_0) + (AC - I)D_F = 0
\end{aligned}$$

dando como resultado,

$$E(Z_F | Z_0) = (AC + A_*)E(Z_F | Z_0) + (AC - I)D_F \quad (110)$$

Luego el error de pronóstico es:

$$\tilde{Z}_{F,D} - Z_{F,D} = (AC + A_*)E(Z_F | Z_0) + AC(\Psi a_F + D_F) + A_*D_F - E(Z_F | Z_0) - \Psi a_F - D_F \quad (111a)$$

Reemplazando (110) en (111a) resulta:

$$\begin{aligned} E(\tilde{Z}_{F,D} - Z_{F,D} | Z_0, Y) &= E(Z_F | Z_0) - (AC - I)D_F + AC(\Psi a_F + D_F) + A_*D_F \\ &\quad - E(Z_F | Z_0) - \Psi a_F - D_F \\ &= -ACD_F + D_F + AC\Psi a_F + ACD_F + A_*D_F - \Psi a_F - D_F \\ &= AC\Psi a_F + A_*D_F - \Psi a_F \\ &= (AC - I)\Psi a_F + A_*D_F \end{aligned}$$

entonces,

$$\begin{aligned} \text{Cov}(\tilde{Z}_{F,D} - Z_{F,D}) &= (AC - I)\Psi\sigma_a^2\Psi'(AC - I)' \\ &= (AC\Psi\sigma_a^2\Psi' - \Psi\sigma_a^2\Psi')(C'A' - I) \\ &= AC\Psi\sigma_a^2\Psi'C'A' - AC\Psi\sigma_a^2\Psi' - \Psi\sigma_a^2\Psi'C'A' + \Psi\sigma_a^2\Psi' \\ &= \sigma_a^2[AC\Psi\Psi'C'A' - AC\Psi\Psi' - \Psi\Psi'C'A' + \Psi\Psi'] \end{aligned}$$

luego,

$$\begin{aligned} \frac{\partial \text{Var}}{\partial A} &= \sigma_a^2[2AC\Psi\Psi'C' - \Psi\Psi'C' - \Psi\Psi'C'] \\ &= \sigma_a^2[2AC\Psi\Psi'C' - 2\Psi\Psi'C'] = 0 \\ &= AC\Psi\Psi'C' - \Psi\Psi'C' = 0 \end{aligned}$$

y así,

$$\hat{A}_D = (\Psi\Psi'C')(C\Psi\Psi'C')^{-1} \quad (112)$$

De esta manera,

$$\begin{aligned} \text{Cov}(\tilde{Z}_{F,D} - Z_{F,D}) &= \sigma_a^2(\hat{A}C - I)\Psi\Psi'(C'\hat{A}'_D - I) \\ &= \sigma_a^2[\Psi\Psi'C'(C\Psi\Psi')^{-1}C\Psi\Psi' - \Psi\Psi'] [C'(C\Psi\Psi'C')^{-1}C\Psi\Psi' - I] \\ &= \sigma_a^2\Psi\Psi'[C'\hat{A}'_D - I] \end{aligned}$$

Por último, reemplazando (110) en (109b) se obtiene

$$\begin{aligned} \hat{Z}_{F,D} &= E(Z_F | Z_0) - (\hat{A}_D C - I)D_F + \hat{A}_D C(\Psi a_F + D_F) \\ &= E(Z_F | Z_0) - \hat{A}_D C D_F + D_F + \hat{A}_D C \Psi a_F + \hat{A}_D C D_F \\ &= E(Z_F | Z_0) + D_F + \hat{A}_D C \Psi a_F \end{aligned}$$

Ejemplo 2.1. Con el fin de aplicar los desarrollos formales presentados en este trabajo, consideramos un proceso $Z_t \sim ARMA(1, 1)$ definido por

$$Z_t = 0,9Z_{t-1} + a_t - 0,8a_{t-1}$$

Haciendo uso del operador polinomial de retardos resulta:

$$\begin{aligned} Z_t &= 0,9LZ_t + a_t - 0,8La_t \\ (1 - 0,9L)Z_t &= (1 - 0,8L)a_t \\ Z_t &= \frac{(1 - 0,8L)}{(1 - 0,9L)}a_t = \frac{\theta(L)}{\phi(L)}a_t = \Psi(L)a_t \end{aligned}$$

pero $\Psi(L) = 1 + \psi_1L + \psi_2L^2 + \dots$, entonces,

$$\begin{aligned} \Psi(L)(1 - 0,9L) &= (1 - 0,8L) \\ (1 + \psi_1L + \psi_2L^2 + \psi_3L^3 + \dots)(1 - 0,9L) &= (1 - 0,8L) \end{aligned}$$

La igualdad de los coeficientes de los polinomios nos permite obtener:

$$\begin{aligned} \psi_1 &= 0,1 \\ \psi_2 &= 0,9\psi_1 = 0,09 \\ \psi_3 &= 0,9\psi_2 = (0,9)^2\psi_1 = 0,081 \\ \psi_4 &= 0,9\psi_3 = (0,9)^3\psi_1 = 0,0729 \\ &\vdots \end{aligned}$$

De este modo, $\psi_j = (0,9)^{j-1}(0,1)$, para $j \geq 1$ y $\psi_0 = 1$. Por lo que

$$\Psi(L) = 0,1 \sum_{j=1}^{\infty} (0,9)^{j-1} L^j$$

Supongamos que la serie es trimestral, entonces, para un horizonte de pronóstico de 1 año, resulta ($H = 4$), entonces:

$$\Psi = \begin{pmatrix} -\Psi_0 & 0 & 0 & 0 \\ -\Psi_1 & -\Psi_0 & 0 & 0 \\ -\Psi_2 & -\Psi_1 & -\Psi_0 & 0 \\ -\Psi_3 & -\Psi_2 & -\Psi_1 & -\Psi_0 \end{pmatrix}$$

Donde,

$$\begin{aligned} \Psi_0 &= 1 \\ \Psi_1 &= -0,1(1) = -0,1 \\ \Psi_2 &= -0,1(0,9) = -0,09 \\ \Psi_3 &= -0,1(0,9)^2 = -0,081 \end{aligned}$$

Entonces,

$$\Psi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0,1 & 1 & 0 & 0 \\ 0,09 & 0,1 & 1 & 0 \\ 0,081 & 0,09 & 0,1 & 1 \end{pmatrix}$$

Pronósticos irrestrictos: Tales pronósticos para el modelo en consideración, se obtienen considerando: $t = 1, 2, \dots, T$, por lo que la serie histórica es $\{Z_1, Z_2, \dots, Z_T\} = Z_0$. Luego,

$$\begin{aligned} E_t[Z_{t+1}|Z_0] &= \widehat{Z}_t(1) = E_t(Z_{t+1}|Z_0) \\ \widehat{Z}_t(1) &= E_t[0,9Z_t + a_{t+1} - 0,8a_t|Z_0] \\ &= 0,9E_t(Z_t) + E_t(a_{t+1}) - 0,8E_t(a_t) \\ &= 0,9Z_t - 0,8E_t[Z_t - 0,9Z_{t-1} - 0,8a_{t-1}] \\ &= 0,9Z_t - 0,8[E_t(Z_t) - 0,9E_t(Z_{t-1}) - 0,8E_t(a_{t-1})] \\ &= 0,9Z_t - 0,8[Z_t - E_t(0,9Z_{t-1} - 0,8a_{t-1})] \\ &= 0,9Z_t - 0,8[Z_t - \widehat{Z}_{t-1}(1)] \\ E_t[Z_{t+2}|Z_0] &= \widehat{Z}_t(2) = E_t(Z_{t+2}|Z_0) \\ &= E_t[0,9Z_{t+1} + a_{t+2} - 0,8a_{t+1}|Z_0] \\ &= 0,9[Z_t - 0,8(Z_t - \widehat{Z}_{t-1}(1))] \\ &= 0,9\widehat{Z}_t(1) \\ E_t[Z_{t+3}|Z_0] &= \widehat{Z}_t(3) = E_t(Z_{t+3}|Z_0) \\ &= E_t[0,9Z_{t+2} + a_{t+3} - 0,8a_{t+2}|Z_0] \\ &= 0,9[0,9\widehat{Z}_t(1)] \\ &= (0,9)^2\widehat{Z}_t(1) \\ E_t[Z_{t+4}|Z_0] &= \widehat{Z}_t(4) = E_t(Z_{t+4}|Z_0) \\ &= E_t[0,9Z_{t+3} + a_{t+4} - 0,8a_{t+3}|Z_0] \\ &= 0,9[(0,9)^2\widehat{Z}_t(1)] \\ &= (0,9)^3\widehat{Z}_t(1) \end{aligned}$$

En conclusión,

$$\begin{aligned} \widehat{Z}_t(1) &= 0,9Z_t - 0,8[Z_t - \widehat{Z}_{t-1}(1)] \\ \widehat{Z}_t(h) &= (0,9)^{h-1}\widehat{Z}_t(1). \quad h \geq 2 \end{aligned}$$

Como se puede observar, los pronósticos para $h \geq 2$ dependen de $\widehat{Z}_t(1)$ pero $\widehat{Z}_t(1)$ a su vez depende de $\widehat{Z}_{t-1}(1)$ y este a su vez de $\widehat{Z}_{t-2}(1)$, obteniéndose de manera recursiva $\widehat{Z}_0(1)$, el cual consideramos igual a Z_1 , es decir:

$$\widehat{Z}_0(1) = Z_1$$

Entonces,

$$\begin{aligned}
\widehat{Z}_1(1) &= 0,9Z_1 - 0,8[Z_1 - \widehat{Z}_0(1)] \\
&= 0,9Z_1 \\
\widehat{Z}_2(1) &= 0,9Z_2 - 0,8[Z_2 - \widehat{Z}_1(1)] \\
&= 0,1Z_2 + 0,8\widehat{Z}_1(1) \\
&= 0,1Z_2 + 0,72Z_1 \\
\widehat{Z}_3(1) &= 0,9Z_3 - 0,8[Z_3 - \widehat{Z}_2(1)] \\
&= 0,1Z_3 + 0,8\widehat{Z}_2(1) \\
&= 0,1Z_3 + 0,8[0,1Z_2 + 0,72Z_1] \\
&= 0,1Z_3 + 0,08Z_2 + 0,576Z_1 \\
&\vdots \\
\widehat{Z}_{T-1}(1) &= 0,9Z_{T-1} - 0,8[Z_{T-1} - \widehat{Z}_{T-2}(1)]
\end{aligned}$$

Pronósticos restringidos: Dada la información adicional acerca de los valores futuros de la serie, establezcamos estas dos restricciones:

$$\begin{aligned}
Y_1 &= Z_{T+4} \\
Y_2 &= Z_{T+4} - Z_{T+1}
\end{aligned}$$

En forma compacta se escribe:

$$\begin{aligned}
Y &= CZ_F \\
\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Z_{T+1} \\ Z_{T+2} \\ Z_{T+3} \\ Z_{T+4} \end{bmatrix} \\
&= \begin{bmatrix} Z_{T+4} \\ Z_{T+4} - Z_{T+1} \end{bmatrix}
\end{aligned}$$

y como se puede observar la matriz C y el vector Z_F tienen la representación:

$$\begin{aligned}
C &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix} \\
Z_F &= [Z_{T+1} \quad Z_{T+2} \quad Z_{T+3} \quad Z_{T+4}]'
\end{aligned}$$

Debido a que los pronósticos restringidos necesitan de los errores aleatorios $\{a_F\}$ observados. Usaremos el Método Recursivo Aproximado para determinarlos.

$$\begin{aligned}
(1 - 0,9L)Z_t &= (1 - 0,8L)a_t \\
Z_t - 0,9Z_{t-1} &= a_t - 0,8a_{t-1}
\end{aligned}$$

Entonces,

$$\hat{a}_t = Z_t - 0,9Z_{t-1} + 0,8\hat{a}_{t-1}$$

Sea $\hat{a}_T = 0$ con $t = 1, 2, \dots, T$. Entonces

$$\begin{aligned}\hat{a}_{T+1} &= \hat{Z}_{T+1} - 0,9Z_T + 0,8\hat{a}_T \\ &= \hat{Z}_{T+1} - 0,9Z_T \\ \hat{a}_{T+2} &= \hat{Z}_{T+2} - 0,9Z_{T+1} + 0,8\hat{a}_{T+1} \\ &= Z_{T+2} - 0,9Z_{T+1} + 0,8(Z_{T+1} - 0,9Z_T) \\ &= Z_{T+2} - 0,9Z_{T+1} + 0,8Z_{T+1} - 0,72Z_T \\ &= \hat{Z}_{T+2} - 0,1\hat{Z}_{T+1} - 0,72Z_T \\ \hat{a}_{T+3} &= \hat{Z}_{T+3} - 0,9Z_{T+2} + 0,8\hat{a}_{T+2} \\ &= \hat{Z}_{T+3} - 0,9\hat{Z}_{T+2} + 0,8(\hat{Z}_{T+2} - 0,1\hat{Z}_{T+1} - 0,72Z_T) \\ &= \hat{Z}_{T+3} - 0,1\hat{Z}_{T+2} - 0,08Z_{T+1} - 0,72Z_T \\ \hat{a}_{T+4} &= \hat{Z}_{T+4} - 0,9Z_{T+3} + 0,8\hat{a}_{T+3} \\ &= \hat{Z}_{T+4} - 0,9Z_{T+3} + 0,8(\hat{Z}_{T+3} - 0,1\hat{Z}_{T+2} - 0,08Z_{T+1} - 0,72Z_T) \\ &= \hat{Z}_{T+4} - 0,1Z_{T+3} - 0,08Z_{T+2} - 0,064\hat{Z}_{T+1} - 0,576\hat{Z}_T\end{aligned}$$

Ahora, sabemos que

$$E[Z_{T+1}|Z_0] = 0,9Z_T - 0,8[Z_T - \hat{Z}_{T-1}(1)]$$

Así,

$$E[Z_F|Z_0] = \left[0,9Z_T - 0,8a_T, 0,9\hat{Z}_T(1), 0,9^2\hat{Z}_T(2), 0,9^3\hat{Z}_T(3) \right]'$$

Ahora, se puede determinar el vector \hat{A} , \hat{Z}_F y verificar la compatibilidad entre la información adicional y la información histórica.

El vector \hat{A} tiene la siguiente forma:

$$\begin{aligned}\hat{A} &= \Psi\Psi'C'(C\Psi\Psi'C')^{-1} \\ &= \begin{pmatrix} 1,0000 & -1,0000 \\ 0,9891 & -0,6606 \\ 0,9890 & -0,3285 \\ 1,0000 & 0,0000 \end{pmatrix}\end{aligned}$$

Y así,

$$\hat{Z}_F = E(Z_F|Z_0) + \hat{A}[Y - CE(Z_F|Z_0)]$$

Se realiza el cálculo anterior de \hat{Z}_F y se verifica la compatibilidad entre la restricción Y y la historia de la serie mediante el estadístico K_{cal} .

3. Pronósticos Restringidos para Series Múltiples

3.1. Modelos VAR y VECM

Como alternativa a los modelos ARIMA, se propone el uso de los modelos en los que se estudien las relaciones de un vector de variables; se pretende entonces incorporar la información adicional proveniente de las metas económicas en forma análoga a lo realizado en el modelo univariado.

Para empezar se hará un breve consideración sobre los modelos VAR, y teniendo en cuenta las relaciones de cointegración que puedan presentarse, se presentan los modelos de Vectores con corrección de errores (VECM).

3.1.1. Aspectos Metodológicos

Consideremos un vector de N variables aleatorias

$$Z_t = (Z_{1t}, \dots, Z_{Nt})' \quad (113)$$

Supongamos $E(Z_{it}) = \mu_i \quad \forall i$, a_t es un ruido blanco Gaussiano con $a_t \sim N_k(\mathbf{0}_k, \Sigma_a)$ y

$$\Sigma_a = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_k^2 \end{pmatrix} \quad (114)$$

donde $\sigma_{ij} = Cov(a_{it}, a_{jt}) \quad i \neq j$ y $\sigma_i^2 = Var(a_{it}) \quad \forall i = 1, 2, \dots, k$.

Para un conjunto de k variables $Z_t = (Z_{1t}, \dots, Z_{Nt})'$ la serie múltiple tiene la siguiente representación en forma $VAR(p)$

$$Z_t = \mu + A_1 Z_{t-1} + A_2 Z_{t-2} + \cdots + A_p Z_{t-p} + a_t \quad (115)$$

donde Z_t es un vector $K \times 1$ conformado por las observaciones de las K variables del sistema, las matrices de coeficientes $A_i (i = 1, \dots, p)$ son de tamaño $(K \times K)$ de coeficientes a estimar y a_t un vector aleatorio conformado por K perturbaciones estocásticamente independientes con $E(a_t) = 0$ y matriz de varianza-covarianza Σ_a .

Haciendo uso del operador de retardos, resulta,

$$\begin{aligned} Z_t - A_1 L Z_t - A_2 L^2 Z_t - \cdots - A_p L^p Z_t &= \mu + a_t \\ (I - A_1 L - A_2 L^2 - \cdots - A_p L^p) Z_t &= \mu + a_t \\ \Pi(L) Z_t &= \mu + a_t \end{aligned}$$

o lo que es lo mismo,

$$\Pi(L)(Z_t - \mu) = a_t \quad (116)$$

con $\Pi(L) = I - A_1L - A_2L^2 - \dots - A_pL^p$.

Si escribimos $\Pi(L)(Z_t - \mu) = a_t$ como $Z_t - \mu = [\Pi(L)]^{-1}a_t$ y hacemos $[\Pi(L)]^{-1} = \Psi(L)$, resulta

$$\Psi(L)\Pi(L) = I \quad \text{con} \quad \Psi(L) = I + \Psi_1(L) + \Psi_2L^2 + \dots$$

Entonces,

$$Z_t - \mu = \Psi(L)a_t \quad (\text{VMA}) \quad (117)$$

Así, se puede pasar de un VAR(p) a una representación VMA pura, igualando los coeficientes de las potencias de L en la ecuación que relaciona los dos polinomios. Cuando $|\Pi(x)| = 0$ tiene sus raíces fuera del círculo unitario, se dice que la serie Z_t es estacionaria.

Decimos que el proceso es estable si

$$\det(I - A_1x - \dots - A_px^p) \neq 0 \quad \text{para} \quad |x| \leq 1 \quad (118)$$

Es decir, el polinomio definido por el determinante del operador de retardo no tiene raíces dentro o sobre el círculo unitario. Si el polinomio en (118) tiene una raíz unitaria (el determinante es cero para $x = 1$), entonces algunas o todas las variables son integradas; por conveniencia se asume que al menos son $I(1)$.

Mediante el concepto de cointegración es posible determinar como se relacionan las variables en el largo y el corto plazo dentro del sistema. Como se citó anteriormente si las variables son de integradas de orden 1, una Combinación lineal de ellas puede resultar de orden cero. Pero en general se puede hacer la siguiente definición.

Definición 3.1. *Los componentes del vector $Z_t = (Z_{1t}, \dots, Z_{nt})$ están cointegrados de orden $(d-b)$, que en forma abreviada se denotará por $CI(d-b)$, si:*

1. *Todos los componentes de Z_t son integrados de orden d .*
2. *Existe un vector $b = (b_1, \dots, b_n)$ para el cual la Combinación lineal:*

$$b_1Z_{1t} + b_2Z_{2t} + \dots + b_nZ_{nt}$$

es integrada de orden $(d-b)$ para $b > 0$.

Si hay alguna relación de cointegración entre las variables que intervienen en el estudio de series que se está haciendo, el modelo (115) puede mejorarse mediante la utilización de un Modelo de Corrección de Error (VECM).

Definición 3.2. *Un modelo VECM se obtiene de un modelo VAR por niveles y tiene la siguiente ecuación:*

$$\Delta Z_t = \Pi Z_{t-p} + \Gamma_1 \Delta Z_{t-1} + \cdots + \Gamma_{p-1} \Delta Z_{t-p+1} + u_t \quad (119)$$

con Z_t vector de variables integradas de orden uno, u_t como vector de error, y:

$$\begin{aligned} \Pi &= -(I_k - A_1 - \cdots - A_p) \\ \Gamma_i &= -(A_{i+1} + \cdots + A_p) \quad \text{para } i = 1, \dots, p-1 \end{aligned}$$

Para llegar a la (119) es necesario sustraer Z_{t-1} en (115) en ambos miembros y reorganizar términos. Si consideramos un modelo VAR(2) con K variables, se obtiene:

$$\begin{aligned} Z_t &= \mu + A_1 Z_{t-1} + A_2 Z_{t-2} + a_t \\ Z_t - Z_{t-1} &= A_1 Z_{t-1} - Z_{t-1} + A_2 Z_{t-2} + a_t \\ \Delta Z_t &= A_1 Z_{t-1} - Z_{t-1} + Z_{t-2} - Z_{t-2} + A_2 Z_{t-2} + a_t \\ \Delta Z_t &= A_1 Z_{t-1} - \Delta Z_{t-1} - Z_{t-2} + A_2 Z_{t-2} + a_t \end{aligned}$$

Ahora, se suma y resta $A_1 Z_{t-2}$ para construir $A_1 \Delta Z_{t-1}$

$$\begin{aligned} \Delta Z_t &= A_1 Z_{t-1} - A_1 Z_{t-2} + A_1 Z_{t-2} - \Delta Z_{t-1} - Z_{t-2} + A_2 Z_{t-2} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} + A_1 Z_{t-2} - \Delta Z_{t-1} - Z_{t-2} + A_2 Z_{t-2} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} - \Delta Z_{t-1} - Z_{t-2} + A_1 Z_{t-2} + A_2 Z_{t-2} + a_t \\ \Delta Z_t &= (-I + A_1) \Delta Z_{t-1} + (-I + A_1 + A_2) Z_{t-2} + a_t \end{aligned}$$

Haciendo $\Gamma_1 = (-I + A_1)$ y $\Pi = (-I + A_1 + A_2)$, obtenemos la forma VECM, así

$$\Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \Pi Z_{t-2} + a_t \quad (120)$$

Para el caso VAR(3) con K variables se obtiene:

$$\begin{aligned} Z_t &= \mu + A_1 Z_{t-1} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ Z_t - Z_{t-1} &= A_1 Z_{t-1} - Z_{t-1} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 Z_{t-1} - Z_{t-1} + Z_{t-2} - Z_{t-2} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 Z_{t-1} - \Delta Z_{t-1} - Z_{t-2} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 Z_{t-1} + A_1 Z_{t-2} - A_1 Z_{t-2} - \Delta Z_{t-1} - Z_{t-2} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} - \Delta Z_{t-1} + A_1 Z_{t-2} - Z_{t-2} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} - \Delta Z_{t-1} + A_1 Z_{t-2} + A_1 Z_{t-3} - A_1 Z_{t-3} - Z_{t-2} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} - \Delta Z_{t-1} + A_1 \Delta Z_{t-2} + A_1 Z_{t-3} - Z_{t-2} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} - \Delta Z_{t-1} + A_1 \Delta Z_{t-2} + A_1 Z_{t-3} - Z_{t-2} + Z_{t-3} - Z_{t-3} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= A_1 \Delta Z_{t-1} - \Delta Z_{t-1} + A_1 \Delta Z_{t-2} + A_1 Z_{t-3} - \Delta Z_{t-2} - Z_{t-3} + A_2 Z_{t-2} + A_3 Z_{t-3} + a_t \\ \Delta Z_t &= (-I + A_1) \Delta Z_{t-1} + (-I + A_1 + A_2) \Delta Z_{t-2} + (-I + A_1 + A_2 + A_3) Z_{t-3} + a_t \end{aligned}$$

Haciendo,

$$\begin{aligned}\Gamma_1 &= (-I + A_1) \\ \Gamma_2 &= (-I + A_1 + A_2) \\ \Pi &= (-I + A_1 + A_2 + A_3)\end{aligned}$$

obtenemos la forma VECM, así

$$\Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \Gamma_2 \Delta Z_{t-2} + \Pi Z_{t-3} + a_t \quad (121)$$

En general, si $Z_t \sim \text{VAR}(p)$ se tiene que

$$\Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \Gamma_2 \Delta Z_{t-2} + \cdots + \Gamma_{p-1} \Delta Z_{t-p+1} + \Pi Z_{t-p} + a_t \quad (122)$$

Donde,

$$\begin{aligned}\Gamma_1 &= (-I + A_1) \\ \Gamma_2 &= (-I + A_1 + A_2) \\ \Gamma_{p-1} &= (-I + A_1 + \cdots + A_{p-1}) \\ \Pi &= (-I + A_1 + A_2 + \cdots + A_p)\end{aligned}$$

El modelo VECM en (122) se puede escribir como

$$\Delta Z_t = \Pi Z_{t-p} + \sum_{i=1}^{p-1} \Gamma_i \Delta Z_{t-i} + a_t \quad (123)$$

La expresión (123) es idéntica a la expresión (122). De esta forma, hemos pasado de un modelo $\text{VAR}(p)$ con K variables a un modelo VECM aplicando diferencias y reorganizando términos.

Proposición 3.1. *Si el proceso $\text{VAR}(p)$ tiene una raíz unitaria, es decir, si $\det(I_k - A_1 L - \cdots - A_p L^p) = 0$ para $L=1$, la matriz $\Pi = -(I_k - A_1 - \cdots - A_p)$ es singular.*

Demostración. Supongamos que el rango(ρ) de Π es r , es decir, $\rho(\Pi) = r$. Entonces Π puede ser escrita como producto de matrices ($K \times r$) α, β con la propiedad de que $\rho(\alpha) = \rho(\beta) = r$, y así, $\Pi = \alpha\beta'$.

Entonces, $\Pi Z_{t-1} = \alpha\beta' Z_{t-1}$ donde $\beta' Z_{t-1}$ es $I(0)$ y contiene las relaciones de cointegración. Se sigue entonces que $\rho(\Pi) = r$; el rango de Π es por lo tanto denominado como el rango de cointegración del sistema y β es la matriz de cointegración. \square

Los modelos (115) y (119) se pueden extender para incorporar la parte determinística, como intercepto, término de tendencia lineal y variables ficticias estacionales, las cuales

podrían dar una mejor representación del conjunto de datos; para incorporar la parte determinística se hace uso de la expresión

$$Z_t = v_t + y_t \quad (124)$$

en donde v_t es la parte determinística y y_t la parte estocástica y está definida como en (115) o (119), es decir, y_t tiene una representación VAR o VECM, en otras palabras $y_t = \mu + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + a_t$ ó $\Delta y_t = \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + \Pi y_{t-p} + a_t$. Si asumimos que v_t es un término de tendencia lineal, entonces $v_t = v_0 + v_1 t$, y así la correspondiente representación VAR sería,

$$Z_t = v_0 + v_1 t + A_1 Z_{t-1} + A_2 Z_{t-2} + \dots + A_p Z_{t-p} + a_t \quad (125)$$

y la representación VECM($p - 1$) toma la forma

$$\Delta Z_t = v_0 + v_1 t + \Gamma_1 \Delta Z_{t-1} + \dots + \Gamma_{p-1} \Delta Z_{t-p+1} + \Pi Z_{t-p} + a_t \quad (126)$$

Además de la parte de determinística, en la práctica se pueden realizar otras generalizaciones, como por ejemplo, se pueden agregar otras variables estocásticas además de la determinística. En representación tipo VECM e incluyendo todos los terminos, tanto estocásticos como determinísticos, tendría la forma

$$\Delta Z_t = \Pi Z_{t-p} + \Gamma_1 \Delta Z_{t-1} + \dots + \Gamma_{p-1} \Delta Z_{t-p+1} + C D_t + B x_t + a_t \quad (127)$$

Donde x_t es una variables estocástica sin modelo, en el sentido de que no hay una ecuación explicativa que la defina, D_t contiene todos los regresores asociados con los términos determinísticos y C , B son matrices de parámetros.

De acuerdo con el enfoque de Engle y Granger si consideramos la variables z_{1t} y z_{2t} , la ruta hacia la búsqueda de la existencia de cointegración entre estas, se determina básicamente por examinar si los residuos de la regresión inicial:

$$z_{1t} = \alpha z_{2t} + u_t$$

son estacionarios, es decir, $u_t \sim I(0)$. Si esto es afirmativo, hay desde luego cointegración. Para hacer esta verificación, se pueden utilizar los contrastes de raíces unitarias habituales como Augmented Dickey-Fuller, ADF; pero teniendo en cuenta que las perturbaciones no son observables, se sugiere que para aceptar la cointegración, los números críticos para este contraste sean más elevados.²

Por su parte el procedimiento de Johansen, implica el uso de dos estadísticos: el de la traza y el del máximo valor propio, como pruebas de cointegración.³ Se hace necesario

²Davidson y Mackinnon (1993)

³Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegrated Vectors in Gaussian Vector Autoregressive Models", *Econometrica* 59, pp. 1551-1580.

considerar para la consideración de la cointegración si el modelo de corrección del error tiene interceptos o tendencia, al igual que el var, como en las ecuaciones dadas por (126) y (127). Después de ejecutados los test, se determina si hay cointegración, y el número de relaciones de cointegración existentes.

Uno de los aspectos más importantes a considerar cuando se utilizan modelos VAR, es el número de rezagos a considerar; si se escogen pocos, no se capturarían las relaciones dinámicas de las variables; por el contrario si se escogen demasiados, podrían aparecer problemas de colinealidad y autocorrelación, con lo que se tendrían parámetros inconsistentes. Para ayudar a determinar el número de rezagos de este modelo, se puede usar la minimización de los Criterios de Schwarz y Akaike.

Es importante la especificación de las variables endógenas y exógenas; se pretende a través de métodos econométricos de evitar la arbitrariedad en dicha asignación y de esta manera, lograr que los datos permitan hacer una estimación correcta. En este caso, se hacen exógenas a las variables haciendo uso de la teoría económica, pero la validación final es un asunto demostrable. Recordemos los siguientes conceptos:

Definición 3.3. *Un conjunto de variables x_t es llamado débilmente exógeno para un vector de parámetros de interés, por ejemplo θ , si al estimar θ dentro de un modelo condicional (condicionado a x_t) no presupone una pérdida de información relativa a la estimación del vector en un modelo que no esta condicionado a x_t .*

Definición 3.4. *Una serie de variables x_t es llamada fuertemente pérdida si es débilmente pérdida para los parámetros del modelo condicional y los pronósticos de Z_t están supeditados a x_t sin pérdida en la precisión de los pronósticos.*

Definición 3.5. *Por último, x_t se denomina super exógena para θ si x_t es débilmente exógena para θ y las acciones de política que afectan el proceso marginal de x_t no afectan los parámetros del proceso condicional. Por lo tanto, débil, fuerte, y super exogeneidad son los conceptos pertinentes para la estimación, previsión, y análisis de políticas, respectivamente. [Ericsson, Hendry y Mizon (1998)].*

En el caso de la exogeneidad débil, permite saber si la información que se está incorporando en el modelo es la correcta y suficiente. Para probar que una variable es débilmente exógena, se impone una restricción al VEC, mediante la asignación del valor cero al elemento i -ésimo en la matriz α o de coeficientes de ajuste para la variable que se está analizando; se prueba la significancia estadística de los coeficientes α restantes haciendo uso del test de radio de verosimilitud, que tiene una distribución χ^2 ; si se cumple, se acepta la hipótesis nula de que la variable representada por ese elemento es exógena respecto al vector de cointegración β . En otras palabras, no hay información adicional relevante en las demás variables de ese vector que contribuya a explicar su comportamiento.

Para la especificación de un modelo VAR, un aspecto a considerar también es la *Causalidad en el sentido de Granger*. Una Prueba de causalidad de este tipo busca determinar

estadísticamente si el pasado de la variable x contiene informaci3n que preceda al comportamiento de la variable y y que por tanto, sirva para explicarla. Suponemos que tenemos dos ecuaciones:

$$\begin{aligned}\Delta y_t &= \eta_0 + \sum_{i=1}^k \eta_{1i} \Delta y_{i-1} + \sum_{i=1}^k \eta_{2i} \Delta y_{t-i} + u_{1t} \\ \Delta x_t &= \kappa_0 + \sum_{i=1}^k \kappa_{1i} \Delta y_{i-1} + \sum_{i=1}^k \kappa_{2i} \Delta y_{t-i} + u_{2t}\end{aligned}$$

En la primera ecuaci3n se busca probar la significancia estadística de las η_{2i} a trav3s de una prueba F . Se plantea como hip3tesis nula que estos coeficientes son cero, es decir que se rechaza la existencia de causalidad en el sentido de Granger contra la alternativa de que son diferentes de cero o que se acepta este tipo de causalidad.

Este mismo procedimiento se aplica para los κ_{1i} de la segunda ecuaci3n. Se asume que u_{1t} y u_{2t} no est3n correlacionados.

3.2. Pron3sticos con restricciones para modelos VAR

Nuevamente, el objeto de inter3s ser3n los pron3sticos; tal y como sucedi3 con los modelos ARIMA, el pron3stico irrestricto con error cuadr3tico m3nimo para modelos VAR ser3 el valor esperado de la variable dada la informaci3n hist3rica. T3ngase en cuenta tambi3n que un modelo VEC puede transformarse a un modelo VAR.

Definici3n 3.6. Si $Z = (Z'_1, Z'_2, \dots, Z'_N)'$ contiene la informaci3n hist3rica de la serie y $Z_F = (Z'_{N+1}, Z'_{N+2}, \dots, Z'_{N+H})'$ denota el vector de $H \geq 1$ valores futuros que se desea pron3sticar para cada serie, el pron3stico lineal 3ptimo con (ECM) m3nimo de Z'_{N+h} esta dado por $E(Z_{N+h}|Z)$ para $h = 1, 2, \dots, H$.

Para determinar el error de pron3stico se tiene en cuenta que,

$$Z_{N+h} - E(Z_{N+h}|Z) = \sum_{j=0}^{\infty} \Psi_j a_{T+h-j} - E \left[\sum_{j=0}^{\infty} \Psi_j a_{T+h-j} \right] \quad (128)$$

$$= \sum_{j=0}^{h-1} \Psi_j a_{T+h-j} + \sum_{j=h}^{\infty} \Psi_j a_{T+h-j} - E \left[\sum_{j=0}^{h-1} \Psi_j a_{T+h-j} - \sum_{j=h}^{\infty} \Psi_j a_{T+h-j} \right] \quad (129)$$

Para $j < h$ $E[\cdot] = 0$, luego la expresi3n anterior se transforma en:

$$\sum_{j=0}^{h-1} \Psi_j a_{T+h-j} - \sum_{j=h}^{\infty} \Psi_j a_{T+h-j} + \sum_{j=h}^{\infty} \Psi_j a_{T+h-j} = \sum_{j=0}^{h-1} \Psi_j a_{T+h-j} \quad (130)$$

Así obtenemos

$$Z_{N+h} - E(Z_{N+h}|Z) = \sum_{j=0}^{h-1} \Psi_j a_{T+h-j} \quad (131)$$

En forma vectorial se obtiene que la expresión (131) puede escribirse como:

$$Z_F - E(Z_F|Z) = \Psi a_F \quad (132)$$

Una expresión similar fue obtenida en el caso univariante; sin embargo la matriz Ψ tendrá la siguiente composición

$$\Psi = \begin{pmatrix} I_k & 0 & 0 & \cdots & 0 \\ \Psi_1 & I_k & 0 & \cdots & 0 \\ \Psi_2 & \Psi_1 & I_k & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Psi_{H-1} & \Psi_{H-2} & \Psi_{H-3} & \cdots & I_k \end{pmatrix}$$

Con Ψ_i matrices de tamaño $k \times k$ que se obtienen a partir de la expresión VAR y como se puede observar Ψ es una matriz diagonal inferior por bloques de dimensión $kH \times kH$, con elementos matriciales I_k en la diagonal.

Proposición 3.2. *El vector de errores $a_F = (a_{T+1}, \dots, a_{T+H})'$ tiene una distribución $N(\mathbf{0}_{kH}, I_H \otimes \Sigma_a)$, con \otimes que denota el producto de Kronecker y $a_{T+1} \dots a_{T+H}$, se distribuye $N_k(0, \Sigma)$.*

Demostración. Sea $a_F = (a_{T+1}, \dots, a_{T+H})'$. Cada uno de los vectores por los que esta conformado a_F es un vector de la forma $a_t = (a_{1t}, a_{2t}, \dots, a_{kt})'$ donde $a_t \sim N_k(0, \Sigma)$ idénticamente distribuidos para $t = 1, \dots, H$ y no correlacionadas.

La función de densidad de a_F es el producto de las densidades de sus componentes debido a la no correlación de las mismas.

$$\begin{aligned} f(a_F) &= \prod_{i=1}^H f a_i \\ &= \prod_{i=1}^H \left[(2\pi)^{-k/2} (\det \Sigma)^{-1/2} e^{\{(-1/2)a_i' \Sigma^{-1} a_i\}} \right] \\ &= (2\pi)^{-kH/2} (\det \Sigma)^{-H/2} e^{-1/2 \sum_{i=1}^H a_i \Sigma^{-1} a_i} \\ &= (2\pi)^{-kH/2} (\det \Sigma)^{-H/2} e^{-1/2 [a_1', \dots, a_H'] \Sigma^{-1} \otimes I [a_1', \dots, a_H']'} \\ &= -(2\pi)^{-kH/2} (\det I^H \det \Sigma^H)^{-1/2} e^{-1/2 A' I \otimes \Sigma^{-1} A} \quad \text{con } A = [a_1', \dots, a_H'] \\ &= -(2\pi)^{-kH/2} \det(I \otimes \Sigma)^{-1/2} e^{-1/2 (A' - 0)(I \otimes \Sigma)^{-1} (A - 0)} \end{aligned}$$

De esta manera se llega a que,

$$a_F \sim N_k H(0, I_H \otimes \Sigma)$$

□

Como se cumplen los supuestos, de igual manera que en el caso univariado se puede escribir una ecuación que permita incorporar la información adicional, así:

$$\hat{Z}_F = E(Z_F|Z) + A[Y - CE(Z_F|Z)] \quad (133)$$

donde A esta dada por,

$$A = \Psi(I_H \otimes \Sigma_a)\Psi' C' [C\Psi(I_H \otimes \Sigma_a)\Psi' C']^{-1} \quad (134)$$

Donde el vector Y contiene la información adicional o restricción que se desea agregar al modelo generador de los pronósticos. Dicho vector se relaciona con el vector de valores futuros de la variable de estudio mediante la combinación lineal

$$Y = CZ_F \quad (135)$$

Además, La matriz C es conocida y su dimension es $M \times kH$ y $\rho(C) = M$, donde $M \leq H$.

Como se verá a continuación, el pronóstico restringido dado por (133), mejora el pronóstico irrestricto determinado por la esperanza condicional de la variable. En este caso se calculará la covarianza del error de pronóstico.

Proposición 3.3. *El pronóstico restringido posee una menor covarianza que el pronóstico dado por el valor esperado de la serie*

Demostración. Si a Z_F le restamos el pronóstico restringido se obtiene:

$$\begin{aligned} Z_F - \hat{Z}_F &= E(Z_F|Z_0) - AY + ACE(Z_F|Z_0) \\ &= Z_F - AY - (I - AC)E(Z_F|Z_0) \\ &= Z_F - ACZ_F - (I - AC)E(Z_F|Z_0) \\ &= (I - AC)(Z_F - E(Z_F|Z_0)) \end{aligned}$$

como $Z_F - E(Z_F|Z_0) = \Psi a_F$, entonces

$$Z_F - \hat{Z}_F = (I - AC)\Psi a_F$$

Además $E(Z_F - \hat{Z}_F) = (I - AC)\Psi E(a_F) = 0$.

Ahora:

$$\begin{aligned} \text{Cov}(Z_F - \hat{Z}_F) &= E[(Z_F - \hat{Z}_F - E(Z_F - \hat{Z}_F))(Z_F - \hat{Z}_F - E(Z_F - \hat{Z}_F))'] \\ &= E[(I - AC)\Psi a_F a_F' \Psi' (I - C' A')] \end{aligned}$$

Como

$$A = \Psi(I \otimes \Sigma_a)\Psi' C'(C\Psi(I \otimes \Sigma_a)\Psi' C')^{-1}$$

Escribamos $A = \xi(C\xi)^{-1}$ donde $\xi = \Psi(I \otimes \Sigma_a)\Psi' C'$.

Entonces,

$$\text{Cov}(Z_F - \hat{Z}_F) = E\{[I - \xi(C\xi)^{-1}C]\Psi a_F a_F' \Psi'[I - C'(\xi' C')^{-1}\xi']\}$$

Ahora,

$$E(a_F a_F') = I \otimes \Sigma_a$$

se realizan los productos indicados y obtenemos

$$\begin{aligned} \text{Cov}(Z_F - \hat{Z}_F) &= [I - \xi(C\xi)^{-1}C][\Psi(I \otimes \Sigma_a)\Psi' - \Psi(I \otimes \Sigma_a)\Psi' C'(\xi' C')^{-1}\xi'] \\ &= [I - \xi(C\xi)^{-1}C][\Psi(I \otimes \Sigma_a)\Psi' - \xi(\xi' C')^{-1}\xi'] \\ &= \Psi(I \otimes \Sigma_a)\Psi' - \xi(C\xi)^{-1}C\Psi(I \otimes \Sigma_a)\Psi' - \xi(\xi' C')^{-1}\xi' + \xi(C\xi)^{-1}C\xi(\xi' C')^{-1}\xi' \\ &= \Psi(I \otimes \Sigma_a)\Psi' - \xi(C\xi)^{-1}C\Psi(I \otimes \Sigma_a)\Psi' - \xi(\xi' C')^{-1}\xi' + \xi(\xi' C')^{-1}\xi' \\ &= \Psi(I \otimes \Sigma_a)\Psi' - AC\Psi(I \otimes \Sigma_a)\Psi' \end{aligned}$$

por su parte, la covarianza del pronóstico irrestricto es:

$$\text{Cov}(Z_F - E(Z_F|Z_0)) = \text{Cov}(\Psi a_F) = E[\Psi a_F a_F' \Psi'] = \Psi(I \otimes \Sigma_a)\Psi'$$

Como $AC\Psi(I \otimes \Sigma_a)\Psi'$ es una matriz semidefinida positiva, entonces:

$$\text{Cov}(Z_F - \hat{Z}_F) \leq \text{Cov}(Z_F - E(Z_F|Z_0))$$

□

Verificación de la compatibilidad de la restricción con la información histórica

Para verificar la compatibilidad entre la información histórica y la información adicional incluida en forma de restricciones lineales, se hace necesario realizar una prueba estadística que utiliza el estadístico calculado

$$K_{calc} = [Y - CE(Z_F|Z)]'[C\Psi(I_H \otimes \Sigma_a)\Psi' C']^{-1}[Y - CE(Z_F|Z)] \quad (136)$$

el cual debe compararse con valores de la distribución χ_{kM}^2 , la cual surge de un resultado asintótico y en el supuesto de que los parámetros son conocidos, por lo que se considera una buena aproximación de la verdadera distribución del estadístico K .

Cuando el estadístico no produce un valor significativo, se llega a la conclusión de que ambas fuentes de información son compatibles. En contraste, cuando se rechace la compatibilidad se debería considerar, como una posible causa de ello, el que ocurrirá un cambio estructural en la serie durante el horizonte de pronóstico en consideración.

Por otro lado, el rechazo de la compatibilidad puede ser ocasionado porque solo algunas de las restricciones impuestas sobre los valores futuros sean incompatibles con los pronósticos irrestrictos correspondientes. Esta idea conduce a realizar pruebas parciales o individuales de compatibilidad, las cuales se basarían en estadísticos calculados que son de la forma

$$K_{m,calc} = [Y_m - c_m E(Z_F|Z)]^2 [c_m \Psi(I_H \otimes \sum_a) \Psi' c_m']^{-1} \quad (137)$$

y que deben compararse con una distribución χ_1^2 . Desde luego, Y_m y c_m deben seleccionarse de manera apropiada, donde $m = 1, \dots, M$ corresponde a cada una de las restricciones originalmente planteadas, pero consideradas de manera individual.

Para ilustrar el proceso de planteamiento del modelo y la obtención de la matriz Ψ , veamos el siguiente ejemplo.

Ejemplo 3.1. *Consideremos un VAR(2) con 2 ecuaciones*

$$\begin{aligned} Z_{1t} &= \alpha_{11,1} Z_{1,t-1} + \alpha_{12,1} Z_{2,t-1} + \alpha_{11,2} Z_{1,t-2} + \alpha_{12,2} Z_{2,t-2} + a_{1t} \\ Z_{2t} &= \alpha_{21,1} Z_{1,t-1} + \alpha_{22,1} Z_{2,t-1} + \alpha_{21,2} Z_{1,t-2} + \alpha_{22,2} Z_{2,t-2} + a_{2t} \end{aligned}$$

En forma matricial, haciendo $Z_t = (Z_{1t}, Z_{2t})'$ tenemos,

$$Z_t = \begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_{11,1} & \alpha_{12,1} \\ \alpha_{21,1} & \alpha_{22,1} \end{bmatrix} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{11,2} & \alpha_{12,2} \\ \alpha_{21,2} & \alpha_{22,2} \end{bmatrix} \begin{bmatrix} Z_{1,t-2} \\ Z_{2,t-2} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \quad (138)$$

En forma compacta,

$$Z_t = A_1 Z_{t-1} + A_2 Z_{t-2} + a_t \quad (139)$$

La forma matricial (138) es idéntica a la forma compacta (139), con

$$A_1 = \begin{bmatrix} \alpha_{11,1} & \alpha_{12,1} \\ \alpha_{21,1} & \alpha_{22,1} \end{bmatrix}, \quad A_2 = \begin{bmatrix} \alpha_{11,2} & \alpha_{12,2} \\ \alpha_{21,2} & \alpha_{22,2} \end{bmatrix}, \quad a_t = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Para obtener los elementos Ψ_i de la matriz Ψ , hacemos uso del operador polinomial de retardos (L) sobre el modelo definido en (139), así:

$$\begin{aligned} Z_t &= A_1 Z_{t-1} + A_2 Z_{t-2} + a_t \\ a_t &= Z_t - A_1 L Z_t - A_2 L^2 Z_t \\ a_t &= (I_2 - A_1 L - A_2 L^2) Z_t \\ Z_t &= \frac{1}{I_2 - A_1 L - A_2 L^2} a_t \end{aligned}$$

Sea $\Psi(L) = \frac{1}{I_2 - A_1 L - A_2 L^2}$, donde $\Psi(L) = I + \Psi_1 L + \Psi_2 L^2 + \Psi_3 L^3 \dots$ Entonces;

$$(I + \Psi_1 L + \Psi_2 L^2 + \Psi_3 L^3 \dots)(I_2 - A_1 L - A_2 L^2) = 1$$

multiplicando, se obtiene

$$I_2 - A_1L - A_2L^2 + \Psi_1I_2L - \Psi_1A_1L^2 - \Psi_1A_2L^3 + \Psi_2I_2L^2 - \Psi_2A_1L^3 - \Psi_2A_2L^4 + \Psi_3I_2L^3 - \Psi_3A_1L^4 - \Psi_3A_2L^5 + \Psi_4I_2L^4 - \Psi_4A_1L^5 - \Psi_4A_2L^6 + \Psi_5I_2L^5 - \Psi_5A_1L^6 - \Psi_5A_2L^7 + \dots = 1$$

Reordenando,

$$I_2 + (\Psi_1 - A_1)L + (\Psi_2 - \Psi_1A_1 - A_2)L^2 + (\Psi_3 - \Psi_2A_1 - \Psi_1A_2)L^3 + \dots = 1$$

Entonces,

$$\begin{aligned} \Psi_1 - A_1 &= 0 \\ \Psi_2 - \Psi_1A_1 - A_2 &= 0 \\ \Psi_3 - \Psi_2A_1 - \Psi_1A_2 &= 0 \\ \Psi_4 - \Psi_3A_1 - \Psi_2A_2 &= 0 \\ \Psi_5 - \Psi_4A_1 - \Psi_3A_2 &= 0 \\ &\vdots \\ \Psi_k - \Psi_{k-1}A_1 - \Psi_{k-2}A_2 &= 0 \end{aligned}$$

De tal manera que,

$$\begin{aligned} \Psi_1 &= A_1 \\ \Psi_2 &= \Psi_1A_1 + A_2 \\ \Psi_3 &= \Psi_2A_1 + \Psi_1A_2 \\ \Psi_4 &= \Psi_3A_1 + \Psi_2A_2 \\ \Psi_5 &= \Psi_4A_1 + \Psi_3A_2 \\ &\vdots \\ \Psi_k &= \Psi_{k-1}A_1 + \Psi_{k-2}A_2 \end{aligned}$$

La matriz Ψ es diagonal inferior por bloques de dimensión $kH \times kH$, con elementos matriciales I_k en la diagonal, Ψ_1 en la primera subdiagonal, Ψ_2 en la segunda subdiagonal y así sucesivamente.

1. Aplicaciones

1.1. Seguimiento al PIB de Colombia 2010

El Producto Interno Bruto (PIB) es el indicador comúnmente utilizado para medir el crecimiento económico de una región (país, estado, ciudad) durante un periodo de tiempo. El PIB es el valor monetario de los bienes y servicios finales producidos por una economía en un período determinado. Se hacen dos mediciones: la del PIB trimestral y la del PIB anual. La primera, demora 1 mes para el cálculo y la segunda, tarda seis meses (fuente: DANE).

El PIB es fundamental para evaluar el crecimiento, la evolución y la estructura de la economía colombiana y la capacidad de ésta para usar eficientemente los recursos disponibles para la producción. El producto interno bruto, mide el desarrollo de la producción, de las prestaciones de servicios y todo lo que genere riquezas en una región determinada, ya sea país o fracción del mismo. Por esto los analistas, los empresarios, el Gobierno y, en general, todos los interesados en la economía están pendientes de su evolución y de las proyecciones de la autoridad económica sobre su crecimiento para cada período de análisis.

Existen tres métodos teóricos para calcular el PIB: Método de la demanda, Método del Ingreso y Método del Valor Agregado.

El método del gasto o por el lado de la demanda: El cálculo se realiza por medio de la suma de todas las demandas finales de productos de la economía; es decir, puesto que la producción nacional puede ser consumida por el Gobierno (gasto), invertida o exportada, este punto de vista calcula el PIB como la suma de todo el consumo (el gasto más la inversión más las exportaciones menos el consumo de productos importados).

El método del valor agregado o por el lado de la oferta: El PIB se calcula sumando, para todos los bienes y servicios, el valor agregado que se genera a medida que se transforma el bien o el servicio en los diferentes sectores de la economía o ramas de la actividad económica. En este caso es útil calcular el PIB sectorial o PIB para cada sector productivo (por ejemplo el PIB del sector de la minería, la agricultura, las comunicaciones, el transporte, la industria manufacturera, la construcción, el sector financiero, etc.).

El método del ingreso o la renta: En este método, el PIB se calcula sumando los ingresos de todos los factores (trabajo y capital) que influyen en la producción. El ingreso sería el dinero o las ganancias que se reciben a través del salario, los arrendamientos, los intereses, etc.

Se trata de aplicar la metodología de pronósticos con restricciones a la serie PIB Colombia; serie que tiene una tendencia lineal y se necesita estacionarizar para poder aplicar la metodología de los pronósticos restringidos.

Se trabajará con datos en millones de pesos (Mp) trimestrales desde 2000 hasta 2009 con la *metodología año base 2000* a precios constantes.

Análisis PIB Metodología Año Base 2000: Se incluyen los datos trimestrales del PIB Colombia desde 2000 hasta 2009, que registra el Banco de la República de Colombia y el DANE, por lo que el tamaño de la muestra es de $N = 40$ datos u observaciones.

En esta ocasión se pretende estacionarizar la serie, generar los pronósticos irrestrictos para un intervalo de $H = 4$ periodos y con estos aplicar la metodología de los pronósticos restringidos para concluir con la verificación de compatibilidad entre la información adicional o restricción y la historia de la serie.

La importancia de este análisis también radica en comprobar si las decisiones de política económica referentes al crecimiento del PIB tomadas por el gobierno del Presidente Alvaro Uribe en cuanto al crecimiento económico propuesto para 2010 son compatibles con este estudio; lo cual se valida comparando los pronósticos obtenidos en este estudio con los datos que arroje el PIB en 2010 luego de verificado la compatibilidad entre las restricciones y la historia de la serie.

Análisis exploratorio de los datos: En primer lugar, la Figura 1 presenta el comportamiento de la serie PIB_t para el periodo comprendido entre el año 2000 y el 2009.

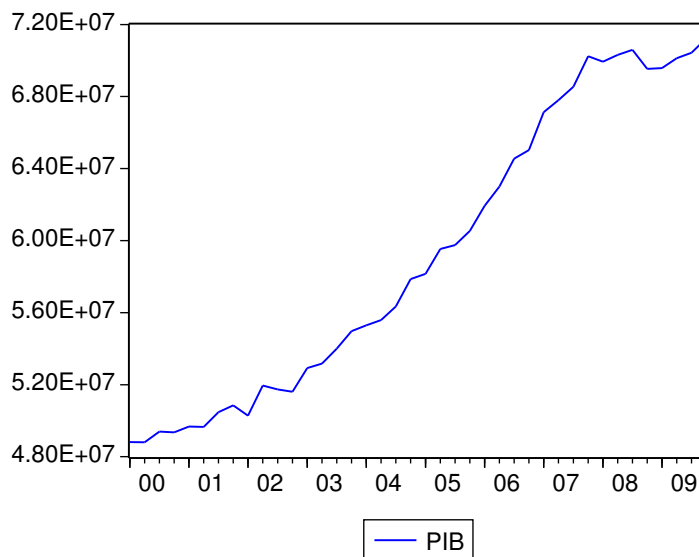


Figura 1: Serie PIB_t

La serie fue desestacionalizada por el banco de la República. Se observa que en hasta el tercer trimestre de 2001 hay una tendencia creciente; sin embargo para el ultimo trimestre

del mismo año, hay una caída en el crecimiento del PIB, lo cual cambia radicalmente para el siguiente trimestre; nuevamente en el segundo trimestre de 2002 el PIB está en descenso. Durante cada primer trimestre de 2003, 2004, 2005, 2006 y 2007 se observa un aumento considerable de los valores de la serie, teniendo en cuenta la naturaleza estacional de la misma, mientras que en los siguientes trimestres de estos años la pendiente de ascenso varía. Esta situación cambia durante el primer semestre del año 2007, cuando hay una disminución en el PIB; en contraste, el segundo semestre muestra un incremento. La curva se vuelve decreciente desde comienzos de 2008 hasta junio; Luego de esta variación, el PIB se muestra en aumento.

En la Figura 2, aparece la serie en *diferencias logarítmicas* estacionaria, esta serie esta doble diferenciada ya que la primera diferencia tenia bastantes indicios de la no estacionariedad en media.

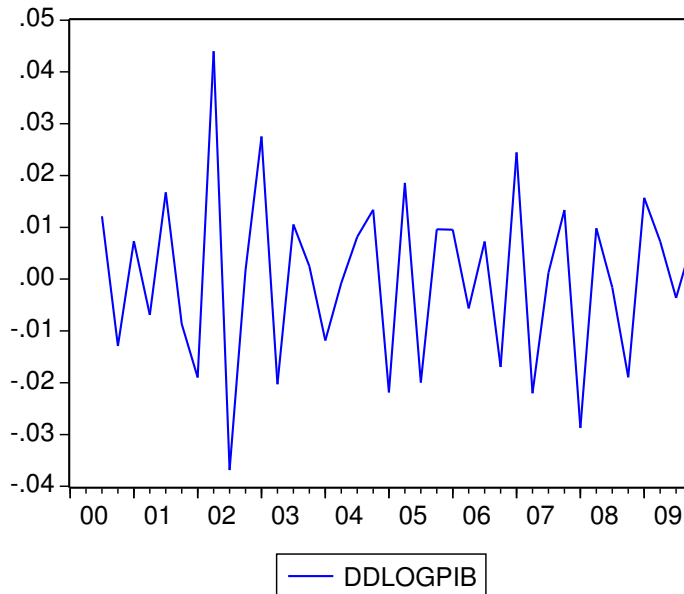


Figura 2: Serie $\nabla^2 \log\{PIB\}$

Un modelo ARIMA que explica el comportamiento histórico de la serie es

$$\nabla^2 \log\{PIB_t\} \quad \text{AR}(1) \quad \text{AR}(2) \quad \text{MA}(4)$$

Cabe aclarar que para verificar si la serie estudiada presenta raíces unitarias, se utiliza el test de Dickey Fuller Aumentado (ADF), en el cual se hace una regresión sobre la siguiente ecuación (asumiendo que Y_t es la variable a verificar y ε_t es una variable ruido blanco):

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \sum_{i=2}^p \beta \Delta Y_{t-i+1} + \alpha_2 t + \varepsilon_t$$

Los parámetros α_0 y α_2 aplican si hay en el proceso intercepto y tendencia determinística respectivamente; El contraste se realiza sobre el parámetro autorregresivo, α_1 . La hipótesis nula es existencia de raíz unitaria contra la alternativa de estacionariedad: $H_0 : \alpha_1 = 0$ vs. $H_1 : \alpha_1 < 0$. Para determinar

Al aplicar el contraste ADF a la serie PIB original se produce un valor del estadístico t de $-2,5899$, por lo que se acepta la hipótesis nula de no estacionariedad en media; al diferenciar una vez la serie, se obtiene un valor de $t = -1,8448$, con lo cual nuevamente se acepta H_0 . Finalmente, cuando se diferencia el proceso por segunda vez, se encuentra que el estadístico t es igual a $-11,1257$, lo cual permite el rechazo de la existencia de raíz unitaria. Así, la serie requiere ser diferenciada dos veces para que sea estacionaria en media.

Por simplicidad en la notación escribiremos: $PIB_t = Y_t$ y la transformación logarítmica se expresará como $\log\{PIB_t\} = TY_t$, obteniéndose el modelo

$$\nabla^2 TY_t = \phi_1 TY_{t-1} + \phi_2 TY_{t-2} + a_t - \theta_4 a_{t-4}$$

Aplicando el operador polinomial de retardos, se escribe

$$TY_t = \frac{1 - \theta_4 L^4}{1 - 2L + L^2 - \phi_1 L - \phi_2 L^2} a_t$$

igual a,

$$TY_t = \frac{1 - \theta_4 L^4}{1 - (2 + \phi_1)L - (\phi_2 - 1)L^2} a_t$$

Sea,

$$\Psi(L) = \frac{1 - \theta_4 L^4}{1 - (2 + \phi_1)L - (\phi_2 - 1)L^2} \quad \text{donde} \quad \Psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \dots$$

Entonces,

$$(1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots)[1 - (2 + \phi_1)L - (\phi_2 - 1)L^2] = 1 - \theta_4 L^4$$

Resolviendo la expresión anterior resulta:

$$\begin{aligned} \psi_1 &= 2 + \phi_1 \\ \psi_2 &= (\phi_2 - 1) + (2 + \phi_1)^2 \\ \psi_3 &= 2(\phi_2 - 1)(2 + \phi_1) + (2 + \phi_1)^3 \\ \psi_4 &= (\phi_2 - 1)^2 + 3(\phi_2 - 1)(2 + \phi_1)^2 + (2 + \phi_1)^4 - \theta_4 \\ \psi_5 &= (\phi_2 - 1)\psi_3 + (2 + \phi_1)\psi_4 \end{aligned}$$

Al estimar el modelo se obtuvieron los siguientes valores para los parámetros

$$\begin{aligned}\widehat{\phi}_1 &= -0,911294 \\ \widehat{\phi}_2 &= -0,479707 \\ \widehat{\theta}_4 &= -0,936804\end{aligned}$$

Con estos valores se obtuvieron las siguientes ponderaciones

$$\begin{aligned}\psi_0 &= 1 \\ \psi_1 &= 1,0887 \\ \psi_2 &= -0,2944 \\ \psi_3 &= -1,9315 \\ \psi_4 &= -0,7304\end{aligned}$$

i	0	1	2	3	4	5	6	7
ψ_i	1	1,0887	-0,2944	-1,9315	-0,7304	2,0629	3,3266	0,5693

Cuadro 1: Ponderaciones del modelo $\{\log(\text{PIB}_t)\}$

Inclusión de la restricción: El Banco de la República proyecta un crecimiento del PIB para 2010 en un rango de (2 a 4) %; así mismo la política económica del gobierno del presidente Uribe, proyecta para 2010 un crecimiento de 2,5 %, cifra con la que planteamos la siguiente restricción.

$$\frac{\text{PIB}_{T+4}}{\text{PIB}_{T+1}} = 1,025$$

Aplicando logaritmos, se obtiene

$$\begin{aligned}\log\{\text{PIB}_{T+4}\} - \log\{\text{PIB}_{T+1}\} &= \log\{1,025\}, \quad \text{igual a} \\ TY_{T+4} - TY_{T+1} &= \log\{1,025\} = Y\end{aligned}$$

Lo anterior se debe a que consideramos que el vector de valores futuros TY_F es de la forma

$$\begin{aligned}TY_F &= (TY_{T+1}, \dots, TY_{T+4}) \\ &= (\log\{\text{PIB}_{T+1}\}, \dots, \log\{\text{PIB}_{T+4}\})\end{aligned}$$

Dado que $Y = C * TY_F$, se deduce que el vector C tiene la siguiente representación

$$C = (-1, 0, 0, 1)$$

La matriz de ponderaciones Ψ para un horizonte de $H = 4$ periodos es

$$\Psi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1,0887 & 1 & 0 & 0 \\ -0,2944 & 1,0887 & 1 & 0 \\ -1,9315 & -0,2944 & 1,0887 & 1 \end{pmatrix} \quad (1)$$

De esta manera se obtiene el vector \widehat{A} cuya expresión es

$$\begin{aligned} \widehat{A} &= \Psi\Psi' C' (C\Psi\Psi' C')^{-1} \\ &= \begin{pmatrix} -0,2698 \\ -0,3208 \\ 0,1501 \\ 0,7302 \end{pmatrix} \end{aligned}$$

Y los pronósticos irrestrictos $E(TY_F|TY_0)$ donde TY_0 son las observaciones de la serie transformada desde $t = (1, 2, \dots, T)$ con $T = 40$, se obtienen de la siguiente forma

$$\widehat{TY}_F = E_t[TY_F|TY_0]$$

pero

$$\begin{aligned} \nabla^2 TY_F &= \phi_1 TY_{t-1} + \phi_2 TY_{t-2} + a_t - \theta_4 a_{t-4} \\ \nabla(TY_t - TY_{t-1}) &= \phi_1 LTY_t + \phi_2 L^2 TY_t + a_t - \theta_4 L^4 a_t \\ TY_t - TY_{t-1} - (TY_{t-1} - TY_{t-2}) &= \phi_1 LTY_t + \phi_2 L^2 TY_t + a_t - \theta_4 L^4 a_t \\ TY_t &= 2TY_{t-1} - TY_{t-2} + \phi_1 TY_{t-1} + \phi_2 TY_{t-2} + a_t - \theta_4 L^4 a_t \end{aligned}$$

resultando,

$$TY_t = (2 + \phi_1)TY_{t-1} + (\phi_2 - 1)TY_{t-2} + a_t - \theta_4 a_{t-4}$$

Entonces,

$$\begin{aligned} E_t[TY_{t+1}|TY_0] &= E_t[(2 + \phi_1)TY_t + (\phi_2 - 1)TY_{t-1} + a_{t+1} - \theta_4 a_{t-3}] \\ &= (2 + \phi_1)TY_t + (\phi_2 - 1)TY_{t-1} - \theta_4 a_{t-3} \\ E_t[TY_{t+2}|TY_0] &= E_t[(2 + \phi_1)TY_{t+1} + (\phi_2 - 1)TY_t + a_{t+2} - \theta_4 a_{t-2}] \\ &= (\phi_2 - 1)TY_t - \theta_4 a_{t-2} \\ E_t[TY_{t+3}|TY_0] &= E_t[(2 + \phi_1)TY_{t+2} + (\phi_2 - 1)TY_{t+1} + a_{t+3} - \theta_4 a_{t-1}] \\ &= -\theta_4 a_{t-1} \\ E_t[TY_{t+4}|TY_0] &= E_t[(2 + \phi_1)TY_{t+3} + (\phi_2 - 1)TY_{t+2} + a_{t+4} - \theta_4 a_t] \\ &= -\theta_4 a_t \end{aligned}$$

Por lo que resulta el vector de pronósticos irrestrictos:

$$E(TY_F|TY_0) = \begin{pmatrix} 18,0932 \\ 18,0962 \\ 18,1070 \\ 18,1163 \end{pmatrix}$$

Por lo que el pronóstico restringido está dado por

$$\widehat{TY}_F = E(TY_F|TY_0) + \widehat{A}[Y - CE(TY_F|TY_0)]$$

obteniendose

$$\widehat{TY}_F = \begin{pmatrix} 18,0921 \\ 18,0959 \\ 18,1071 \\ 18,1168 \end{pmatrix}$$

Dado que $\widehat{TY}_F = \log\{\widehat{PIB}_F\}$, deshaciendo esta transformación se obtiene $\widehat{Y}_F = \exp\{\log \widehat{PIB}_F\}$.

De esta manera,

$$\widehat{Y}_F = \widehat{PIB}_F = \begin{pmatrix} 71996000 \\ 72271000 \\ 73081000 \\ 73796000 \end{pmatrix}$$

En el cuadro 2 aparece el PIB real para 2010 hasta el tercer trimestre y los pronósticos irrestrictos para la misma serie, con sus correspondientes errores de pronóstico

Pronósticos irrestrictos de la serie PIB_t			
Período	PIB real	Error	Irrestrictos
2010q1	71955594.41	0.0756 %	72010000
2010q2	72723776.22	0.5978 %	72289000
2010q3	73112020.98	0.0533 %	73073000
2010q4			73755000

Cuadro 2: Pronósticos irrestrictos de PIB_t

De la misma manera se presentan los pronósticos restringidos para la serie PIB_t y los errores de pronóstico, cuadro 3:

Como se puede observar en el cuadro 2 y 3, los errores de pronósticos restringidos para el primer trimestre y el segundo son menores que los errores para los pronósticos irrestrictos, debido al uso de la información adicional en forma de metas económicas. Se espera que el error de pronóstico para el cuarto trimestre también sea menor en el caso de los restringidos.

Pronósticos restringidos de la serie PIB_t			
Período	PIB real	Error	Restringidos
2010q1	71955594.41	0.0561 %	71996000
2010q2	72723776.22	0.6225 %	72271000
2010q3	73112020.98	0.0424 %	73081000
2010q4			73796000

Cuadro 3: Pronósticos restringidos de la serie PIB_t

Compatibilidad entre la restricción $\{Y\}$ y la historia de la serie $\{\log(PIB_t)\}$:

La idea es comparar el estadístico K_{calc} con una ji-cuadrada con m grados de libertad, donde $m = 1$ ya que se introdujo una sola restricción.

Para la serie $TY_t = \log(PIB_t)$ se obtuvo un valor de $K_{\text{calc}} = 0,4548$, mientras que para un nivel de confianza de 95 % que es con el que usualmente se trabaja, se tiene que una $\chi_1^2 = 3,84$, por lo que se cumple que,

$$K_{\text{calc}} \leq \chi_1^2$$

lo cual indica compatibilidad entre la restricción y los datos históricos.

La Figura 3 muestra el comportamiento de la serie PIB real (en azul) hasta el cuarto trimestre de 2009 y los pronósticos restringidos (en rojo).

Se concluye entonces que con datos que cubrían desde el 2000 hasta el cuarto trimestre de 2009, la meta de crecimiento del PIB planteada por el gobierno del presidente Uribe, es razonable y valida ya que la restricción es compatible con la historia de la serie, por lo que es de esperarse una trayectoria parecida a la de los pronósticos restringidos en la figura 8.

Inclusión de la restricción del 3 % para el crecimiento del PIB: En líneas anteriores se ha hecho referencia al cambio de la meta de crecimiento por parte de las autoridades económicas, en este caso, los analistas económicos que representan al gobierno del Dr Álvaro Uribe Vélez. Tal cambio tuvo su origen en los buenos resultados de la economía en el primer trimestre del año 2010. En el cual el PIB se ubicó por encima del 4 %.

Por tal razón, se hace necesario verificar la viabilidad estadística de la nueva meta de crecimiento del 3 % para el crecimiento del PIB.

En primer lugar, se generan los pronósticos restringidos incluyendo la restricción del 3 %, luego se aplica el test de verificación de la compatibilidad entre la restricción y la historia de la serie.

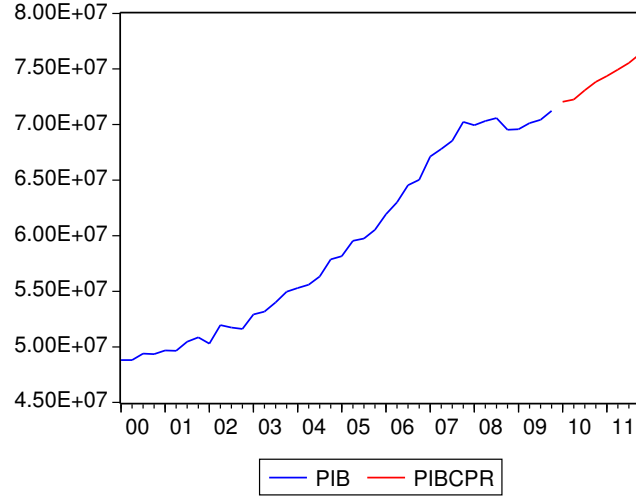


Figura 3: PIB (Datos observados y pronósticos restringidos)

Sea la restricción:

$$\frac{\text{PIB}_{T+4}}{\text{PIB}_{T+1}} = 1,030$$

entonces,

$$\log(\text{PIB}_{T+4}) - \log(\text{PIB}_{T+1}) = \log(1,030)$$

Por lo que,

$$Y = \log(1,030) \quad \text{y así,}$$

$$C = (-1, 0, 0, 1)$$

Sabemos que la matriz de ponderaciones Ψ y los pronósticos irrestrictos $E(TY_F|TY_0)$ del modelo son:

$$\Psi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1,0887 & 1 & 0 & 0 \\ -0,2944 & 1,0887 & 1 & 0 \\ -1,9315 & -0,2944 & 1,0887 & 1 \end{pmatrix} \quad E(TY_F|TY_0) = \begin{pmatrix} 18,0932 \\ 18,0962 \\ 18,1070 \\ 18,1163 \end{pmatrix}$$

De esta manera se obtiene el vector \hat{A} cuya expresión es

$$\hat{A} = \Psi\Psi' C'(C\Psi\Psi' C')^{-1}$$

$$= \begin{pmatrix} -0,2698 \\ -0,3208 \\ 0,1501 \\ 0,7302 \end{pmatrix}$$

De esta manera, se calcula los pronósticos restringidos mediante la expresión:

$$\widehat{TY}_F = E(TY_F|TY_0) + \widehat{A}[Y - CE(TY_F|TY_0)]$$

$$\widehat{TY}_F = \begin{pmatrix} 18,0915 \\ 18,0941 \\ 18,1080 \\ 18,1210 \end{pmatrix}$$

Al deshacer la transformación logarítmica, es decir, al hacer uso de la expresión $\exp(\widehat{TY}_F)$, resulta:

$$\widehat{Y}_F = \begin{pmatrix} 71948000 \\ 72141000 \\ 73146000 \\ 74107000 \end{pmatrix}$$

En el cuadro 4 aparecen los pronósticos restringidos para la serie (PIB_t), para el periodo 2010 y los correspondientes errores de pronóstico, según la meta del 3%.

Pronósticos y errores de pronóstico de la serie PIB_t			
Período	PIB Real	Error	Restringidos
2010q1	71955594.41	0.0105 %	71948000
2010q2	72723776.22	0.8013 %	72141000
2010q3	73112020.98	0.0464 %	73146000
2010q4			74107000

Cuadro 4: Pronósticos de PIB_t

Verificación de compatibilidad:

$$\widehat{\sigma}_a^2 = \frac{a_0 * a'_0}{T - K} = 0,00010498$$

entonces,

$$K_{cal} = \frac{[Y - CE(Y_F|Y_0)]'(C\Psi\Psi'C')^{-1}[Y - CE(Y_F|Y_0)]}{\widehat{\sigma}_a^2}$$

$$= 0,0366$$

Dado que se cumple la desigualdad $K_{cal} \leq \chi_1^2$, concluimos que la restricción es compatible con la historia de la serie, por lo que un crecimiento del 3% como lo pronosticó el gobierno es factible y tal valor se ubica en el centro de lo esperado por el Banco de la República, el cual estableció un rango meta entre 2% y 4% para el crecimiento del PIB.

Conclusiones

Luego de aplicar la metodología de pronósticos con restricciones al seguimiento del PIB en Colombia y luego de hacer una detallada revisión de los alcances de la misma, podemos plantear las siguientes conclusiones.

La meta del 2,5% y 3% fijadas por las autoridades económicas del gobierno de turno es compatible con la historia de la serie como se demuestra en la aplicación mediante el test de compatibilidad del estadístico K_{calc} .

Según los datos arrojados del PIB por el DANE hasta el tercer trimestre de 2010, y al compararlos con los pronósticos restringidos, se observa una buena aproximación ya que los errores de pronóstico no superan el 0,8%.

Se resalta la utilización de un modelo ARIMA, el cual se ajustó a la serie original en su parte histórica, explica el comportamiento de la serie y produjo buenos pronósticos irrestrictos; a dicho modelo se le incluyó la restricción (Meta de crecimiento) y se procedió a calcular los pronósticos restringidos.

Se obtuvieron resultados satisfactorios al aplicar la metodología de pronósticos restringidos (caso univariado) al seguimiento del PIB Colombiano, los pronósticos obtenidos se ajustan bastante bien a lo esperado para el PIB en el año 2010.

Referencias Bibliográficas

- [1] Amado, Peiró. Ezequiel, Uriel. *Introducción al análisis de series temporales*. Editorial AC. (2000) 1–56.
- [2] Anchuelo Álvaro. *Series integradas y cointegradas: una introducción*, Revista de Economía aplicada, Número 1, Volumen 1 1 (1993), 151–164.
- [3] Correa, Victor. Escandón, Antonio. Luengo, René. *Empalme de series anuales y trimestrales del PIB*. Departamento de Cuentas Nacionales, Banco Central de Chile. (2003).
- [4] Diaz Vernón, Javier. *Empalme series de PIB regionales 1960-2001, Base 1996*. División de planificación regional, Gobierno de Chile. (2004), 3–39.
- [5] Guerrero M, Victor. *Pronósticos restringidos con modelos de series de tiempo múltiples y su aplicación para evaluar metas de política macroeconómica en México*. Estudios Económicos, Redalyc. Vol. 22, número 002. (2007). 242–305.
- [6] Guerrero M, Victor. *Pronósticos con restricciones en series de tiempo univariadas: Aplicación al seguimiento del PIB de México en 2001*, Revista Mexicana de Economía y Finanzas 1 (2002), 15–38.
- [7] Guerrero M, Victor. *Análisis estadístico de series de tiempo económicas*. México: UAM-Iztapalapa (2003) 67–358.
- [8] Guerrero M, Victor. Pena, Bernardo. Serna, Eva. Alegria, Alejandro. *Restricted forecasting with a VEC Model: Validating the feasibility of economic targets*. ITAM, México.
- [9] Guisán, Carmen. *Causalidad y cointegración en modelos econométricos: Aplicaciones a los países de la OCDE y limitaciones de los tests de cointegración*. Econometrics Working Paper Series Economic Development n° 61. University of Santiago de Compostela, Faculty of Economics.
- [10] Lucas, Robert Jr. *Econometric policy evaluation: A critique*. Carnegie-Rochester Conference Series on Public Policy (1976), Elsevier, vol. 1(1), pages 19-46, January.
- [11] Lütkepohl, Helmut. Krätzig, Markus. *Applied time series econometrics*. Cambridge University Press. (2004).
- [12] Melo V, Luis Fernando. *Pronósticos condicionados para modelos VAR*. Estudios económicos del Banco de la República. (1996).

- [13] Misas A. Martha, Oliveros C. Hugo. *Cointegración, Exogeneidad y Crítica de Lucas: Funciones de demanda de dinero en Colombia: un ejercicio más*, Borradores de Economía del Banco de la República **1** (1997), 2–12.
- [14] Perdomo Strauch, Álvaro Andrés. *Inversión pública sectorial y crecimiento económico: Una aproximación desde la metodología VAR*. Departamento Nacional de Planeación, Archivos de Economía. 2002.
- [15] Kirchgässner, Gebhard y Wolters, Jürgen. *Introduction to Modern Time Series Analysis*. Springer, (2007), 165,171.
- [16] Mahadeva, Lavan y Robinson, Paul. *Prueba de Raíz unitaria para ayudar a construir un modelo*. Centro de Estudios monetarios latinoamericanos. <http://www.cemla.org/pdf/pub-en-76.pdf>. pag 30-36