



CEE

Centro de Estudios Económicos

www.colmex.mx

El Colegio de México, A.C.

Serie documentos de trabajo

**OBTENCIÓN DE PRONÓSTICOS ÓPTIMOS, SUJETOS A
RESTRICCIONES, CON MODELOS ARIMA**

Víctor M. Guerrero

DOCUMENTO DE TRABAJO

Núm. I - 1988

OBTENCION DE PRONOSTICOS OPTIMOS, SUJETOS A RESTRICCIONES,
CON MODELOS ARIMA

Víctor M. Guerrero*
Banco de México

RESUMEN

En este artículo se presenta un pronóstico aplicable a series de tiempo univariadas, que surge al combinar información histórica de la serie con información adicional acerca de su comportamiento futuro. Se derivan también algunos estadísticos y procedimientos de inferencia en relación con dicho pronóstico. De manera explícita se consideran dos situaciones básicas: (i) cuando la información adicional impone una restricción que debe ser satisfecha exactamente por el pronóstico y (ii) cuando la información adicional es sólo una conjetura acerca de los valores futuros de la serie, o bien un valor pronosticado a partir de un modelo alternativo.

PALABRAS CLAVE: Combinación de pronósticos; Error medio cuadrático mínimo; Minimización cuadrática; Modelos ARIMA; Series de Tiempo.

1. INTRODUCCION

Existen diversas situaciones en las cuales un analista de series de tiempo enfrenta el problema de restringir los pronósticos de un modelo ARIMA para una cierta variable, de tal manera que se satisfaga alguna restricción acerca de los valores futuros de dicha variable. Por ejemplo, Vera y Guerrero (1981) consideran el caso en que el saldo del financiamiento otorgado por el Sistema Bancario Mexicano debía alcanzar un valor determinado al final del año siguiente, pero los saldos mensuales para dicho año no estaban restringidos, excepto por el comportamiento histórico de la serie.

Cholette (1982) menciona otras situaciones y en particular señala que el caso de una restricción sobre el nivel anual de la serie, podría ser contemplada como el de un escenario o bien como un pronóstico que provenga de un mo-

* Víctor M. Guerrero es Jefe de Investigación Estadística en la Dirección de Investigación Económica del Banco de México, México DF, D.F.

delo econométrico. Aquí, de nuevo, se desea obtener la trayectoria (el pronóstico) de la variable para los meses o trimestres dentro del año dado. Por lo común, las restricciones acerca de los valores futuros de la variable se proporcionan como combinaciones lineales (es decir, son de la forma de promedios, diferencias entre dos valores específicos o bien valores aislados).

El enfoque adoptado en este artículo permite combinar las restricciones lineales, con los pronósticos obtenidos a partir de un modelo ARIMA. La fórmula del pronóstico que aquí se deduce, ha aparecido de manera implícita en el artículo de Doan, Litterman y Sims (1983) y en el de Cholette (1982), aunque apareció más explícitamente en un trabajo no-publicado de Pérez-Porrúa (1984); en este sentido, el presente artículo presenta una recopilación de metodologías. Sin embargo, en ninguno de los trabajos mencionados se enfatizan las propiedades estadísticas del pronóstico, ni se indica la posible adecuación del procedimiento asociado. Los resultados del presente artículo pretenden cubrir esta carencia y además se intenta reconciliar a las aparentemente distintas metodologías (particularmente las de Cholette y Pérez-Porrúa) dentro de un marco teórico unificado y con la notación de Box-Jenkins.

En el artículo de Doan, et.al.(1983) se obtiene un pronóstico condicionado mediante la aplicación de la Ley de Proyecciones Iteradas, dentro de un contexto multivariado bastante general en el cual se detalla la derivación matemática, pero se hace caso omiso de los aspectos estadísticos involucrados. Cholette (1982) utiliza la representación autorregresiva pura del modelo ARIMA y da solución a un problema de minimización cuadrática que proporciona un pronóstico condicionado, el cual podría pensarse como generalización del resultado de Doan, et.al.(1983) en el sentido de que la restricción impuesta pueda considerarse como una conjetura que no se requiere satisfacer exactamente. Por último Pérez-Porrúa (1984) resuelve otro problema general que aquí se demostrará es idéntico al planteado por Cholette, sólo que ahora Pérez-Porrúa emplea la representación de promedios móviles pura del modelo ARIMA y sigue un procedimiento similar al de Doan, et.al.

La formulación que ahora se presenta está más orientada hacia la interpretación estadística del resultado y hacia el uso de algunos estadísticos que

permitan juzgar la adecuación de los procedimientos sugeridos. La siguiente sección presenta el marco teórico dentro del cual se deriva el pronóstico condicionado óptimo (lineal, con error cuadrático medio mínimo). En la sección 3 se considera el problema de verificar la compatibilidad entre la restricción impuesta y la información histórica, para lo cual se deriva un estadístico de prueba. La sección 4 muestra un ejemplo teórico cuya finalidad es la de apreciar la naturaleza de los cálculos involucrados y los resultados que se obtienen. La generalización presentada por Pérez-Porrúa se justifica en la sección 5, en donde también se indica cómo elegir la matriz de varianza-covarianza involucrada y se demuestra la equivalencia formal entre los resultados de Pérez-Porrúa y de Cholette. En la sección 6 se muestran algunas aplicaciones que hacen uso de datos reales con fines ilustrativos exclusivamente.

2. DERIVACION DE LOS PRONÓSTICOS OPTIMOS CONDICIONADOS

Sea $\{Z_t\}$ una serie de tiempo observada durante el período $t=1, \dots, N$, para la cual se puede construir un modelo ARIMA. La representación de promedios móviles para dicha serie es

$$Z_t = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j} \quad (2.1)$$

en donde las variables a_t forman una sucesión de choques aleatorios con media cero y varianza σ_a^2 , y las ponderaciones ψ_j son constantes conocidas.

Box y Jenkins (1970) demuestran que el pronóstico óptimo (lineal con error cuadrático medio mínimo) de Z_{N+h} , dado el vector de observaciones $Z_0 = (Z_1, \dots, Z_N)'$ es

$$E(Z_{N+h} | Z_0) = \mu + \sum_{j=h}^{\infty} \psi_j a_{N+h-j}, \quad h = 1, \dots, H \quad (2.2)$$

cuyo error de pronóstico está dado por

$$Z_{N+h} - E(Z_{N+h} | Z_0) = \sum_{j=0}^{h-1} \psi_j a_{N+h-j} \quad (2.3)$$

En notación matricial se tiene

$$\underline{z}_F - E(\underline{z}_F | \underline{z}_0) = \underline{\psi} \underline{a}_F \quad (2.4)$$

con

$$\underline{z}_F = (z_{N+1}, \dots, z_{N+H})', \quad \underline{a}_F = (a_{N+1}, \dots, a_{N+H})'$$

y ψ es la matriz de dimensiones $H \times H$

$$\psi = \begin{bmatrix} \psi_0 & 0 & \dots & 0 \\ \psi_1 & \psi_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{H-1} & \psi_{H-2} & \dots & \psi_0 \end{bmatrix} \quad \text{con } \psi_0 = 1$$

en donde $E(\underline{a}_F | \underline{z}_0) = \underline{0}$ y $E(\underline{a}_F \underline{a}_F' | \underline{z}_0) = \sigma_a^2 \mathbf{I}$.

Supóngase que la información adicional está dada mediante la combinación lineal

$$\underline{y} = \underline{C} \underline{z}_F \quad (2.5)$$

con \underline{C} una matriz $m \times H$ ($m \leq H$) de constantes. Entonces, de acuerdo con (2.4) se sigue que

$$\underline{y} = \underline{C} E(\underline{z}_F | \underline{z}_0) + \underline{C} \underline{\psi} \underline{a}_F \quad (2.6)$$

2.1. El Pronóstico Optimo Sujeto a Restricciones

Para obtener un pronóstico lineal que incorpore la restricción (2.5) se requiere de la existencia de una matriz de constantes A , de dimensiones $H \times m$, tal que

$$\begin{aligned}\tilde{z}_F &= AY \\ &= AC[E(z_F|z_0) + \psi a_F]\end{aligned}\quad (2.7)$$

además, para que el error de pronóstico, dado z_0 , sea insesgado, se debe pedir que

$$\begin{aligned}0 &= E[(\tilde{z}_F - z_F)|z_0] \\ &= (AC - I)E(z_F|z_0)\end{aligned}\quad (2.8)$$

por lo cual

$$\tilde{z}_F - z_F = (AC - I)\psi a_F \quad (2.9)$$

y

$$\text{Cov}[(\tilde{z}_F - z_F)|z_0] = \sigma_a^2(AC\psi\psi'C'A' - AC\psi\psi' - \psi\psi'C'A' + \psi\psi') \quad (2.10)$$

por lo tanto, para que \tilde{z}_F tenga error cuadrático medio mínimo, A debe seleccionarse de forma tal que se minimice la varianza generalizada definida como

$$\text{Var}_F = \text{tr} \{ \text{Cov}[(\tilde{z}_F - z_F)|z_0] \} \quad (2.11)$$

entonces A se encuentra como solución de la ecuación $0 = (\partial \text{Var}_F / \partial A)|_{\hat{A}}$; esto es

$$\hat{A} = \psi\psi'C'(C\psi\psi'C')^{-1} \quad (2.12)$$

A partir de (2.7) y (2.8) se sigue que el pronóstico óptimo debe ser

$$\hat{z}_F = E(z_F | z_0) + \hat{A}C\psi_a' \quad (2.13)$$

en consecuencia, al hacer uso de (2.6) se obtiene

$$\begin{aligned} \hat{z}_F &= E(z_F | z_0) + \hat{A}[Y - CE(z_F | z_0)] \\ &= E(z_F | z_0) + \psi\psi'C'(C\psi\psi'C')^{-1}[Y - CE(z_F | z_0)] \end{aligned} \quad (2.14)$$

Por otro lado, para obtener bandas de confianza, dada la información z_0 y Y , se hará uso de la distribución del error de pronóstico

$$z_F - \hat{z}_F = (I - \hat{A}C)\psi_a' \quad (2.15)$$

que sigue una distribución normal con

$$E[(z_F - \hat{z}_F) | z_0, Y] = 0 \text{ y } \text{Cov}[(z_F - \hat{z}_F) | z_0, Y] = \sigma_a^2 \psi\psi'(I - \hat{A}C)' \quad (2.16)$$

Un hecho importante que debe resaltarse es que el error del pronóstico sin restricciones $E(z_F | z_0)$ tiene asociada la matriz de varianza-covarianza

$$\text{Cov}(\psi_a' | z_0) = \sigma_a^2 \psi\psi' \quad (2.17)$$

la cual excede a la matriz de (2.16) asociada con \hat{z}_F , por

$$\sigma_a^2 \psi\psi'C'\hat{A}' = \sigma_a^2 \psi\psi'C'(C\psi\psi'C')^{-1}C\psi\psi' \quad (2.18)$$

que es una matriz semidefinida positiva. Esto implica que los pronósticos condicionados a la información z_0 y Y serán más precisos que aquellos obtenidos al condicionar sólo a z_0 .

2.2 Enfoques Alternativos para Obtener el Pronóstico Óptimo

Un primer enfoque alternativo de derivación es mediante minimización cuadrática, para lo cual se considera el problema

$$\min_{\underline{z}_F} \{ \underline{a}_F' \underline{a}_F \} \quad \text{sujeto a } \underline{Y} = C \underline{z}_F \quad (2.19)$$

en donde, por (2.4), se sabe que

$$\underline{a}_F = \psi^{-1} [\underline{z}_F - E(\underline{z}_F | \underline{z}_0)]$$

y cuya solución resulta ser idéntica a (2.14).

Otro enfoque similar consiste en resolver el problema

$$\min_{E(\underline{a}_F | \underline{z}_0)} \{ E(\underline{a}_F | \underline{z}_0)' E(\underline{a}_F | \underline{z}_0) \} \quad \text{sujeto a } C \psi E(\underline{a}_F | \underline{z}_0) = \underline{Y} - C E(\underline{z}_F | \underline{z}_0) \quad (2.20)$$

el cual proporciona como resultado

$$\hat{E}(\underline{a}_F | \underline{z}_0, \underline{Y}) = \psi' C' (C \psi C')^{-1} [\underline{Y} - C E(\underline{z}_F | \underline{z}_0)] \quad (2.21)$$

que al sustituirse por \underline{a}_F en (2.4) conduce de nuevo a (2.14). En particular éste enfoque es mencionado por Doan, et.al. (1983).

3. PRUEBA ESTADÍSTICA DE LA HIPOTESIS $H_0: \underline{Y} = C \underline{z}_F$

Una vez que el modelo ARIMA para la serie $\{ \underline{z}_t \}$ ha sido construido, uno podría sustituir la matriz ψ por su estimación y, dada la información adicional \underline{Y} , se podría calcular entonces el pronóstico condicionado. Es de notar, sin embargo, que al usar la información provista por \underline{Y} uno presupone que la relación (2.5) es válida, en donde cada elemento del vector \underline{z}_F es generado por

el mismo proceso descrito por (2.1). Para decidir si Y puede ser considerado compatible con la información histórica, uno podría hacer uso de la prueba estadística sugerida a continuación.

Supóngase que la representación (2.1) es válida para $t=1, \dots, N+H$ y que cada a_t sigue una distribución normal, entonces se tiene

$$E(\underline{a}_F | \underline{z}_0) = \underline{0} \quad \text{y} \quad \text{Cov}(\underline{a}_F | \underline{z}_0) = \sigma_a^2 I \quad (3.1)$$

así pues, dado \underline{z}_0 , se sigue que

$$\begin{aligned} \underline{e}_Y &= \underline{Y} - C E(\underline{z}_F | \underline{z}_0) \\ &= C \underline{v}_F \sim N_m(\underline{0}, \sigma_a^2 C \psi \psi' C') \end{aligned} \quad (3.2)$$

de aquí, si σ_a^2 fuese conocida

$$K_1 = \underline{e}_Y' (C \psi \psi' C')^{-1} \underline{e}_Y / \sigma_a^2 \sim \chi_m^2 \quad (3.3)$$

pero, como regularmente ocurre en la práctica,

$$\hat{\sigma}_a^2 = \underline{a}_0' \underline{a}_0 / (N-k) \quad (3.4)$$

con $\underline{a}_0 = (a_1, \dots, a_N)'$ y k el número de parámetros en el modelo ARIMA, entonces, si la hipótesis nula es cierta, se obtiene

$$K_2 = \underline{e}_Y' (C \psi \psi' C')^{-1} \underline{e}_Y / (m \hat{\sigma}_a^2) \sim F_{m, N-k} \quad (3.5)$$

A partir de este resultado, la hipótesis nula de compatibilidad entre la restricción y la historia, o sea $H_0: Y = CZ_F$ puede ser llevada a cabo mediante el estadístico calculado

$$\hat{K}_2 = [\underline{Y} - C E(\underline{z}_F | \underline{z}_0)]' (C \psi \psi' C')^{-1} [\underline{Y} - C E(\underline{z}_F | \underline{z}_0)] / (m \hat{\sigma}_a^2) \quad (3.6)$$

al comparar su valor con tablas de una distribución F con un m y N-k grados de libertad, al nivel de significación deseado.

Debe notarse que el resultado (3.5) es únicamente una aproximación cuando ψ no es conocida, sino estimada a partir de los datos históricos, lo cual es el caso usual. Por este motivo, en lugar de \hat{K}_2 uno debería calcular el estadístico (3.3), en la forma

$$\hat{K}_1 = [Y - CE(Z_F | Z_0)]' (C \hat{\psi} \hat{\psi}' C')^{-1} [Y - CE(Z_F | Z_0)] / \hat{\sigma}_a^2 \quad (3.7)$$

y, como aproximación, referir este valor a tablas de la distribución Ji-cuadrada con m grados de libertad. El mismo argumento es utilizado por Box y Tiao (1976), quienes justifican el uso de un estadístico similar a (3.7), cuando N es grande. Además, Box y Tiao realizaron una breve investigación con el fin de observar la bondad de la aproximación utilizada cuando el modelo ARIMA es autorregresivo de orden p y encontraron que los errores de estimación elevan el valor de la χ^2 calculada de acuerdo con un factor cercano a $1+p/N$.

Si el estadístico calculado \hat{K}_1 (o bien \hat{K}_2) no muestra evidencia en contra de la hipótesis nula de que Y pudo haber sido generado por el proceso original, entonces se puede usar con validez la expresión (2.14). Por otro lado, si la hipótesis $H_0: Y = CZ_F$ es rechazada por la evidencia empírica, uno podría considerar la alternativa sugerida en la sección 5.

4. UNA ILUSTRACION TEORICA

Con el fin de comprender mejor la manera en que la fórmula (2.14) funciona, considérese el siguiente ejemplo teórico:

$$Z_t \sim \text{ARMA}(1,1) \text{ con } E(Z_t) = 0 \text{ y } C = (0,0,0,1)$$

Aquí se tiene entonces

$$\begin{aligned} Z_t &= \phi Z_{t-1} + a_t - \theta a_{t-1} \\ &= \sum_{j=1}^{\infty} \phi^{j-1} (\phi - \theta) a_{t-j} \end{aligned} \quad (4.1)$$

y por lo tanto, para $h=1, \dots, 4$

$$E(Z_{N+h} | Z_0) = \phi^{h-1} \hat{Z}_N(1) \quad \text{con} \quad \hat{Z}_N(1) = \phi Z_N - \theta a_N \quad (4.2)$$

así que

$$E(Z_F | Z_0) = \hat{Z}_N(1) (1, \phi, \phi^2, \phi^3)' \quad (4.3)$$

Considérese ahora la restricción $Y = Z_{N+4}$, de donde

$$Y - CE(Z_F | Z_0) = Y - \phi^3 \hat{Z}_N(1) \quad (4.4)$$

y

$$\hat{A} = \ell^{-1} (\phi^2 (\phi - \theta), \phi (\phi - \theta) [1 + \phi (\phi - \theta)], (\phi - \theta) [1 + \phi (\phi - \theta) + \phi^3 (\phi - \theta)], \ell)' \quad (4.5)$$

con

$$\ell = 1 + (\phi - \theta)^2 + \phi^2 (\phi - \theta)^2 + \phi^4 (\phi - \theta)^2 \quad (4.6)$$

por lo cual, el pronóstico sujeto a la restricción Y , dada la información histórica Z_0 , resulta ser

$$\hat{Z}_F = \lambda^{-1}$$

$$\left[\begin{array}{l} (\lambda - \phi^5(\phi - \theta)) \hat{Z}_N(1) + \phi^2(\phi - \theta) Y \\ (\lambda - \phi^3(\phi - \theta)[1 + \phi(\phi - \theta)]) \phi \hat{Z}_N(1) + \phi(\phi - \theta)[1 + \phi(\phi - \theta)] Y \end{array} \right] \quad (4.7)$$

$$(\lambda - \phi(\phi - \theta)[1 + \phi(\phi - \theta) + \phi^3(\phi - \theta)]) \phi^2 \hat{Z}_N(1) + (\phi - \theta)[1 + \phi(\phi - \theta) + \phi^3(\phi - \theta)] Y$$

λY

mientras que el estadístico calculado

$$\hat{K}_1 = \hat{K}_2 = [Y - \phi^3 \hat{Z}_N(1)]^2 / (\lambda \sigma_a^2) \quad (4.8)$$

debe ser comparado con tablas de χ_1^2 (o bien, se le debe comparar con tablas de una distribución F con 1 y N-2 grados de libertad) para verificar la adecuación del uso de (4.7).

Si uno asigna valores específicos a los parámetros ϕ y θ se pueden observar algunas propiedades interesantes del pronóstico, por ejemplo el efecto que la memoria del proceso tiene sobre los resultados; en particular si se considera una caminata aleatoria en logaritmos (o sea $\phi = 1$ y $\theta = 0$, con Z_t el logaritmo de la variable de interés) se obtendría el caso analizado por Vera y Guerrero (1981) para una serie con tasa de crecimiento aproximadamente constante.

5. UNA GENERALIZACION

El trabajo de Pérez-Porrúa (1984) se abocó a resolver el problema de la sección 2 con el supuesto adicional de que las restricciones impuestas pueden no ser ciertas por completo y convendría incluso denominarlas conjeturas. Así pues, Pérez-Porrúa deriva un pronóstico sujeto a la restricción

$$Y - CZ_F = u \quad (5.1)$$

en donde u es un vector aleatorio que satisface

$$u \sim N_H(0, U), \quad E(Z_0 u') = E(a_F u') = 0 \quad (5.2)$$

entonces, mediante una técnica similar a la empleada por Doan, et.al(1983), se obtiene la expresión (véase el Apéndice)

$$\begin{aligned} \hat{Z}_{F,U} &= E(Z_F | Z_0, Y) \\ &= E(Z_F | Z_0) + \hat{A}_U [Y - CE(Z_F | Z_0)] \end{aligned} \quad (5.3)$$

donde

$$\hat{A}_U = \psi \psi' C' (C \psi \psi' C' + U / \sigma_a^2)^{-1} \quad (5.4)$$

que son las expresiones generalizadas correspondientes a (2.14) y (2.12) respectivamente.

Dentro del contexto del presente artículo, la expresión (5.1) podría ser justificada empíricamente por la prueba descrita en la sección 3, ya que al rechazar la hipótesis nula $H_0: Y = CZ_F$, podría optarse por incluir el término aleatorio u y considerar así (5.1). Pueden existir también casos en que la inclusión de u se considere apropiada aún sin realizar la prueba estadística, e.g. cuando la información adicional Y proviene como pronóstico de un modelo

econométrico que usa datos anuales, mientras que la serie $\{Z_t\}$ consiste de datos mensuales; esta aplicación constituiría de hecho una combinación de pronósticos de dos modelos distintos (para información respecto a combinación de pronósticos véase por ejemplo, Clemen y Winkler, 1986).

5.1 Selección de la Matriz U

En la generalización que ahora se considera, se requiere especificar la matriz de varianza U que aparece en (5.2). Esto puede lograrse ya sea empleando la misma fuente externa de información que proveyó \underline{Y} (e.g. un modelo econométrico) o bien haciendo uso de la evidencia interna provista por los datos de la serie de tiempo. Para este último caso se sugiere emplear el método que sigue: ya que

$$\underline{e}_{Y,U} = \underline{Y} - CE(\underline{Z}_F | \underline{Z}_0) = C\psi \underline{a}_F + u \quad (5.5)$$

entonces $\underline{e}_{Y,U}$ se distribuye como una normal con

$$E(\underline{e}_{Y,U} | \underline{Z}_0) = \underline{0}, \text{Cov}(\underline{e}_{Y,U} | \underline{Z}_0) = \sigma_a^2 (C\psi\psi' C' + U) \quad (5.6)$$

y además

$$K_U = \underline{e}_{Y,U}' (C\psi\psi' C' + U / \sigma_a^2)^{-1} \underline{e}_{Y,U} / \sigma_a^2 \sim \chi_m^2 \quad (5.7)$$

por lo tanto, para N suficientemente grande, podría calcularse el estadístico

$$\hat{K}_U = [\underline{Y} - CE(\underline{Z}_F | \underline{Z}_0)]' (C\psi\psi' C' + U / \hat{\sigma}_a^2)^{-1} [\underline{Y} - CE(\underline{Z}_F | \underline{Z}_0)] / \hat{\sigma}_a^2 \quad (5.8)$$

el cual sirve para especificar la matriz U, puesto que si se desea:

$$\hat{K}_U < \chi_m^2(\alpha) \quad (5.9)$$

con $\chi_m^2(\alpha)$ el punto porcentual superior α de la distribución Ji-cuadrada con m grados de libertad, entonces la hipótesis

$$\underline{y} - C\underline{z}_F = \underline{u} \quad \text{con} \quad \underline{u} \sim N(\underline{0}, U) \quad (5.10)$$

no sería rechazada por los datos.

Ya que la matriz U ha sido elegida, ya sea mediante el uso del estadístico \hat{K}_U , o por algún otro método que se juzgue apropiado, uno puede obtener los pronósticos óptimos $\hat{\underline{z}}_{F,U}$ por medio de (5.3)-(5.4). Adicionalmente podrían hacerse inferencias acerca de tales pronósticos si se usa el hecho de que

$$\begin{aligned} \underline{z}_F - \hat{\underline{z}}_{F,U} &= (I - \hat{A}_U C)(\psi_{\underline{z}_F} + \underline{u}) \\ &\sim N(\underline{0}, \sigma_a^2 \psi \psi' (I - \hat{A}_U C)') \end{aligned} \quad (5.11)$$

5.2 El Enfoque de Minimización Cuadrática

A continuación se presenta una breve enumeración de los pasos seguidos por Cholette (1982) para obtener una fórmula que le permitiera combinar las restricciones con los pronósticos de un modelo ARIMA, además se demuestra que esta solución es equivalente a la anterior.

Supóngase que Z_t admite la representación autorregresiva

$$a_t = Z_t - \sum_{i=1}^{t-1} \pi_i Z_{t-i}, \quad t = 1, \dots, N \quad (5.12)$$

en donde π_1, \dots, π_{t-1} son constantes conocidas, entonces

$$Z_{N+h} = \sum_{i=1}^{N+h-1} \pi_i Z_{N+h-i} + a_{N+h} \quad (5.13)$$

así que, para $h = 1, \dots, H$ se tiene

$$E(Z_{N+h}|Z_0) = \pi_1 E(Z_{N+h-1}|Z_0) + \dots + \pi_{h-1} E(Z_{N+1}|Z_0) + \pi_h Z_N + \dots + \pi_{N+h-1} Z_1 \quad (5.14)$$

Por otro lado, de (2.4) se sabe que

$$a_F = \pi[Z_F - E(Z_F|Z_0)] \text{ con } \pi = \psi^{-1} \quad (5.15)$$

y además

$$\begin{aligned} \pi E(Z_F|Z_0) &= \begin{bmatrix} 1 & 0 & \dots & 0 \\ -\pi_1 & 1 & \dots & 0 \\ & \dots & & \\ -\pi_{H-1} & -\pi_{H-2} \dots & & 1 \end{bmatrix} \begin{bmatrix} E(Z_{N+1}|Z_0) \\ E(Z_{N+2}|Z_0) \\ \dots \\ E(Z_{N+H}|Z_0) \end{bmatrix} \\ &= \begin{bmatrix} \pi_N & \pi_{N-1} & \dots & \pi_1 \\ \pi_{N+1} & \pi_N & \dots & \pi_2 \\ \dots & \dots & \dots & \dots \\ \pi_{N+H-1} & \pi_{N+H-2} & \dots & \pi_H \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_N \end{bmatrix} \\ &= \pi_* Z_0 \end{aligned} \quad (5.16)$$

El problema planteado por Cholette fue

$$\min_{Z_F} \{ a'_F a_F + (Y - CZ_F)' G (Y - CZ_F) \} \quad (5.17)$$

en donde G es una matriz diagonal $m \times m$ de ponderaciones a ser elegida por el analista. Este problema puede ser expresado en términos de Z_F exclusivamente si se nota que

$$\begin{aligned} \hat{a}_{F,G} &= [Z_F - E(Z_F|Z_0)]' \pi' \pi [Z_F - E(Z_F|Z_0)] \\ &= (\pi Z_F - \pi_* Z_0)' (\pi Z_F - \pi_* Z_0) \end{aligned} \quad (5.18)$$

y así la solución resulta ser

$$\begin{aligned} \hat{Z}_{F,G} &= (\pi' \pi + C' G C)^{-1} (\pi' \pi_* Z_0 + C' G Y) \\ &= [(\pi' \pi + C' G C)^{-1} \pi' \pi_* \quad (\pi' \pi + C' G C)^{-1} C' G] \begin{bmatrix} Z_0 \\ Y \end{bmatrix} \end{aligned} \quad (5.19)$$

esta ecuación aparece como expresión (9) en el artículo de Cholette, sólo hay que reconocer que $P_1 = \pi$, $B = C$, $P_2 = \pi_*$ y $X_0 = Z_0$. Una aclaración es que Cholette restringe innecesariamente la suma que aparece en (5.13) hasta el sumando N -ésimo, hecho que conduce a que la matriz π_* de (5.16) tenga todos sus elementos debajo del término π_N en la diagonal, iguales a cero; esto podría afectar los resultados cuando aparezcan parámetros de promedios móviles en el modelo ARIMA, puesto que entonces no se haría uso de todos los datos disponibles.

Existen dos dificultades prácticas en lo que toca a la solución (5.19) que obtiene Cholette: primero, la manipulación computacional de la matriz π_* con dimensiones $(N+H) \times (N+H)$ cuando N es grande y segundo, la interpretación (y especificación) de la matriz de ponderaciones G . Para vencer estas dificultades, nótese que (5.19) se puede describir como

$$\begin{aligned} \hat{Z}_{F,G} &= (\pi' \pi + C' G C)^{-1} [\pi' \pi E(Z_F|Z_0) + C' G Y] \\ &= E(Z_F|Z_0) + \hat{A}_G [Y - C E(Z_F|Z_0)] \end{aligned} \quad (5.20)$$

con

$$\hat{A}_G = (\pi' \pi + C' G C)^{-1} C' G \quad (5.21)$$

la cual resulta ser de la forma (5.3). Adicionalmente, al manipular algebraicamente la ecuación (5.21) se obtiene lo siguiente

$$\begin{aligned} \hat{A}_G &= [I + (\pi' \pi)^{-1} C' G C]^{-1} (\pi' \pi)^{-1} C' G \\ &= (I + \psi \psi' C' G C)^{-1} \psi \psi' C' G \\ &= \{I - \psi \psi' C' G [(G^{-1} + C \psi \psi' C')]^{-1} C\} \psi \psi' C' G \\ &= \psi \psi' C' G - \psi \psi' C' (G^{-1} + C \psi \psi' C')^{-1} C \psi \psi' C' G \\ &= \psi \psi' C' G - \psi \psi' C' [I - (G^{-1} + C \psi \psi' C')^{-1} G^{-1}] G \\ &= \psi \psi' C' (C \psi \psi' C' + G^{-1})^{-1} \end{aligned} \quad (5.22)$$

la cual tiene la forma (5.4) con $G^{-1} = U/\sigma_a^2$

Por consiguiente, no es necesario trabajar con las matrices π y π' para obtener $\hat{\beta}_{F,G}$, sino única y exclusivamente la matriz $\psi \psi'$ de dimensiones $H \times H$. Además, ya que G es igual al inverso de la matriz de varianza-covarianza U/σ_a^2 , que aparece en (5.4), se obtiene entonces una manera de interpretar dicha matriz, y mediante los resultados de la sección 5.1 se tiene también un método empírico para especificarla en la práctica.

5.3 Un Ejemplo Teórico

Al igual que en la sección 4, a continuación se presenta una ilustración teórica cuyo fin es el de apreciar mejor los cálculos involucrados en la obtención del pronóstico. Considérese pues el caso en que

$$Z_t \sim \text{ARMA}(1,1), E(Z_t) = 0, C = (0,0,0,1) \text{ y } U = \sigma_u^2$$

entonces se tiene $\hat{Z}_N(1)$ igual que en (4.2), $\hat{A}_U = \hat{A} \ell (\ell + \sigma_u^2 / \sigma_a^2)^{-1}$ con \hat{A} y ℓ dados por (4.5)-(4.6) y

$$\hat{Z}_{F,U} = (\ell + \sigma_u^2 / \sigma_a^2)^{-1} \begin{bmatrix} [\ell + \sigma_u^2 / \sigma_a^2 - \phi^5(\phi - \theta)] \hat{Z}_N(1) + \phi^2(\phi - \theta) Y \\ (\ell + \sigma_u^2 / \sigma_a^2 - \phi^3(\phi - \theta) [1 + \phi(\phi - \theta)]) \phi \hat{Z}_N(1) + \phi(\phi - \theta) [1 + \phi(\phi - \theta)] Y \\ (\ell + \sigma_u^2 / \sigma_a^2 - \phi(\phi - \theta) [1 + \phi(\phi - \theta) + \phi^3(\phi - \theta)]) \phi^2 \hat{Z}_N(1) + (\phi - \theta) [1 + \phi(\phi - \theta) + \phi^3(\phi - \theta)] Y \\ (\sigma_u^2 / \sigma_a^2) \phi^3 \hat{Z}_N(1) + (1 - \sigma_u^2 / \sigma_a^2) Y \end{bmatrix} \quad (5.23)$$

En este caso el estadístico calculado es

$$\hat{K}_U = [Y - \phi^3 \hat{Z}_N(1)]^2 / (\ell \hat{\sigma}_a^2 + \sigma_u^2) \quad (5.24)$$

el cual se debe comparar con tablas de $\chi_1^2(\alpha)$ al nivel de significación deseado.

Conviene apreciar que el pronóstico (5.23) indica que la conjetura Y no se satisface exactamente a menos que $\sigma_u^2 = 0$, lo que conduciría de hecho al caso de la sección 2. Por otro lado, la expresión (5.9) conduce a elegir σ_u^2 de tal manera que se satisfaga la relación

$$\frac{[Y - \phi^3 \hat{Z}_N(1)]^2}{\chi_1^2(\alpha)} - \ell \hat{\sigma}_a^2 < \sigma_u^2 \quad (5.25)$$

6. ILUSTRACIONES EMPIRICAS

Con el fin de proporcionar una ilustración numérica de los procedimientos sugeridos, a continuación se considerará la serie de la tasa de crecimiento anual del Índice de Producción Manufacturera (IPM), denotada por $\{R_t^{IPM}\}$. El modelo ARIMA construido para esta serie, de acuerdo con la estrategia de Box-Jenkins, resultó ser

$$\nabla R_t^{IPM} = (1-\theta B)(1-\theta B^{12})a_t$$

en donde los parámetros estimados, con datos de enero de 1975 a septiembre de 1986 (N=129) son

$$\hat{\theta} = 0.1031, \hat{\theta} = 0.8111 \quad y \quad \hat{\sigma}_a^2 = 0.9313$$

por lo cual las ponderaciones ψ_1, ψ_2, \dots se estimaron como

$$\hat{\psi}_1 = \dots = \hat{\psi}_{11} = 0.896876, \quad \hat{\psi}_{12} = 0.085819, \quad \hat{\psi}_{13} = \hat{\psi}_{14} = 0.169458.$$

6.1 Pronósticos Sujetos a Una Restricción

Supóngase que existe interés en obtener un escenario mensual para 1987 de acuerdo con la restricción de que la tasa de crecimiento del IPM alcance el valor de 7% en diciembre de 1987. Aquí se tiene $H=15$, $m=1$, $C=(0,0,\dots,0,1)$ y $Y=7$. Es deseable aplicar el procedimiento de la sección 2 (llámese procedimiento A), a menos de que exista evidencia en contra de que la mencionada restricción es compatible con los datos históricos; por este motivo conviene realizar la prueba descrita en la sección 3. Al hacer esto se obtiene $K_1=K_2=5.781$, de tal manera que la restricción no es compatible con la historia, al nivel de significación del 5%. De aquí que convenga hacer uso del procedimiento de la sección 5 (procedimiento A_U).

Para poder especificar la matriz U de dimensión 1×1 , de hecho un escalar, que corresponda al procedimiento A_U , debe resolverse la desigualdad (5.9), o sea

$$[Y - \hat{R}_N^{IPM}(15)]^2 (\hat{\sigma}_a^2 \hat{C} \hat{\psi}' \hat{C}' + U)^{-1} < \chi_1^2(\alpha)$$

de la cual se obtiene

$$U > 49.7025 / \chi_1^2(\alpha) - 8.5978$$

así que, para $\alpha=0.05$, se debe elegir $U > 4.3456$.

Feb 197 -1.31

Feb 197 -2.04

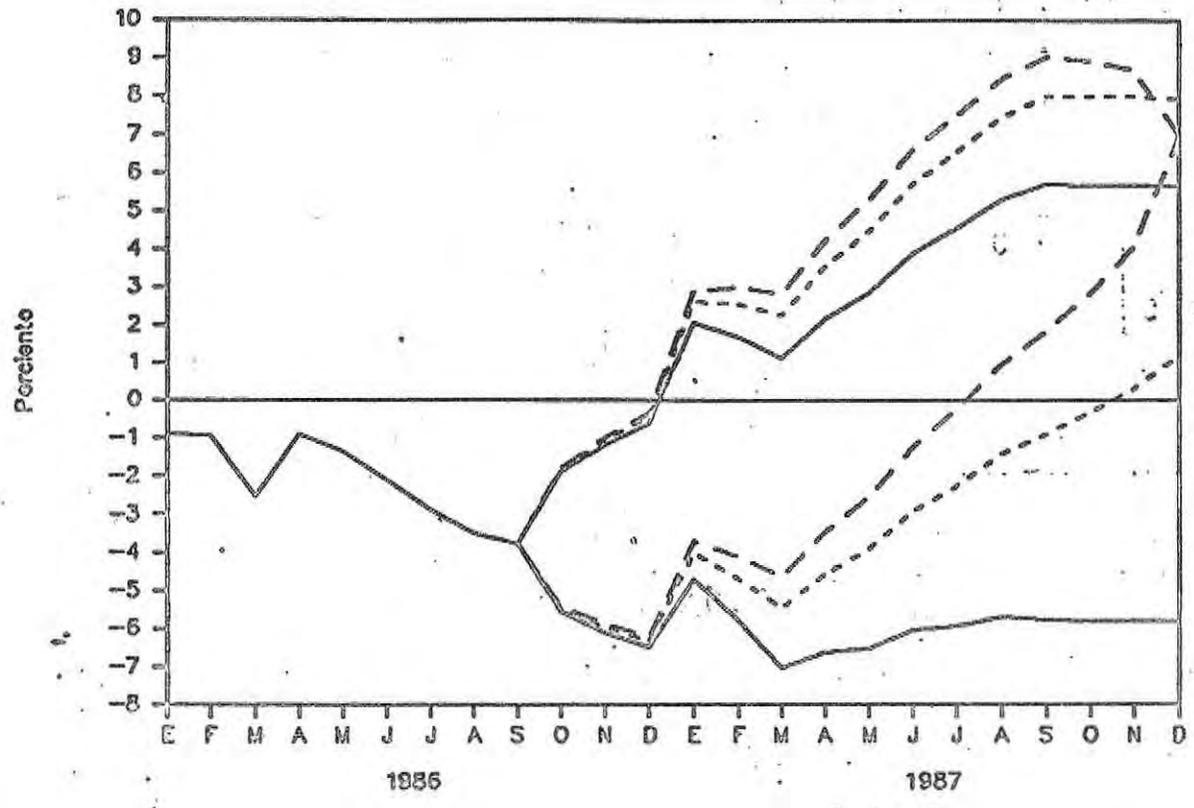
Los pronósticos ARIMA originales, así como los pronósticos que surgen de los procedimientos A y A_U (para el cual se asignó el valor $U=4.5$, de forma que $\hat{K}_U=3.795$) se muestran en el cuadro 1: En este cuadro se muestran también los errores estándar para los pronósticos respectivos.

Cuadro 1
Pronósticos para la Tasa de Crecimiento de IPM, con $\gamma=7$

Mes	ARIMA		Procedimiento A		Procedimiento A _U	
	Pronóstico	Error Estándar	Pronóstico	Error Estándar	Pronóstico	Error Estándar
Oct '86	-3.72	0.931	-3.60	0.930	-3.64	0.930
Nov	-3.66	1.251	-3.43	1.247	-3.51	1.249
Dic	-3.55	1.504	-3.27	1.500	-3.37	1.501
Ene '87	-1.31	1.721	-0.40	1.679	-0.71	1.693
Feb	-2.04	1.913	-0.56	1.811	-1.07	1.846
Mar	-2.94	2.087	-0.89	1.904	-1.59	1.969
Abr	-2.23	2.248	0.39	1.965	-0.51	2.067
May	-1.82	2.398	1.38	1.996	0.29	2.143
Jun	-1.07	2.539	2.70	1.998	1.40	2.199
Jul	-0.71	2.673	3.63	1.971	2.14	2.238
Ago	-0.19	2.801	4.72	1.915	3.04	2.259
Sep	-0.03	2.923	5.46	1.827	3.57	2.264
Oct	-0.05	2.924	5.91	1.551	3.86	2.125
Nov	-0.05	2.928	6.39	1.179	4.18	1.964
Dic	-0.05	2.932	7.00	0.000	4.58	1.719

El patrón de las bandas de confianza asociados con los pronósticos del cuadro 1 se puede apreciar visualmente en la gráfica 1. Aquí sobresale en particular el hecho de que la información adicional incrementa la precisión de los pronósticos, además se nota la diferencia entre considerar a la información adicional como una conjetura o como una restricción exacta (en este último caso la banda de confianza se cierra al final del horizonte de pronóstico).

TASA DE CRECIMIENTO DE IPM



Gráfica 1. Bandas del 95% de Confianza para los Pronósticos obtenidos con los Diferentes Métodos.

- ARIMA originales
- - - Procedimiento A
- · · Procedimiento Au

6.2 Otra Ilustración

Una situación ligeramente más complicada ocurre cuando se tiene más de una restricción. Por ejemplo, supóngase que la tasa de crecimiento del IPM debe alcanzar el valor 7% en diciembre de 1987 y el promedio de las tasas durante el año 1987 debe ser también del 3%. En este caso se tiene $H=15$, $m=2$,

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 1/12 & \dots & 1/12 & 1/12 \end{bmatrix}, Y = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$$

Los estadísticos calculados en este caso resultan ser $\hat{K}_1 = 5.871$ y $\hat{K}_2 = 2.891$, los cuales al compararse con $\chi_2^2(0.1)=4.61$ y $F_{2,127}(0.1)=2.35$ respectivamente, conducen al rechazo de la hipótesis de compatibilidad entre información adicional e histórica al nivel de significación $\alpha=0.1$. Por consiguiente conviene aplicar el procedimiento A_U .

La especificación de la matriz U de dimensiones 2×2 no es tan simple ahora como lo fue en el caso previo, no obstante, por ensayo y error, y en el supuesto de que $U = \sigma_U^2 \cdot I$, con I la matriz identidad, se encontró que el valor $\sigma_U^2 = 3.5$ era suficientemente grande como para satisfacer el criterio (5.9), ya que se obtuvo

$$\hat{K}_U = 4.606 < \chi_2^2(0.1)$$

y en consecuencia $U = 3.5 \cdot I$ se consideró adecuada. Los resultados obtenidos al aplicar los procedimientos A y A_U se resumen en el cuadro 2.

Cuadro 2

Pronóstico para la Tasa de Crecimiento de IPM con $\gamma=(7,3)^1$

Mes	ARIMA		Procedimiento A		Procedimiento A	
	Pronóstico	Error Estándar	Pronóstico	Error Estándar	Pronóstico	Error Estándar
Oct '86	-3.72	0.931	-3.59	0.850	-3.40	0.870
Nov	-3.66	1.251	-3.42	0.996	-3.03	1.061
Dic	-3.55	1.504	-3.25	0.915	-2.63	1.091
Ene '87	-1.31	1.721	-0.38	0.953	0.23	1.118
Feb	-2.04	1.913	-0.54	0.931	0.06	1.093
Mar	-2.94	2.087	-0.86	0.890	-0.31	1.042
Abr	-2.23	2.248	0.42	0.848	0.91	0.979
May	-1.82	2.398	1.41	0.818	1.80	0.918
Jun	-1.07	2.539	2.73	0.808	3.01	0.873
Jul	-0.71	2.673	3.66	0.821	3.80	0.855
Ago	-0.19	2.801	4.75	0.850	4.73	0.871
Sep	-0.03	2.923	5.48	0.883	5.28	0.919
Oct	-0.05	2.924	5.93	1.056	5.36	1.184
Nov	-0.05	2.928	6.40	0.995	5.45	1.331
Dic	-0.05	2.932	7.00	0.000	5.56	1.332
Prom '87	1.041	-	3.000	-	2.990	-

7. CONCLUSIONES

Se presentaron dos procedimientos estadísticos para obtener pronósticos sujetos a restricciones. Ambos procedimientos se derivan de un modelo ARIMA que se construye a partir de datos observados, pero la forma en que la restricción lineal se aplica: (i) como un hecho dado o (ii) como una conjetura, produce los diferentes resultados, cuya adecuación para la serie en estudio puede juzgarse por medio de una prueba estadística que aquí se describió.

Conviene apreciar que la generalización de la sección 5 proporciona una solución al problema de combinar pronósticos de dos modelos diferentes, uno de los cuales es un modelo ARIMA. Esto es más claro si se elige la matriz C como la matriz identidad de dimensión H en (5.1).

Finalmente, los cálculos necesarios para aplicar los procedimientos descritos no deben imponer limitantes, ya que la parte más complicada es obtener el valor numérico de la matriz $(C\Psi'C' + U/\sigma_a^2)^{-1}$ de dimensiones $m \times m$, en donde m es comúnmente un número pequeño. Por otra parte, las ilustraciones mostradas, tanto teórica como empíricas, permiten observar la utilidad potencial de los procedimientos.

AGRADECIMIENTOS

Se desea agradecer el interés mostrado por J.M. Pérez-Porrúa y J. Romo en este trabajo. Asimismo es de agradecer el apoyo de cómputo que brindo S. Portela.

APENDICE. SOLUCION AL PROBLEMA GENERALIZADO

En este apéndice se muestran los pasos básicos que siguió Pérez-Porrúa para derivar las expresiones (5.3) y (5.4). Este método es esencialmente el usado por Doan, et.al.(1983), pero la forma explícita de las expresiones obtenidas permite entender mejor el resultado. La notación empleada aquí es distinta de la que usó Pérez-Porrúa porque se trató de cuidar la consistencia de la notación usada en el presente artículo.

La idea principal es emplear la fórmula de la esperanza iterada

$$E(Z_F | Z_0, Y) = E(Z_F | Z_0) + E[(Z_F - E(Z_F | Z_0)) | Y - E(Y | Z_0)]$$

que se obtiene de la Ley de Proyecciones Iteradas. Entonces, de acuerdo con (5.1) y (2.4) se obtiene

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

así que

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

ahora bien, por el supuesto de normalidad se sabe que esta esperanza debe ser lineal, o sea

$$E(\psi_{a_F} | C\psi_{a_F} + u) = A_U(C\psi_{a_F} + u)$$

con A_U una matriz $H \times m$ elegida de tal forma que se satisfaga la condición de ortogonalidad

$$E\{(C\psi' + U)[a_F'\psi' - (a_F'\psi' C' + U')\hat{A}_U']\} = 0$$

lo cual conduce a la solución

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

como se indica en (5.4).

REFERENCIAS

- Box, G.E.P. y Jenkins, G.M. (1970) Time Series Analysis: Forecasting and control. San Francisco: Holden-Day
- Box, G.E.P. y Tiao, G.C. (1976) "Comparison of Forecast and Actuality", Applied Statistics 25, 195-200.
- Cholette, P.A. (1982) "Prior Information and ARIMA Forecasting", Journal of Forecasting 1, 375-383.
- Clemen, R.T. y Winkler, R.L. (1986) "Combining Economic Forecasts", Journal of Business and Economic Statistics 4, 39-46.
- Doan, T., Litterman, R., y Sims, Ch. (1983) "Forecasting and Conditional Projection using Realistic Prior Distributions", Working Paper 243, Research Department; Federal Reserve Bank of Minneapolis.
- Pérez-Porrúa, J.M. (1984) "Pronósticos ARIMA que Incorporan Información Externa". Unpublished Manuscript, Dirección de Investigación Económica, Banco de México.

Vera, G. y Guerrero, V.M. (1981) "Estimation of Seasonal Factors using Both Traditional Methods and Box-Jenkins Techniques", en Time Series Analysis, O.D. Anderson and M.R. Perryman eds., Amsterdam: North-Holland.