

01168
5
20J



Universidad Nacional Autónoma
de México

División de Estudios de Postgrado
FACULTAD DE INGENIERIA

PRONOSTICOS CON VECTORES AUTORREGRESIVOS;
APLICACIONES ECONOMICAS

T E S I S

Que para obtener el grado de:
MAESTRO EN INGENIERIA
(Investigación de Operaciones)

P R E S E N T A :

Gabriel Corral Alcala

México, D.F.

1993

TESIS CON
FALLA DE ORIGEN



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

INDICE

	Pag.
RESUMEN	I
INTRODUCCION	1
ANTECEDENTES	4
1. PROCESOS AUTORREGRESIVOS	5
1.1 ESTIMACION DE LOS PROCESOS AUTORREGRESIVOS	10
1.2 AUTOCORRELACION PARCIAL	15
1.3 EJEMPLO NUMERICO	19
2. VECTORES AUTORREGRESIVOS	25
2.1 CONCEPTOS BASICOS	29
2.2 ESTACIONARIEDAD	31
2.3 ESTIMACION DE VECTORES AUTORREGRESIVOS CON ORDEN CONOCIDO P	35
2.4 SELECCION DEL ORDEN DE UN VECTOR AUTORREGRESIVO	51
3. APLICACIONES ECONOMICAS	53
4. PRONOSTICO	59
5. CONCLUSIONES Y RECOMENDACIONES	66
6. REFERENCIAS	68

RESUMEN

Los vectores autorregresivos son una generalización de los procesos autorregresivos y constituyen toda una metodología que surge de la necesidad de describir la generación de cierta información económica mediante dos o más ecuaciones. El objetivo de este trabajo es establecer dicha metodología para pronosticar valores de un sistema de ecuaciones simultaneas dinámicas de carácter económico. Para llevar a cabo este fin se definen en primer lugar los procesos autorregresivos donde se analizan condiciones de estacionariedad, dando posteriormente el método de mínimos cuadrados para la estimación de estos procesos. La autocorrelación parcial es utilizada para la determinación del orden del proceso por lo que existe un apartado especial para este tema. Después de haber establecido la teoría de los procesos autorregresivos se definen los vectores autorregresivos y las condiciones necesarias para poderlos plantear, como son la estacionariedad y la transformación de la información a una forma reducida. Se hace un ejemplo considerando el orden de los vectores conocido para posteriormente establecer un método de obtención del orden cuando este sea desconocido para plantear una aplicación económica, la cual será utilizada como base para el tema de pronóstico.

INTRODUCCION

Uno de los principales aspectos del análisis de información económica es la necesidad de predecir valores futuros de variables de interés, por ejemplo, en la Bolsa de Valores, si se conociera el valor que una acción tendrá en un año, solucionaría todos los problemas del inversionista, porque sabría si debe invertir en ella o no. Sin embargo, es evidente que esto no es posible, pero si se conoce el comportamiento diario de la acción, digamos, de los dos últimos dos años, se podría hacer un pronóstico de su valor, con cierto grado de confianza, dando oportunidad para tomar alguna decisión; Las empresas se plantean anualmente metas de ventas, presupuestos, gastos, etc., donde generalmente, se utilizan métodos de pronóstico que son presentados a los consejos de administración; El gobierno, constantemente proporciona índices estimados a largo plazo de gran importancia para las diferentes actividades económicas del país, como son, Balanza Comercial para los dos próximos años (94 - 95), índice de inflación estimada para la próxima quincena, población económicamente activa a final de año, etc.

Una forma de obtener estos pronósticos es establecer un modelo paramétrico de la información disponible utilizando procedimientos estadísticos, suponiendo que la información es generada por un proceso estocástico.

Con frecuencia es indispensable determinar valores futuros de dos o más variables que tienen fuerte dependencia entre sí, por ejemplo, el consumo de un país en un año dado, digamos 1993, depende del consumo de 1992 y del ingreso obtenido en 1993, pero además, el ingreso de 1993 depende del ingreso y el consumo de 1992, la relación anterior entre consumo e ingreso se puede modelar con un sistema de dos ecuaciones con fuerte dependencia. Este trabajo se enfoca a esta clase de problemas, por lo tanto el objetivo es establecer la metodología de los vectores autorregresivos para pronosticar valores de un sistema de ecuaciones simultáneas dinámicas de carácter económico.

El desarrollo de este trabajo es como sigue: En el siguiente capítulo se establecen los antecedentes del uso Vectores Autorregresivos en pronósticos de indicadores económicos, así como los principales investigadores de este tema; Este trabajo trata sobre modelos multivariados, por lo que para llevar una secuencia lógica, en el capítulo 3 se definen los Procesos Autorregresivos para modelos univariados. En este mismo capítulo, se presenta un apartado sobre autocorrelaciones parciales, herramienta útil en la determinación del orden del modelo propuesto, terminando con un ejemplo numérico para analizar en detalle los procedimientos teóricos. El capítulo 4, es de hecho la parte medular de esta exposición, aquí se presentan los fundamentos teóricos de los Vectores Autorregresivos, analizando las condiciones de estacionariedad, además se

estima Vectores Autorregresivos para un orden p conocido y se presenta un ejemplo. Posteriormente, se analiza un método para seleccionar el orden del modelo cuando este sea desconocido. En el capítulo 5 es presentada una aplicación económica de acuerdo a lo tratado en la teoría de los capítulos anteriores. El capítulo 6 desarrolla la teoría de pronóstico, utilizando el ejemplo tratado en el capítulo 5 y se verá la aplicación para este tema. Finalmente, se presentará un último capítulo de conclusiones y recomendaciones.

ANTECEDENTES

Los vectores autorregresivos son una generalización de los procesos autorregresivos para 2 o más variables, donde se utilizan los valores observados del pasado para predecir valores futuros, sin embargo, los vectores autorregresivos van mucho más allá de una simple generalización, ya que su metodología incluye el plantear variables endógenas en términos de un conjunto de variables exógenas. Hay que notar que las variables exógenas juegan un papel muy importante en materia de política económica, como el tipo de cambio de una moneda dada, emisión de circulante, etc. Además, es importante mencionar que el modelado con vectores autorregresivos depende de muchas hipótesis que habrá que considerar.

El estudio de estos vectores es realmente reciente y hay varios estudios de este tema, pero a partir de 1987 se han incrementado, debidos principalmente a G. Judge y Helmut Lutkepohl (Judge, Lutkepohl 1985, 1987 [6], [7], [8]).

1. PROCESOS AUTORREGRESIVOS

Los procesos Autorregresivos representan la expresión más sencilla para determinar una relación de una variable aleatoria Y en el período T , con sus propios valores observados en períodos anteriores a T y ponderados de acuerdo a los coeficientes Autorregresivos (θ).

El proceso más elemental es:

$$Y_T = \theta Y_{T-1} + e_T .$$

La observación Y_T depende de la observación anterior (Y_{T-1}) más un error e_T , el cual se trata ampliamente más adelante. Por lo general, el proceso generador de una serie de tiempo $Y_1, Y_2, Y_3, \dots, Y_T$ será desconocido y puede ser tan complicado que el proceso que se acaba de establecer sería muy simple, por ejemplo, Y_T puede depender en una forma lineal no sólo de Y_{T-1} , sino también de Y_{T-2} y así sucesivamente.

En general, se puede escribir el proceso de la siguiente manera :

$$Y_T = \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p} + e_T ,$$

donde e_T es ruido blanco, esto es, $E[e_T]=0$ para toda T , $E[e_T e_S]=0$ para toda $T \neq S$ y $E[e_T^2]=\sigma^2$ para toda $T=S$. Esto último quiere decir que las covarianzas de los errores son cero y por lo tanto no están correlacionados, teniendo como consecuencia que no exista relación lineal alguna entre e_T y

es para toda $T \neq S$, y por otro lado, las varianzas son finitas e iguales en cualquier punto. Utilizando la notación de operador de retraso, este proceso lo podemos escribir como:

$$(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_P L^P) Y_T = e_T .$$

Un proceso estocástico de esta forma es llamado un Proceso Autorregresivo de orden P , cuya abreviación es $AR(P)$. Suponiendo que el proceso empieza en un pasado infinito y que Y_T tiene medias y varianzas acotadas, puede mostrarse que el proceso es estacionario si todas las raíces Z_0 del polinomio:

$$\theta_P(Z) = 1 - \theta_1 Z - \theta_2 Z^2 - \dots - \theta_P Z^P ,$$

tienen modulo $|Z_0| > 1$.

Por el momento se considera que un proceso es estacionario si tiene media y varianzas constantes.

Para que un proceso $AR(1)$ $Y_T = \theta Y_{T-1} + e_T$ sea estacionario, se requiere que $|\theta| < 1$.

A continuación se muestra esta aseveración:

Sea un proceso $AR(1)$: $Y_T = \theta Y_{T-1} + e_T$,

$$Y_T(1 - \theta L) = e_T ,$$

$$Y_T = (1 - \theta L)^{-1} e_T ,$$

$$\text{pero } (1 - \theta L)^{-1} = \sum_{i=0}^{\infty} (\theta L)^i ,$$

$$\text{así : } Y_T = (1 - \theta L)^{-1} e_T = \sum_{i=0}^{\infty} (\theta L)^i e_T ,$$

$$= e_T + \theta L e_T + \theta^2 L^2 e_T + \dots ,$$

$$= e_T + \theta e_{T-1} + \theta^2 e_{T-2} + \dots ,$$

Esta serie converge si $|\theta| < 1$ y diverge si $|\theta| \geq 1$.

NOTA: Esta condición de convergencia es debida al siguiente teorema:

TEOREMA.- Si una serie converge, entonces su n-ésimo término tiende a cero, cuando n crece indefinidamente.

Demostración.- Sea la serie $u_1 + u_2 + \dots + u_n + \dots$ convergente, es decir, tiene lugar la igualdad $\lim_{n \rightarrow \infty} S_n = S$, donde S es la suma de la serie (o sea un número finito fijo), pero entonces la igualdad $\lim_{n \rightarrow \infty} S_{n-1} = S$, puesto que cuando $n \rightarrow \infty$, también $(n-1) \rightarrow \infty$. Sustrayendo término a término la segunda igualdad de la primera, resulta:

$$\lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S_{n-1} = 0 ,$$

$$\lim_{n \rightarrow \infty} (S_n - S_{n-1}) = 0 ,$$

$$\text{pero } S_n - S_{n-1} = U_n ,$$

$$\text{Por lo tanto } \lim_{n \rightarrow \infty} U_n = 0 .$$

En este caso, se necesita que si la serie $\sum_{i=0}^n (\theta L)^i e_T$ converge, el límite del n-ésimo término cuando n tiende a infinito sea igual a cero:

$$\lim_{n \rightarrow \infty} (\theta)^i e_T = 0 \Rightarrow \lim_{n \rightarrow \infty} (\theta)^i = 0 , \text{ esto sucede si } \theta < 1 .$$

Por lo tanto , la serie efectivamente converge si $\theta < 1$.

Ahora, continuando con un proceso AR(1) se tiene que:

$$E[Y_T] = E[e_T] + \theta E[e_{T-1}] + \theta^2 E[e_{T-2}] + \dots = 0 .$$

$$\text{Por lo tanto, } E[Y_T] = 0 .$$

Se calcula ahora la varianza del proceso Y_T :

$$\text{Var}(Y_T) = \gamma_0 = \text{Var}[e_T] + \text{Var}[\theta e_{T-1}] + \text{Var}[\theta^2 e_{T-2}] + \dots ,$$

$$\text{Var}(Y_T) = \gamma_0 = \text{Var}[e_T] + \theta^2 \text{Var}[e_{T-1}] + \theta^4 \text{Var}[e_{T-2}] + \dots ,$$

$$\text{Var}(Y_T) = \gamma_0 = \sigma_a^2 (1 + \theta^2 + \theta^4 + \dots) = \sigma_a^2 / (1 - \theta^2) .$$

$$\text{Por lo tanto, } \text{Var}(Y_T) = \sigma_a^2 / (1 - \theta^2) .$$

Asimismo, las autocovarianzas pueden obtenerse a partir de la expresión:

$$\begin{aligned} \gamma_K &= E[(Y_T - \mu)(Y_{T+K} - \mu)] , \\ &= E[(e_T + \theta e_{T-1} + \theta^2 e_{T-2} + \dots)(e_{T+K} + \theta e_{T+K-1} + \dots)] , \\ &= E[\theta^k e_t^{2+\theta k} + \theta e_{t-1}^{2+\theta k} + \theta^2 e_{t-2}^{2+\theta k} + \dots] + E[\text{Prod. Cruzados de} \\ &\quad \theta^i e_{T-i} \text{ y } \theta^j e_{T-j}] , \\ &= \sigma_a^2 (\theta^k + \theta^k + \theta^k + \theta^k + \dots) , \\ &= \sigma_a^2 \left(\sum_{i=0}^{\infty} \theta^i \theta^k + 1 + \theta^k \right) , \\ &= \sigma_a^2 \theta^k \sum_{i=0}^{\infty} \theta^i , \quad k=1, 2, \dots , \end{aligned}$$

pero debido a que $|\theta| < 1$,

$$\gamma_K = \sigma_a^2 \theta^k / (1 - \theta^2) = \theta \gamma_{K-1} ,$$

ya que:

$$\gamma_{K-1} = \theta \sigma_a^2 \theta^{k-1} / (1 - \theta^2) = \sigma_a^2 \theta^k / (1 - \theta^2) ,$$

y debido a que $\gamma_K = \gamma_{-K}$ se obtiene en general:

$$\gamma_K = \sigma_a^2 \theta^{|k|} / (1 - \theta^2) , \text{ con } k = \dots, -2, -1, 0, 1, 2, \dots ,$$

de donde se sigue que las autocorrelaciones deben ser de la forma

$$\Gamma_k = \gamma_k / \gamma_0 = \theta^{|k|} .$$

Lo cual indica que, conforme $k > 0$ crece, la función de autocorrelación (FAC) tiende a cero, con decaimiento del tipo exponencial cuando $0 < \theta < 1$, y con signos alternados cuando $-1 < \theta < 0$, o lo que es lo mismo, la memoria del proceso muere en el tiempo. Por ejemplo, saber lo que pasó hace 30 ó 40 años previos no sirve mucho para predecir un valor, digamos anual.

En esta sección se planteó los procesos de grado 1, determinando su esperanza, varianzas, covarianzas y autocorrelación. En la siguiente sección se estimará los parámetros de estos modelos.

1.1 ESTIMACION DE LOS PROCESOS AUTORREGRESIVOS

Suponiendo que la muestra $Y_1, Y_2, Y_3, \dots, Y_T$ es generada por un proceso autorregresivo, es necesario encontrar el orden P y los valores de los parámetros $\theta_1, \theta_2, \dots, \theta_p$ que describen el proceso. Si se conoce el orden P , la ecuación $Y_T = \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p} + e_T$ sugiere como encontrar los valores $\theta_1, \theta_2, \dots, \theta_p$, ya que esta ecuación tiene precisamente la forma de un modelo estadístico lineal.

Hay que notar, sin embargo, que los regresores $Y_{T-1}, Y_{T-2}, \dots, Y_{T-p}$ son variables estocásticas. Si los e_T son ruido blanco Gaussiano, un e_T individual representa un choque aleatorio, el cual es sumado al sistema en el tiempo T y es independiente de las variables aleatorias en puntos previos. Más aun, como los coeficientes θ son parámetros desconocidos del modelo y aun conociendo sus valores, la combinación lineal, digamos $Y_T = \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p}$, no determinaría exactamente el valor Y_T , ya que por razones económicas Y_T como se había dicho, será considerada una variable estocástica y no exacta. Por lo tanto, e_T es una medida de discrepancia entre la combinación lineal y cualquier realización muestral de valores de Y . De esta manera, los regresores en una ecuación particular son independientes del error. Entonces, se estimará $\theta_p = (\theta_1, \theta_2, \dots, \theta_p)'$ por el método de mínimos cuadrados.

Remplazando las variables aleatorias por sus variables observadas se tiene:

$$Y_{p+1} = \theta_1 Y_p + \dots + \theta_p Y_1 + e_{p+1} ,$$

$$Y_{p+2} = \theta_1 Y_{p+1} + \dots + \theta_p Y_2 + e_{p+2} ,$$

.

.

.

$$Y_T = \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p} + e_T .$$

o, en notación matricial: $Y_p = X_p \theta_p + e$,

donde $Y_p = (Y_{p+1}, Y_{p+2}, \dots, Y_T)'$, $e = (e_{p+1}, e_{p+2}, \dots, e_T)'$, y

$$X_p = \begin{pmatrix} Y_p & Y_{p-1} & \dots & Y_1 \\ Y_{p+1} & Y_p & \dots & Y_2 \\ \vdots & \vdots & \dots & \vdots \\ Y_{T-1} & Y_{T-2} & \dots & Y_{T-p} \end{pmatrix} .$$

Sí e_T tiene una distribución Normal, Y_T "hereda" dicha distribución, y el estimador de θ_p por mínimos cuadrados se calcula de la siguiente manera:

Sea el modelo : $Y_p = X_p \theta_p + e$.

Sea θ_* , cualquier elemento k del vector θ_p , así:

$$e_* = Y_p - X_p \theta_* .$$

El método de mínimos cuadrados consiste en minimizar la suma de los cuadrados de los residuales $e_*' e_*$, esto es, minimizar:

$$\begin{aligned}
e_*'e_* &= (Y_P - X_P\theta_*)'(Y_P - X_P\theta_*) , \\
&= Y_P'Y_P - 2\theta_*'X_P'Y_P + \theta_*'X_P'X_P\theta_* .
\end{aligned}$$

ahora, $\frac{\partial(e_*'e_*)}{\partial\theta_P} = -2X_P'Y_P + 2X_P'X_P\theta_* \dots(3.1) .$

La condición necesaria para encontrar los puntos estacionarios que se buscan, es que (3.1) sea igual al vector 0. Así,

$$-2X_P'Y_P + 2X_P'X_P\theta_* = 0 ,$$

$$X_P'X_P\theta_* = X_P'Y_P ,$$

Por lo tanto, $\hat{\theta}_P = (X_P'X_P)^{-1}X_P'Y_P .$

El siguiente objetivo es determinar la distribución de θ_P .

Se tiene que :

$$\hat{\theta}_P = (X_P'X_P)^{-1}X_P'Y_P \dots(3.2) .$$

Donde X_P es una matriz de regresores "estocásticos".

Para calcular la esperanza de θ_P se sustituirá

$$Y_P = X_P\theta_P + e \text{ en (3.2),}$$

$$\text{Así, } \hat{\theta}_P = (X_P'X_P)^{-1}X_P'X_P\theta_P + (X_P'X_P)^{-1}X_P'e ,$$

$$\hat{\theta}_P = \theta_P + (X_P'X_P)^{-1}X_P'e .$$

Hay que notar que e es el vector error aleatorio, el cual se distribuye independientemente de la matriz X_P estocástica. Esta independencia de e y X_P implica que:

$$E[\mathbf{e}|\mathbf{X}_P] = E[\mathbf{e}] = \mathbf{0} ,$$

Y

$$E[\mathbf{e}\mathbf{e}'|\mathbf{X}_P] = E[\mathbf{e}\mathbf{e}'] = \sigma^2\mathbf{I}_T ,$$

ahora,

$$E[\hat{\theta}_P] = E[\theta_P] + E[(\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P\mathbf{e}] ,$$

por independencia de \mathbf{e} y \mathbf{X}_P se tiene que:

$$E[\hat{\theta}_P] = \theta_P + E[(\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P]E[\mathbf{e}] ,$$

por lo tanto,

$$E[\hat{\theta}_P] = \theta_P .$$

Este resultado implica que el estimador por mínimos cuadrados es insesgado

La Varianza de $\hat{\theta}_P$ se estima de la siguiente manera:

$$E[(\hat{\theta}_P - \theta_P)(\hat{\theta}_P - \theta_P)'] = E[(\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P'\mathbf{e}\mathbf{e}'\mathbf{X}_P(\mathbf{X}_P'\mathbf{X}_P)^{-1}] ,$$

pero

$$\hat{\theta}_P = \theta_P + (\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P\mathbf{e} ,$$

así,

$$\hat{\theta}_P - \theta_P = (\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P\mathbf{e} .$$

$$\begin{aligned} E[(\hat{\theta}_P - \theta_P)(\hat{\theta}_P - \theta_P)'] &= E[E[(\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P'\mathbf{e}\mathbf{e}'\mathbf{X}_P(\mathbf{X}_P'\mathbf{X}_P)^{-1}|\mathbf{X}_P]] , \\ &= E[(\mathbf{X}_P'\mathbf{X}_P)^{-1}\mathbf{X}_P'\sigma_e^2\mathbf{I}_T\mathbf{X}_P(\mathbf{X}_P'\mathbf{X}_P)^{-1}] , \\ &= \sigma_e^2E[(\mathbf{X}_P'\mathbf{X}_P)^{-1}] . \end{aligned}$$

Para la obtención de la varianza se utilizó la siguiente regla de esperanza matemática: si x y y son dos variables aleatorias, entonces la esperanza de x es igual a la esperanza de la esperanza condicional de x dado y , como $E[x] = E_y[E[x|y]]$.

El parámetro escalar σ_e^2 puede ser estimado por:

$$\begin{aligned}\hat{\sigma}_e^2 &= \frac{\mathbf{e}'\mathbf{e}}{(T-2P)}, \\ &= \frac{(\mathbf{Y}_P - \mathbf{X}_P \hat{\theta}_P)' (\mathbf{Y}_P - \mathbf{X}_P \hat{\theta}_P)}{T - 2P}.\end{aligned}$$

Finalmente, la matriz de covarianzas $\Sigma_{\hat{\theta}_P}$ del estimador $\hat{\theta}_P$ obtenido por mínimos cuadrados puede ser estimado por:

$$\Sigma_{\hat{\theta}_P} = \hat{\sigma}_e^2 / (\mathbf{X}_P' \mathbf{X}_P)^{-1}.$$

Hasta aquí se ha determinado los parámetros para un proceso con orden P conocido. A continuación se analizará las autocovarianzas parciales que serán utilizadas para determinar el orden P cuando este sea desconocido.

1.2 AUTOCORRELACIONES PARCIALES

Una manera de identificar el orden adecuado de un proceso autorregresivo para un conjunto de datos dados, es estimar procesos en orden ascendente K y probar la significancia de θ_K . Este coeficiente es llamado el k -ésimo coeficiente de autocorrelación parcial y será denotado por θ_{KK} , ya que es el k -ésimo coeficiente de un proceso autorregresivo de orden k . El coeficiente de autocorrelación parcial mide la correlación entre los residuales de cada variable después de que el efecto lineal de todas las demás variables explicatorias ha sido removido.

Sea X_2 una variable explicatoria que influye en Y , y X_3 también una variable explicatoria que influye en Y . El coeficiente de correlación entre Y y X_2 es por definición:

$$r_{12} = \frac{\sum Y X_2}{\sqrt{\sum Y^2} \sqrt{\sum X_2^2}}$$

Parte de la variación en Y es realmente debida a X_3 , y así r_{12} no mediría correctamente la correlación atribuida a X_2 . Por lo que, habrá que eliminar el efecto de X_3 sobre Y para obtener la influencia exacta de X_2 en Y . El coeficiente de correlación parcial

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2} \sqrt{1-r_{23}^2}}$$

Mide la correlación entre Y y X_2 después de que el efecto lineal de X_3 ha sido removido tanto de X_2 como de Y . La notación $r_{12.3}$, quiere decir coeficiente de correlación parcial entre la primera variable, en este caso Y , y la segunda variable X_2 , sin considerar el efecto lineal de la tercera variable X_3 .

Ahora, introducido ya el concepto de autocorrelación parcial, se aplicará este concepto a un proceso AR(1) : $Y_T = \theta Y_{T-1} + e_T$ para el cual, se sabe que $\Gamma_k = \theta^{|k|}$ con $k = \dots, -2, -1, 0, 1, 2, \dots$. Supongase que existe el interes de cuantificar la dependencia entre Y_T (1) y Y_{T-2} (2) sin tener en cuenta a Y_{T-1} (3). Así, $r_{12} = r_2 = \theta^2$ y $r_{13} = r_{32} = r_3 = \theta$, por lo tanto $r_{12.3} = 0$. De hecho las autocorrelaciones parciales que excluyan a Y_{T-1} son cero, como era de esperarse, pues la única variable 'independiente' que aparece en el modelo AR(1) es Y_{T-1} .

De la misma manera, considerese el proceso AR(2) : $Y_T = \theta_1 Y_{T-1} + \theta_2 Y_{T-2} + e_T$, en el cual intervienen Y_{T-1} y Y_{T-2} como variables 'independientes' y para las cuales:

$$r_1 = \frac{\theta_1}{1-\theta_2} \quad \text{y} \quad r_2 = \frac{\theta_2 + \theta_1^2}{1-\theta_2} .$$

$$\text{Así : } r_{12.3} = \frac{r_2 - r_1^2}{1 - r_1^2} = \theta_2 ,$$

es decir, la contribución de Y_{T-2} para explicar Y_T se mide a través de θ_2 .

El objetivo es determinar el orden P del proceso autorregresivo y se obtendrá de la siguiente manera:

Se calculará en forma sucesiva los coeficientes $\theta_{11}, \theta_{22}, \dots, \theta_{pp}$ y el orden del proceso autorregresivo será igual a P , si las autocorrelaciones parciales son distintas de cero para un proceso de orden P . Para analizar más claramente esta aseveración, ya se verificó que para un proceso $AR(1)$: $Z_T = \theta Z_{T-1} + e_T$, la única variable independiente es Z_{T-1} , por lo tanto, el coeficiente de correlación de primer orden de un proceso $AR(1)$ es distinto de cero, mientras cualquier correlación parcial de mayor orden es igual a cero. Si el proceso es $AR(2)$: $Z_T = \theta_1 Z_{T-1} + \theta_2 Z_{T-2} + e_T$, las únicas variables independientes son Z_{T-1} y Z_{T-2} , por lo que cualquier autocorrelación parcial de orden 2 de un proceso $AR(2)$ es distinto de cero y cualquier autocorrelación parcial de mayor orden es igual a cero. Así que, cuando una autocorrelación parcial de orden $P+1$ sea igual a cero para un proceso $AR(P+1)$, y siendo la autocorrelación parcial de orden P distinta de cero para un proceso $AR(P)$, implica que el orden del proceso autorregresivo es P .

La secuencia de θ_{kk} para $k=1,2,\dots$, de autocorrelaciones parciales será llamada la función de autocorrelación parcial. Como un modelo para el proceso que genera los datos, se escogerá un proceso $AR(P)$ tal que:

$$\theta_{KK} \begin{cases} \neq 0 & \text{para } k=p \\ = 0 & \text{para } k>p \end{cases}$$

Para probar la significancia de θ_{KK} se necesita saber la distribución de su estimador. El correspondiente estimador $\hat{\theta}_{KK}$ se obtiene de la última coordenada de

$$\hat{\theta}_K = (X_K' X_K)^{-1} X_K' Y_K ,$$

donde se supone que la media muestral es cero. Se puede mostrar que para muestras de tamaño grande, si el orden del proceso autorregresivo es de hecho P , las correlaciones parciales $\hat{\theta}_{KK}$ estimadas tienen una distribución aproximadamente Normal con media cero y varianza $1/T$ para $K > P$, donde T es, como siempre, el tamaño de la muestra. En consecuencia, para verificar la significancia de $\hat{\theta}_{KK}$ al 95% aproximadamente, se tomará el intervalo de confianza:

$$(\hat{\theta}_{KK} - 2/T, \hat{\theta}_{KK} + 2/T) .$$

1.3 EJEMPLO NUMERICO.

Se considera la muestra del Cuadro No.1 como base para este análisis. La tendencia de estos datos se muestran en la Gráfica No.1.

El primer objetivo será determinar el orden P del proceso autorregresivo que se propondrá. Después de haber determinado el orden, se tendrá que estimar los parámetros del proceso autorregresivo propuesto. En esta sección se determinó la manera de obtener el orden P por medio de autocorrelaciones parciales, lo interesante del caso es que en el proceso mediante el que se obtienen, se van estimando al mismo tiempo los parámetros del proceso autorregresivo. A continuación se verá el cálculo de las 3 primeras autocorrelaciones parciales y se determinará el proceso autorregresivo propuesto, así como la estimación de los parámetros.

La primera consideración que se debe tomar en cuenta, es que el estimador de las autocorrelaciones parciales es:

$$\hat{\theta}_K = (X_K' X_K)^{-1} X_K' Y_K ,$$

donde $Y_K = (Y_{K+1}, Y_{K+2}, \dots, Y_T)'$ y

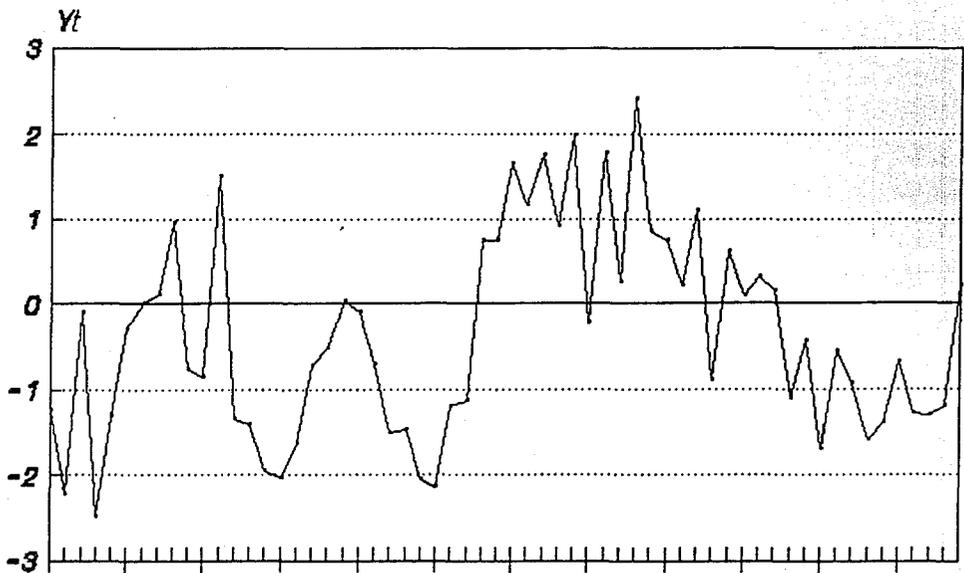
$$X_K = \begin{bmatrix} Y_K & Y_{K-1} & \dots & Y_1 \\ Y_{K+1} & Y_K & \dots & Y_2 \\ \vdots & \vdots & \ddots & \vdots \\ Y_{T-1} & Y_{T-2} & \dots & Y_{T-P} \end{bmatrix} .$$

CUADRO NO.1

I	YI	I	YI	I	YI	I	YI
1	-1.245203	26	-2.155142	51	-1.701559	76	-0.448346
2	-2.229019	27	-1.205098	52	-0.557646	77	-1.407836
3	-0.102889	28	-1.125508	53	-0.935374	78	-1.416262
4	-2.504927	29	0.745601	54	-1.590731	79	-2.654282
5	-1.295782	30	0.745816	55	-1.392477	80	-1.568367
6	-0.28721	31	1.652314	56	-0.702117	81	-1.182442
7	-0.000095	32	1.170398	57	-1.279344	82	0.377995
8	0.098003	33	1.748881	58	-1.310539	83	-1.028879
9	0.946282	34	0.915977	59	-1.206204	84	-1.402405
10	-0.775019	35	1.988901	60	0.213225	85	-0.865423
11	-0.868487	36	-0.225835	61	-0.61447	86	0.314383
12	1.500287	37	1.772048	62	0.430978	87	-0.569566
13	-1.351182	38	0.256119	63	1.281028	88	0.834981
14	-1.408163	39	2.398788	64	1.625546	89	1.322164
15	-1.945642	40	0.846766	65	2.695052	90	1.643056
16	-2.029809	41	0.736914	66	1.794567	91	1.302254
17	-1.642363	42	0.219055	67	1.082367	92	1.794244
18	-0.742933	43	1.099941	68	2.351032	93	2.026379
19	-0.519426	44	-0.906636	69	1.586859	94	0.345364
20	0.0308898	45	0.620321	70	2.131216	95	-0.614447
21	-0.109768	46	0.079424	71	0.253485	96	0.979466
22	-0.720097	47	0.32415	72	-0.510527	97	2.082886
23	-1.523058	48	0.146697	73	-0.909131	98	2.011572
24	-1.481303	49	-1.114253	74	-1.097981	99	0.422207
25	-2.070795	50	-0.43741	75	-1.305027	100	1.239364

FUENTE : JUDGE, HILL, & LUTKEPOHL.
 INTRODUCTION TO THEORY & PRACTICE OF ECONOMETRICS.
 PAG. 686.
 ADDISON WELY 1987.

Grafica No.1



Así, para K=1

Se tiene que $Y_K = (Y_{K+1} = -2.229019, Y_{K+2} = -0.0102889, \dots, Y_T = 1.239364)$ ' y

$$X_K = \begin{pmatrix} Y_K & = & -1.245203 \\ Y_{K+1} & = & -2.229019 \\ \vdots & & \vdots \\ Y_{T-1} & = & 0.422207 \end{pmatrix} .$$

de estos datos se obtiene:

$$(X_K'X_K) = 169.0401, \quad (X_K'X_K)^{-1} = .005916 \text{ y } X_K'Y_K = 112.4409.$$

$$\text{Por lo tanto, } \hat{\theta}_K = (X_K'X_K)^{-1}X_K'Y_K = 0.665173 .$$

Para K=2

Se tiene que $Y_K = (Y_{K+1} = -0.102889, Y_{K+2} = -2.5044927, \dots, Y_T = 1.239364)$ ' y

$$X_K = \begin{pmatrix} Y_K & = & -2.229019 & & Y_{K-1} & = & -1.245203 \\ Y_{K+1} & = & -0.102889 & & Y_K & = & -2.229019 \\ \vdots & & \vdots & & \vdots & & \vdots \\ Y_{T-1} & = & 0.422207 & & Y_{T-2} & = & 2.011572 \end{pmatrix} ,$$

$$(X_K'X_K) = \begin{pmatrix} .010552 & 0.006980 \\ .006980 & 0.010541 \end{pmatrix} , \quad X_K'Y_K = \begin{pmatrix} 112.4409 \\ 103.4680 \end{pmatrix} ,$$

$$(X_K'X_K)^{-1} = \begin{bmatrix} .010541 & -0.006980 \\ -0.00698 & 0.010552 \end{bmatrix} .$$

Así, $\hat{\theta}_K = (X_K'X_K)^{-1}X_K'Y_K = \begin{bmatrix} 0.462397 \\ 0.306271 \end{bmatrix} .$

Por lo tanto, $\hat{\theta}_{KK} = 0.306271 .$

Para K=3

Se tiene que $Y_K = (Y_{K+1} = -2.504927, Y_{K+2} = -1.295782, \dots, Y_T = 1.239364)'$,

$$X_K = \begin{bmatrix} Y_K & = & -0.102889 & Y_{K-1} & = & -2.229019 & Y_{K-2} & = & -1.245203 \\ Y_{K+1} & = & -2.504927 & Y_K & = & -0.102889 & Y_{K-1} & = & -2.229019 \\ Y_{T-1} & = & 0.422207 & Y_{T-2} & = & 2.011572 & Y_{T-3} & = & 2.082886 \end{bmatrix} ,$$

$$(X_K'X_K) = \begin{bmatrix} 169.04016 & 111.9177 & 100.9749 \\ 111.91770 & 168.8619 & 111.0684 \\ 100.97490 & 111.0684 & 164.8154 \end{bmatrix} , \quad X_K'Y_K = \begin{bmatrix} 0.489196 \\ 0.349622 \\ -0.092911 \end{bmatrix} ,$$

$$(X_K'X_K)^{-1} = \begin{bmatrix} 0.01153 & -0.00538 & -0.00343 \\ -0.00538 & 0.01314 & -0.00556 \\ -0.00343 & -0.00556 & 0.01192 \end{bmatrix} .$$

Así, $\hat{\theta}_K = (X_K'X_K)^{-1}X_K'Y_K = \begin{bmatrix} 0.489196 \\ 0.349622 \\ -0.092911 \end{bmatrix} .$

Por lo tanto : $\hat{\theta}_{KK} = -0.092911$.

Hasta aquí, se ha calculado el proceso autorregresivo para grados 1, 2 y 3. El proceso en realidad es de grado $P = 2$, ya que se puede verificar que las dos primeras autocorrelaciones salen del intervalo de confianza para θ_{KK} . A continuación se muestra esta situación:

<u>K</u>	<u>θ_{KK}</u>		
1	0.67	∉	(-.024 , .024) ;
2	0.31	∉	(-.024 , .024) ;
3	-0.10	∈	(-.024 , .024) ;
4	-0.02	∈	(-.024 , .024) ;
5	-0.16	∈	(-.024 , .024) ;
6	0.02	∈	(-.024 , .024) ;
7	-0.17	∈	(-.024 , .024) ;

Por lo tanto, el proceso según las pruebas sobre θ_{KK} , es de grado $P = 2$.

Por lo que, el proceso determinado será para un parámetro estimado:

$$\hat{\theta} = \begin{bmatrix} 0.46 \\ 0.31 \end{bmatrix} .$$

Con este ejemplo se da por terminado el planteamiento de procesos autorregresivos. En el siguiente capítulo se analizarán los vectores autorregresivos.

2. VECTORES AUTORREGRESIVOS

El término *DINAMICO* aplicado al análisis económico, ha tenido diferentes significados en diferentes épocas y para diferentes economistas. Hoy en día, sin embargo, el término se refiere a un tipo de análisis en el cual el objetivo es estudiar el comportamiento en el tiempo de variables o de determinar cuando, dado suficiente tiempo, estas variables tenderán a converger a un cierto (equilibrio) valor. En este capítulo se analizará sistemas dinámicos más complejos que los estudiados en el capítulo anterior de procesos autorregresivos, ya que no siempre una sola ecuación puede describir sistemas dinámicos, por lo que se recurrirá a varias ecuaciones (Dinámicas) para determinar adecuadamente un proceso para la generación de información. Por ejemplo, usando variables trimestrales, se puede suponer que el consumo de bienes C_T en el periodo T depende del ingreso Y_T y en el consumo de los bienes en el periodo anterior (C_{T-1}). Esto es, una función de consumo de la forma:

$$C_T = N_1 + \alpha Y_T + \beta C_{T-1} + e_{1T} .$$

Como el aumento del consumo podrá estimular el crecimiento económico y generar un incremento en el consumo futuro, una ecuación de ingreso resulta:

$$Y_T = N_2 + \delta C_{T-1} + \phi Y_{T-1} + e_{2T} .$$

Aquí, se supone que el ingreso depende de la suma de un consumo retrasado y el ingreso del periodo pasado. Estas dos ecuaciones juntas pueden ser consideradas como una forma estructural de un sistema de ecuaciones simultaneas, y determinan una relación intertemporal entre variables endógenas (*), por lo tanto, es un sistema dinámico.

En el estudio de los sistemas de ecuaciones simultaneas dinámicas la Forma Reducida es de una gran utilidad, la cual consiste en que las variables endógenas sean reducidas a expresiones explícitas en terminos de variables independientes. Así, para el sistema propuesto anteriormente bastaría substituir $N_2 + \delta C_{T-1} + \phi Y_{T-1} + e_{2T}$ por Y_T en la primera ecuación, con lo que el sistema quedaría:

(*) NOTA : Propiamente construido, un modelo económico puede ser resuelto para que se obtenga los valores solución de un cierto conjunto de variables, como por ejemplo, la maximización del nivel de ganancia de la venta de un producto. Dichos valores son conocidos en economía como variables endógenas (Originadas desde adentro), sin embargo, el modelo puede tener variables que son determinadas por fuerzas externas al mismo, y cuyas magnitudes son aceptadas unicamente como valores dados. Estas variables son conocidas como exógenas (Originadas desde afuera).

$$C_T = (N_1 + \alpha N_2) + (\beta + \alpha \delta) C_{T-1} + \alpha \phi Y_{T-1} + (e_{1T} + \alpha e_{2T}),$$

$$Y_T = N_2 + \delta C_{T-1} + \phi Y_{T-1} + e_{2T}.$$

Ahora, el sistema consiste solamente por variables predeterminadas en el lado derecho y por lo tanto, se tiene una forma reducida.

Algunas veces será razonable incluir más de una variable de retraso en el modelo. En ese caso, una manera general del modelo será:

$$C_T = N_1 + \alpha_1 C_{T-1} + \dots + \alpha_p C_{T-p} + \beta_1 Y_{T-1} + \dots + \beta_q Y_{T-q} + V_{1T},$$

$$Y_T = N_2 + \delta_1 C_{T-1} + \dots + \delta_p C_{T-p} + \phi_1 Y_{T-1} + \dots + \phi_q Y_{T-q} + V_{2T}.$$

Este modelo en notación vectorial y matricial queda de la siguiente manera:

$$Y_T = v + \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p} + V_T \dots (4.1)$$

donde

$$Y_T = \begin{bmatrix} C_T \\ Y_T \end{bmatrix}, \quad v = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}, \quad \theta_i = \begin{bmatrix} \alpha_i & \beta_i \\ \delta_i & \phi_i \end{bmatrix} \quad \text{y} \quad V_T = \begin{bmatrix} V_{1T} \\ V_{2T} \end{bmatrix}.$$

Donde se tiene el supuesto de que $p \geq q$ y β_1 y ϕ_1 son igual a cero para $i > q$.

La ecuación (4.1) se parece mucho a un proceso autorregresivo de orden p [AR(p)], el cual se estudio en el capítulo anterior. La única diferencia es que la variable Y_T

del modelo AR(P) es ahora reemplazada por un vector Y_T , el proceso de ruido blanco e_T es reemplazado por el vector V_T , y los coeficientes autorregresivos θ_i son ahora matrices θ_i . Además, un vector de intersección V es incluido en el proceso. Como un vector de variables Y_T esta relacionado con vectores de retraso Y_{T-1}, Y_{T-2}, \dots se designará al modelo (4.1) como VECTORES AUTORREGRESIVOS DE ORDEN P [VAR(P)]. A continuación se define formalmente este modelo.

2.1 CONCEPTOS BASICOS

Como se mencionó anteriormente, los vectores autorregresivos de orden P [VAR(P)] para un sistema de M variables $Y_T = (Y_{1T}, \dots, Y_{MT})'$, podrá ser definido en la misma manera que para el sistema bivariado visto en el principio de esta sección, pero en forma general:

$$Y_T = v + \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p} + v_T \quad .$$

En este sistema de M ecuaciones, $v = (v_1, \dots, v_M)'$ es un vector de dimensión M y la matriz:

$$\theta_i = \begin{bmatrix} \theta_{11,i} & \dots & \theta_{1M,i} \\ \vdots & & \vdots \\ \theta_{M1,i} & \dots & \theta_{MM,i} \end{bmatrix} \quad ,$$

es de dimensión $M \times M$ y $v_T = (v_{1T}, \dots, v_{MT})'$, tiene las siguientes propiedades estocásticas: $E[v_T] = 0$ y $\Sigma_v = E[v_T v_T']$ será la matriz de covarianzas (no singular) para toda T, y además, v_T y v_S no están correlacionados para toda $T \neq S$. Un vector v_T con estas propiedades es llamado *vector de ruido blanco*, en analogía con los términos aplicados para el error e_T de un proceso autorregresivo.

Generalmente, los parámetros $v, \theta_1, \theta_2, \dots, \theta_p$ y Σ_v serán desconocidos en la práctica, por lo que tienen que ser estimados de los datos disponibles, antes de que el proceso

pueda ser usado para pronóstico. Se discutirá la estacionariedad de los vectores autorregresivos en primer lugar, para continuar con la estimación de los parámetros del modelo, ya que esta propiedad será muy útil para estudiar las propiedades asintóticas de los estimadores de los parámetros.

2.2 ESTACIONARIEDAD

La estacionariedad es una propiedad estadística de suma importancia, ya que asegura que las medias, varianzas y autocovarianzas sean constantes a través del tiempo. Se considerará que una colección de vectores aleatorios de dimensión M ..., Y_{T-1} , Y_T , Y_{T+1} ,... son vectores estocásticos, de tal forma que por el momento se restringirá el estudio a vectores autorregresivos, que son un ejemplo particular de vectores estocásticos.

Los vectores estocásticos son estacionarios si:

i) Todos los vectores aleatorios tiene el mismo vector-media μ , esto es, $E[Y_T]=\mu$ para toda T .

ii) Las varianzas de todas las variables aleatorias involucradas son finitas, así $\text{Var}(Y_{mT}) < \infty$ para $m = 1, \dots, M$ y para toda T .

iii) Las matrices de covarianzas de los vectores Y_T y Y_{T+K} que están separadas K periodos no dependen de T sino solamente de K :

$$\text{Cov}(Y_T, Y_{T+K}) = E[(Y_T - \mu)(Y_{T+K} - \mu)'] = \Gamma_K \text{ para toda } T.$$

Esta propiedad implica que, para $K=0$ todos los vectores Y_T tienen la misma matriz de covarianza, esto es, $E[(Y_T -$

$\mu)(Y_T - \mu)'] = \Sigma_Y$ para toda T. Para fines prácticos, estas condiciones implican que las series de tiempo bajo consideración no deben tener tendencias, patrones temporales o varianzas que cambian con el transcurso del tiempo. A menudo será necesario hacer modificaciones a la información para obtener estas propiedades. Además, es claro que casi nunca la información con la que se cuenta (conjunto de datos económicos), tendrán la propiedad de estacionariedad. Existe toda una literatura acerca de procesos no estacionarios y no es el objeto de este trabajo analizarlos, por lo que aquí sólo se tratará con procesos que si sean estacionarios.

Se puede mostrar que los vectores autorregresivos de orden P como el que se definió anteriormente, es estacionario si tiene acotadas sus medias y matrices de covarianzas, y el polinomio definido por el determinante

$$\text{Det}(1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_P z^P) \dots (4.2) ,$$

tiene todas sus raíces fuera del círculo complejo unitario (comparar con lo visto en el capítulo de procesos autorregresivos AR(1)), por ejemplo, para vectores autorregresivos bivariados de orden 1 [VAR(1)]:

$$\begin{bmatrix} Y_{1T} \\ Y_{2T} \end{bmatrix} = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} + \begin{bmatrix} \theta_{11,1} & \theta_{12,1} \\ \theta_{21,1} & \theta_{22,1} \end{bmatrix} \begin{bmatrix} Y_{1,T-1} \\ Y_{2,T-1} \end{bmatrix} + \begin{bmatrix} V_{1T} \\ V_{2T} \end{bmatrix} \dots (4.3).$$

el determinante (4.2) resultaría ser:

$$\text{Det} \left(\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} \theta_{11,1} & \theta_{12,1} \\ \theta_{21,1} & \theta_{22,1} \end{pmatrix} Z \right) =$$

$$= 1 - (\theta_{11,1} + \theta_{22,1})Z + (\theta_{11,1}\theta_{22,1} - \theta_{21,1}\theta_{12,1})Z^2 \dots (4.4),$$

el cual, es un polinomio de segundo grado.

Los vectores autorregresivos descritos en (4.3) serán estacionarios, si este polinomio no tiene raíces en o dentro del círculo unitario, por ejemplo si:

$$\theta_1 = \begin{pmatrix} 0.008 & 0.461 \\ 0.232 & 0.297 \end{pmatrix},$$

el polinomio (4.4) sería:

$$1 - (0.008 + 0.297)Z + (0.008 \cdot 0.297 - 0.232 \cdot 0.461)Z^2 =$$

$$1 - 0.305 Z - 0.104576 Z^2,$$

$$\text{Así: } Z_1 = \frac{0.305 + \sqrt{(0.305)^2 - 4(-0.104576)}}{2(-0.104576)} = -4.877.$$

$$Z_2 = \frac{0.305 - \sqrt{(0.305)^2 - 4(-0.104576)}}{2(-0.104576)} = 1.961.$$

lo que implica que $Z_1 > |1|$ y $Z_2 > |1|$. En consecuencia, el proceso correspondiente es estacionario.

En la siguiente sección se estimarán los parámetros de los vectores autorregresivos.

2.3 ESTIMACION DE VECTORES AUTORREGRESIVOS CON ORDEN CONOCIDO P

En la definición de vectores autorregresivos se propusó el sistema bivariado:

$$Y_T = v + \theta_1 Y_{T-1} + \dots + \theta_p Y_{T-p} + v_T .$$

Ahora, se considerará la m-ésima ecuación de este sistema en forma reducida como se planteó en la introducción de este capítulo.

$$Y_{mt} = V_m + \theta_{m1,1} Y_{1,t-1} + \dots + \theta_{mM,1} Y_{M,t-1} + \dots + \theta_{m1,p} Y_{1,t-p} + \dots + \theta_{mM,p} Y_{M,t-p} + v_{mt} \dots (4.5).$$

Hay que hacer aquí una pausa y aclarar en donde se encuentra el análisis que se está llevando a cabo. Se ha transformado el sistema propuesto a la **forma reducida**, de tal manera, que se tiene hasta ahora, una observación que consta de **M** ecuaciones (variables endógenas), en la que cada una de ellas depende de **M** variables retrasadas **p** periodos, de tal manera que determinán un conjunto de vectores autorregresivos.

Ahora, se pasará a estudiar **T** observaciones y ya no sólo una observación. La consecuencia inmediata es que el proceso se hace bastante complejo en cuanto a notación.

Suponiendo que se tiene T observaciones y P valores muestrales para cada una de las variables, se puede definir los siguientes vectores:

$$y^m = \begin{pmatrix} y_{m,1} \\ y_{m,2} \\ \vdots \\ y_{m,T} \end{pmatrix}, \quad y_{-i}^m = \begin{pmatrix} y_{m,1-i} \\ y_{m,2-i} \\ \vdots \\ y_{m,T-i} \end{pmatrix},$$

(donde por ejemplo, $y_{m,2}$ quiere decir la m -ésima ecuación de la 2a. observación), i tiene valores: $1, \dots, p$ y $m = 1, \dots, M$. En otras palabras, y_{-i}^m contiene las variables de los vectores de retraso por i periodos; Por otro lado, se define $v^m = (v_{m1}, \dots, v_{mT})'$. Usando esta notación (4.5) se puede escribir como:

$$y^m = v_m J + \theta_{m1,1} y_{-1}^1 + \dots + \theta_{mM,1} y_{-1}^M + \dots + \theta_{m1,p} y_{-p}^1 + \dots + \theta_{mM,p} y_{-p}^M + v^m,$$

donde J es un vector de dimension $(T \times 1)$ que consiste en puros unos.

En notación compacta se tiene:

$$y^m = x \theta_m + v^m,$$

donde $x = [J, y_{-1}^1, \dots, y_{-1}^M, y_{-2}^1, \dots, y_{-2}^M, \dots, y_{-p}^1, \dots, y_{-p}^M]$

$$Y \theta_m = [v_m, \theta_{m1,1}, \dots, \theta_{mM,1}, \theta_{m1,2}, \dots, \theta_{mM,2}, \dots, \theta_{m1,p}, \dots, \theta_{mM,p}]'$$

es el vector de coeficientes en la m-ésima ecuación del sistema.

Ahora, se pasará a estudiar ya no sólo una ecuación, sino las M ecuaciones en conjunto:

La t-ésima observación es:

$$y_{1t} = v_1 + \theta_{11,1}x_{1,t-1} + \dots + \theta_{1M,1}x_{M,t-1} + \dots + \theta_{11,p}x_{1,t-p} + \dots + \theta_{1M,p}x_{M,t-p} + v_{1t}$$

$$y_{mt} = v_m + \theta_{m1,1}x_{1,t-1} + \dots + \theta_{mM,1}x_{M,t-1} + \dots + \theta_{m1,p}x_{1,t-p} + \dots + \theta_{mM,p}x_{M,t-p} + v_{mt}$$

$$y_{Mt} = v_M + \theta_{M1,1}x_{1,t-1} + \dots + \theta_{MM,1}x_{M,t-1} + \dots + \theta_{M1,p}x_{1,t-p} + \dots + \theta_{MM,p}x_{M,t-p} + v_{Mt}$$

La T-ésima (Última) observación es:

$$y_{1T} = v_1 + \theta_{11,1}x_{1,T-1} + \dots + \theta_{1M,1}x_{M,T-1} + \dots + \theta_{11,p}x_{1,T-p} + \dots + \theta_{1M,p}x_{M,T-p} + v_{1T}$$

$$y_{mT} = v_m + \theta_{m1,1}y_{1,T-1} + \dots + \theta_{mM,1}y_{M,T-1} + \dots + \theta_{m1,p}y_{1,T-p} + \dots + \theta_{mM,p}y_{M,T-p} + v_{mT} ,$$

$$y_{MT} = v_M + \theta_{M1,1}y_{1,T-1} + \dots + \theta_{MM,1}y_{M,T-1} + \dots + \theta_{M1,p}y_{1,T-p} + \dots + \theta_{MM,p}y_{M,T-p} + v_{MT} .$$

Obtenido el resultado :

$$y^m = x\theta_m + v^m ,$$

que representa la m-ésima ecuación de cada una de las T observaciones; Ahora, se puede mostrar que el sistema de M ecuaciones de T observaciones puede ser representado por:

$$Y = (I_M \otimes X) \theta + v \dots (4.6)$$

donde la matriz X es la misma para cada una de las M ecuaciones de cada observación. Por otro lado, \otimes representa el producto de Kronecker. Este resultado puede ser verificado facilmente al hacer las operaciones segun lo indica (4.6), esto es:

$$\begin{matrix} y_1 = \\ y_2 = \\ \vdots \\ y_M = \end{matrix} \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix} + \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_M \end{pmatrix} + \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^M \end{pmatrix}$$

donde:

$$Y = \begin{pmatrix} Y_{11} & Y_{21} & \dots & Y_{M1} \\ Y_{12} & Y_{22} & \dots & Y_{M2} \\ \vdots & \vdots & \dots & \vdots \\ Y_{1T} & Y_{2T} & \dots & Y_{MT} \end{pmatrix} = (y^1 \ y^2 \ \dots \ y^M) ,$$

y

$$V = \begin{pmatrix} V_{11} & V_{21} & \dots & V_{M1} \\ V_{12} & V_{22} & \dots & V_{M2} \\ \vdots & \vdots & \dots & \vdots \\ V_{1T} & V_{2T} & \dots & V_{MT} \end{pmatrix} = (v^1 \ v^2 \ \dots \ v^M) .$$

Considerando los supuestos de estacionariedad se tiene que la matriz de covarianzas de V es $E[VV'] = \Sigma_V \otimes I_T$.

Esto es debido a que:

$$\begin{aligned} E[V_i] &= 0 \quad \text{para toda } i=1,2,\dots,M \\ \text{y } E[V_i V_i'] &= \sigma_{ii} I_T = \sigma_i^2 I_T \quad \text{para toda } i=1,2,\dots,M \\ \text{y } E[V_i V_j'] &= \sigma_{ij} I_T \quad \text{para toda } i \neq j \text{ y } i,j=1,2,\dots,M, \end{aligned}$$

o, compactamente

$$E[V_i V_j'] = \sigma_{ij} I_T \text{ para toda } i, j=1, 2, \dots, M.$$

Como $V' = [V_1, V_2, \dots, V_M]$, entonces

$$E[VV'] = \begin{pmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_M \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_M \end{bmatrix}' \\ \\ \end{pmatrix} = \begin{pmatrix} \sigma_{11}I & \sigma_{12}I & \dots & \sigma_{1M}I \\ \sigma_{21}I & \sigma_{22}I & \dots & \sigma_{2M}I \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1}I & \sigma_{M2}I & \dots & \sigma_{MM}I \end{pmatrix} = \Sigma_V \otimes I_t.$$

Para obtener el estimador θ se utilizará el método de mínimos cuadrados. Se partirá de que se tiene el sistema en forma reducida

$$Y = (I_M \otimes X) \theta + v \dots (4.7).$$

Si se elige la i -ésima ecuación (observación) de este sistema

$$y_i = x\theta_i + v_i,$$

Por mínimos cuadrados se tiene que :

$$\hat{\theta}_i = (x'x)^{-1}x'y_i.$$

Este es un estimador para la i -ésima observación, pero el mejor estimador generalizado por mínimos cuadrados es:

$$\theta = (X' \Sigma_V^{-1} X)^{-1} X' \Sigma_V^{-1} Y .$$

Aplicando las ecuaciones

$$Y = (I_M \otimes X) \theta + v \quad y \quad E[vv'] = \Sigma_V \otimes I_t,$$

Se tiene :

$$\begin{aligned} \hat{\theta} &= [(I_M \otimes X)' (\Sigma_V \otimes I_t)^{-1} (I_M \otimes X)]^{-1} (I_M \otimes X)' (\Sigma_V \otimes I_t)^{-1} Y \\ &= [I_M \otimes (X'X)^{-1} X']' Y . \end{aligned}$$

Por lo tanto:

$$\hat{\theta} = [I_M \otimes (X'X)^{-1} X']' Y .$$

Ahora, se procederá a investigar una propiedad (Consistencia) de este estimador, que es de suma importancia, supongase primero que

$$v_T \sim N(0, \Sigma_V) .$$

y v_T es independiente de v_S para toda $S \neq T$. Si además, v_T es un proceso estacionario se puede mostrar que:

$$p\text{Lim}_{T \rightarrow \infty} \frac{1}{T} (X'X) = Q ,$$

donde Q es una matriz que representa la media de los cuadrados y la media de los productos cruzados de las variables explicatorias. Si esta matriz es constante en muestras repetidas, entonces

$$\lim_{T \rightarrow \infty} \frac{1}{T} (X'X) = \frac{1}{T} (X'X).$$

Ahora, si las variables explicatorias son estocásticas, los momentos muestrales convergerán probabilísticamente a los momentos de la población.

Por lo tanto,

$$p\lim_{T \rightarrow \infty} \frac{1}{T} (X'X) = Q.$$

donde Q es una matriz simétrica, finita y positivamente definida. Por otro lado

$$p\lim_{T \rightarrow \infty} \frac{1}{T} X'v^m = 0.$$

esto es debido a que:

$$p\lim_{T \rightarrow \infty} \frac{1}{T} X'v^m = \begin{pmatrix} p\lim_{T \rightarrow \infty} \left(\frac{1}{T} \sum v^m \right) \\ p\lim_{T \rightarrow \infty} \left(\frac{1}{T} \sum x_{2T} v^m \right) \\ \vdots \\ p\lim_{T \rightarrow \infty} \left(\frac{1}{T} \sum x_{KT} v^m \right) \end{pmatrix}$$

Pero es evidente que como $T \rightarrow \infty$, se tiene que el límite de probabilidad es igual a cero.

Bajo estos supuestos :

$$\begin{aligned}
 p\text{Lim } \hat{\theta}_m &= p\text{Lim } (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}^m , \\
 &= p\text{Lim } [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\theta_m + \mathbf{v}^m)] , \\
 &= \theta_m + p\text{Lim } \frac{1}{T}(\mathbf{X}'\mathbf{X})^{-1} p\text{Lim } \frac{(\mathbf{X}'\mathbf{v}^m)}{T} , \\
 &= \theta_m .
 \end{aligned}$$

En consecuencia, cada θ_m es consistente y se sigue que

$$p\text{Lim } \hat{\theta} = \theta,$$

Más aun, debido a que se supone que el error \mathbf{v}_t tiene una distribución Normal, $\hat{\theta}$ es asintóticamente equivalente al estimador por Máxima Verosimilitud y por tanto, asintóticamente eficiente y con distribución Normal:

$$\sqrt{T} (\hat{\theta} - \theta) \sim N(0, \Sigma_{\hat{\theta}}) \dots (4.8).$$

(*) Nota: Esta propiedad asintótica es llamada *Convergencia en distribuciones*. La idea consiste en estudiar la distribución de un estimador conforme el tamaño de la muestra se hace muy grande. Hay que notar que el estimador cuya distribución es desconocida en pequeñas muestras, puede a veces conocerse con muestras de gran tamaño. Un conocido y

buen ejemplo de esto es el Teorema del Limite Central : Sean Y_1, Y_2, \dots, Y_T variables aleatorias independientes e idénticamente distribuidas con $E[Y_i] = \beta$ y $\text{Var}[Y_i] = \sigma^2$. Asimismo, sea $\hat{\beta}_t = \sum Y_i/T$, entonces $\sqrt{T} (\hat{\beta}_t - \beta)$ converge en distribución a una Variable Aleatoria Normal $N(0, \sigma^2)$. Este resultado a menudo se escribe como:

$$\sqrt{T} (\hat{\beta}_t - \beta) \sim N(0, \sigma^2) .$$

Este importante resultado dice, que dada una muestra aleatoria de observaciones de una población infinita, cualquier distribución, con media y varianzas finitas, entonces una función simple de la media muestral $\sqrt{T} (\hat{\beta}_t - \beta)$, tiene una distribución en el límite que es normal, de hecho $N(0, \sigma^2)$.

La matriz de Covarianzas de una distribución asintótica puede mostrarse que es:

$$\Sigma_0 = \Sigma_V \otimes Q^{-1} .$$

Para estimar consistentemente la matriz, se necesita un estimador consistente de Σ_V y, en vista de que:

$$p\text{Lim}_{T \rightarrow \infty} \frac{1}{T} (X'X) = Q ,$$

se puede utilizar a $(X'X/T)^{-1}$ como un estimador consistente de Q^{-1} . Como en el capítulo de procesos

autorregresivos, se estimará el ij -ésimo elemento de σ_{ij} de Σ_v por

$$\hat{\sigma}_{ij} = \frac{(y_i - x_i \hat{\theta}_i)' (y_j - x_j \hat{\theta}_j)}{T - MP - 1}$$

donde en el denominador, el número de parámetros $MP + 1$ en cada ecuación, es restada del tamaño de la muestra T . Denotando $\hat{\Sigma}_v$, la matriz con elemento ij -ésimo $\hat{\sigma}_{ij}$, se tiene que un estimador para la matriz de covarianzas de θ es:

$$\Sigma_{\theta} = \hat{\Sigma}_v \otimes (X'X)^{-1}$$

E J E M P L O:

Considerese el sistema bivariado que consiste en cambios de cada cuatro meses en consumo de bienes (y_1) de Estados Unidos y el ingreso disponible (y_2). Los datos para el segundo cuatrimestre del año 1951 al cuarto cuatrimestre de 1969 se encuentran listados en el cuadro No.2. Los datos de cada una de las series se encuentran representadas en la gráfica No.2 y No.3 respectivamente. Las dos gráficas muestran claramente que no existe tendencia alguna o patrones temporales regulares, por lo tanto, se supondrá que los datos están generados por un *proceso estacionario*. Sólo se utilizarán 71 observaciones para propósitos de estimación.

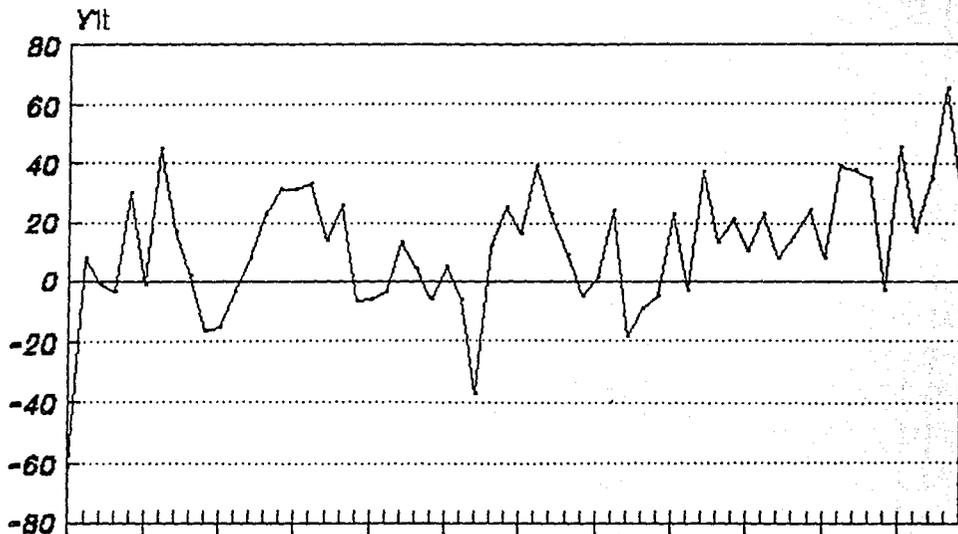
CUADRO No.2

t	Y1t	Y2t	t	Y1t	Y2t	t	Y1t	Y2t
1	-61	42	26	5	1	51	8	30
2	8	-1	27	-6	-20	52	39	47
3	-1	-11	28	-37	-35	53	38	75
4	-4	-12	29	12	6	54	35	27
5	30	16	30	25	45	55	-3	23
6	-1	41	31	16	25	56	46	22
7	45	14	32	39	6	57	17	32
8	17	17	33	23	32	58	35	76
9	2	26	34	9	-30	59	65	47
10	-17	-20	35	-5	10	60	29	17
11	-16	-10	36	1	6	61	-2	6
12	-4	-11	37	24	6	62	22	27
13	8	-23	38	-19	-12	63	0	21
14	23	29	39	-9	-23	64	15	38
15	31	36	40	-5	13	65	31	21
16	31	8	41	23	28	66	7	16
17	33	43	42	-3	17	67	6	17
18	14	31	43	37	38	68	54	36
19	26	29	44	13	14	69	30	43
20	-7	8	45	21	16	70	54	-7
21	-6	9	46	10	3	71	8	9
22	-4	2	47	23	1	72	21	-2
23	13	20	48	8	15	73	9	19
24	4	-10	49	15	17	74	9	47
25	-6	5	50	24	19	75	16	10

FUENTE : JUDGE, HILL & LUTKEPOHL.
 INTRODUCTION TO THEORY & PRACTICE OF ECONOMETRICS.
 PAG. 759.
 ADDISON WEELEY 1987.

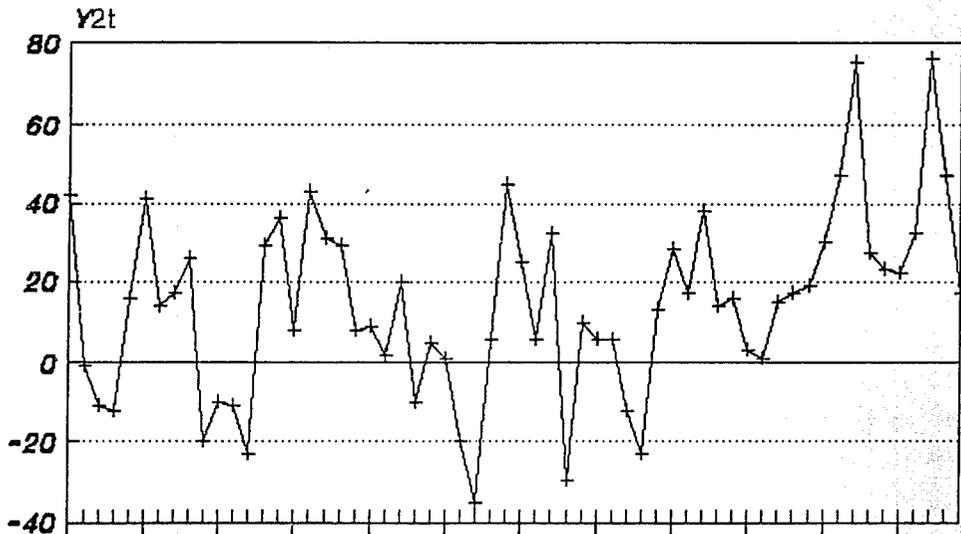
Grafica No.2

Consumo de Bienes en E.U.A.
II 1961- IV 1989 (Cuatrimestres)



Gráfica No.3

Ingreso Disponible en E.U.A.
II 1961- IV 1969 (Cuatrimestres)



Antes de que se estime el modelo bivariado de vectores autorregresivos para estas dos series, se debe especificar el orden P, porque en este modelo al menos, el orden es conocido. En este caso, P=4 fue seleccionado porque esto significa que los valores de retraso de un año previo de cada una de las variables, están incluidas en cada ecuación. Como cuatro retrasos están involucrados, las primeras cuatro observaciones de cada una de las dos variables del cuadro No.2 son consideradas como valores premuestrales. Por tanto, se tiene un tamaño efectivo de la muestra de 67 observaciones (T=71-4=67).

Utilizando el método de mínimos cuadrados se obtiene la estimación del modelo:

$$\begin{matrix} Y_{1t} \\ (3) \\ Y_{2t} \\ 4 \end{matrix} = \begin{bmatrix} 8 \\ (0.15) & (0.12) \\ 9 \\ (0.19) & (0.15) \end{bmatrix} + \begin{bmatrix} -.07 & 0.50 \\ (0.15) & (0.12) \\ 0.23 & 0.33 \\ (0.19) & (0.15) \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} 0.15 & -0.03 \\ (0.17) & (0.15) \\ -0.05 & -0.11 \\ (0.20) & (0.18) \end{bmatrix} \begin{bmatrix} Y_{1,t-2} \\ Y_{2,t-2} \end{bmatrix}$$

$$\begin{bmatrix} 0.11 & -0.16 \\ (0.17) & (0.15) \\ 0.04 & 0.01 \\ (0.20) & (0.18) \end{bmatrix} \begin{bmatrix} Y_{1,t-3} \\ Y_{2,t-3} \end{bmatrix} + \begin{bmatrix} -0.19 & 0.07 \\ (0.12) & (0.13) \\ 0.10 & -0.06 \\ (0.15) & (0.16) \end{bmatrix} \begin{bmatrix} Y_{1,t-4} \\ Y_{2,t-4} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$$

$$\hat{\Sigma}_v = \begin{bmatrix} 284 & 195 \\ 195 & 428 \end{bmatrix}$$

Los números entre parentesis son los errores estandar estimados (matriz de covarianzas de θ).

(EN LA SIGUIENTE SECCION SE HARA EL CALCULO DE ESTE MODELO PARA UN ORDEN $P=1$, POR LO QUE SE VERA CLARAMENTE LA FORMA EN QUE SE APLICAN LOS PROCEDIMIENTOS DE ESTIMACION. EN ESTE EJEMPLO SE DAN LOS RESULTADOS SIMPLEMENTE, DEBIDO A QUE LA FORMA DE CALCULO ES MUY COMPLEJO PARA ORDEN $P=4$, SIENDO EL METODO DE ESTIMACION EXACTAMENTE EL MISMO PARA CUALQUIER ORDEN P .)

Hay problemas obvios con algunas especificaciones. Uno, es la forma arbitraria en que fue escogido el orden del modelo. A continuación se presenta algunas opciones para obtener dicho orden.

2.4 SELECCION DEL ORDEN DE UN VECTOR AUTORREGRESIVO

Como se vio en capítulo de procesos autorregresivos, para escoger el orden adecuado del proceso, se utilizó las autocorrelaciones parciales. En el caso de vectores autorregresivos, las autocorrelaciones parciales son matrices; Escoger el orden de este modelo por inspección requiere bastante experiencia. Por supuesto, exámenes formarles pueden ser utilizados para decidir sobre su significancia. Un enfoque alternativo es el de usar criterios diseñados para ayudar a escoger el orden de los vectores autorregresivos .

A continuación se define el criterio de Akaike (Judge 1988 [1]), que se denota por AIC(n):

$$AIC(n) = \text{Ln det}(\Sigma_n) + \frac{2M^2n}{T} .$$

Y el criterio de Schwarz (Judge 1988 [1]), denotado por SC(n) :

$$SC(n) = \text{Ln det}(\Sigma_n) + \frac{M^2n \text{Ln}(T)}{T} .$$

donde M es el número de variables en el sistema, T es el tamaño de la muestra, y Σ_n es un estimado de la matriz de covarianzas residuales Σ_v obtenida con un modelo de vectores autorregresivos de orden n. Los elementos de Σ_n se calculan de la siguiente manera:

$$\hat{\sigma}_{ij} = \frac{(y^i - x\hat{\theta}_i)'(y^j - x\hat{\theta}_j)}{T}$$

Esto es, en contraste con los estimados de σ_{ij} de Σ_V :

$$\hat{\sigma}_{ij} = \frac{(y^i - x\hat{\theta}_i)'(y^j - x\hat{\theta}_j)}{T - MP - 1}$$

La suma de los cuadrados o productos cruzados, es dividido por el tamaño de la muestra y no por los grados de libertad. El orden P del modelo se escoje de tal manera que los criterios AIC o SC sean minimizados. Esto es, modelos con orden $n = 0, 1, \dots, P$ son estimados con P , como una cota superior prescrita para el orden del modelo. Entonces, las matrices $\tilde{\Sigma}_n$ para $n = 0, 1, \dots, P$ y sus correspondientes valores AIC(n) y SC(n) son calculados. El valor $p(\text{AIC})$, es el orden que minimiza AIC(n) para $n = 0, 1, \dots, P$ y $p(\text{SC})$ es el orden que minimiza SC(n). En el proceso, el tamaño de la muestra T es fijo. En otras palabras, P observaciones para cada variable son tratadas como variables premuestrales.

3. APLICACIONES ECONOMICAS

Como ejemplo, se considera el Consumo e Ingreso en los Estados Unidos, tabulados en el cuadro No.2 y la gráfica No.2 y No.3 representarán respectivamente cada serie. Las matrices Σ_n para $n = 0, 1, \dots, P$ están dadas en el cuadro No.3, con sus correspondientes valores AIC y SC. Las observaciones 1,2,3 y 4, han sido utilizadas como variables muestrales y la muestra utilizada para efectos de estimación, consiste de la quinta observación hasta la numero 71, tal como fue considerada en el modelo de orden 4 de la sección anterior. En el cuadro No.3, se puede observar que ambos criterios son minimizados para orden $P=1$, esto es, $p(\text{AIC}) = p(\text{SC}) = 1$. Con el objeto de utilizar los datos muestrales con eficiencia, a continuación se muestra la estimación del modelo para un orden $P=1$, considerando el primer valor de ambas series como valor premuestral. De tal manera que se tendrá una muestra de tamaño $T=70$ y $M=2$ (Modelo Bivariado) :

$$X = [J, Y_{-1}^1, \dots, Y_{-1}^M, Y_{-2}^1, \dots, Y_{-2}^M, \dots, Y_{-p}^1, \dots, Y_{-p}^M] .$$

Como $M=2$, $P=1$ entoces, $X = [J, Y_{-1}^1, Y_{-1}^2]$.

Así :

CUADRO No. 3

PROCESO	VAR		AIC(N)	SC(N)
Orden			*	
(n)				
0	373	247	112669	11.63
	247	465		
1	266	161	74711	11.34
	161	378		
2	265	162	72990	11.44
	162	374		
3	255	163	68644	11.49
	163	373		
4	246	168	62601	11.52
	168	370		

* $AIC(N) = \ln \det(\Sigma) + 8n/67$

** $SC(n) = \ln \det(\Sigma) + (4n \ln 67)/67$

$$X = \begin{bmatrix} 1 & -61 & 42 \\ 1 & 8 & -1 \\ 1 & -1 & -11 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & 54 & -7 \end{bmatrix}$$

$$Y_{-1}^1 = \begin{bmatrix} -61 \\ 8 \\ -1 \\ \vdots \\ \vdots \\ -54 \end{bmatrix} \quad Y \quad Y_{-1}^2 = \begin{bmatrix} 42 \\ -1 \\ -11 \\ \vdots \\ \vdots \\ -7 \end{bmatrix}$$

Por otro lado:

$$Y^1 = \begin{bmatrix} 8 \\ -1 \\ \vdots \\ \vdots \\ -54 \\ 8 \end{bmatrix}, \quad Y^2 = \begin{bmatrix} -1 \\ -11 \\ \vdots \\ \vdots \\ -7 \\ 9 \end{bmatrix}$$

Ahora, se estimará $\hat{\theta}$:

$$\hat{\theta} = [I_M \otimes (X'X)^{-1}X']'Y$$

Se tiene que:

$$(X'X) = \begin{bmatrix} 70 & 928 & 1090 \\ 928 & 43496 & 29993 \\ 1090 & 29993 & 50548 \end{bmatrix}$$

Asi:

$$(X'X)^{-1} = \begin{bmatrix} 0.023031 & -0.000250 & -0.000340 \\ -0.000250 & 0.000041 & -0.000010 \\ -0.000340 & -0.000010 & 0.000038 \end{bmatrix} .$$

$$(X'X)^{-1}X' = \begin{bmatrix} 0.023828 & 0.021362 & . & . & . & 0.118500 \\ -0.003600 & 0.000100 & . & . & . & 0.002133 \\ 0.002455 & -0.000540 & . & . & . & -0.001650 \end{bmatrix}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ \vdots \\ -54 \\ 8 \\ -1 \\ -11 \\ \vdots \\ 9 \end{bmatrix}$$

$$[I_M \otimes (X'X)^{-1}X'] =$$

$$\begin{bmatrix} 0.023828 & 0.021362 & . & . & . & 0.118500 & 0 & 0 & 0 & 0 & \dots & 0 \\ -0.003600 & 0.000100 & . & . & . & 0.002133 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0.002455 & -0.000540 & . & . & . & -0.001650 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0.023828 & 0.021362 & . & . & . & 0.118500 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & -0.003600 & 0.000100 & . & . & . & 0.002133 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0.002455 & -0.000540 & . & . & . & -0.001650 \end{bmatrix}$$

$$[I_M \otimes (X'X)^{-1}X']'Y = \begin{bmatrix} 6.964268 \\ 0.007653 \\ 0.460916 \\ 7.391297 \\ 0.232436 \\ 0.297163 \end{bmatrix}$$

Por lo tanto, el modelo estimado quedaría de la siguiente manera:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 6.96 \\ (2.51) \\ 7.39 \\ (3.00) \end{bmatrix} + \begin{bmatrix} 0.008 & 0.461 \\ (0.107) & (0.103) \\ 0.232 & 0.297 \\ (0.127) & (0.123) \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix},$$

$$\hat{\Sigma}_v = \begin{bmatrix} 275 & 167 \\ 167 & 390 \end{bmatrix}$$

La forma en que fue determinado $\hat{\Sigma}_v$ fue directamente obtenida de la expresión

$$\hat{\Sigma}_{ij} = \frac{(y^i - x_0^i)'(y^j - x_0^j)}{T - MP - 1},$$

donde $i, j = 1, 2$.

Cabe mencionar que ya se había comprobado en el capítulo de estacionariedad que la matriz de coeficientes es efectivamente estacionaria.

Analizados los vectores autorregresivos, se procederá a investigar el tema de pronóstico en la siguiente sección.

4. PRONOSTICO

Para propósitos de pronóstico se supondrá que el proceso de generación de un conjunto de variables es un sistema de vectores autorregresivos conocido. Bajo esta condición el pronóstico óptimo es la esperanza condicional dada la información hasta el periodo en el cual el pronóstico es hecho. Aquí óptimo significa que el error mínimo cuadrado del pronóstico de cada variable es minimizado. Si el proceso de generación es de orden P conocido, entonces es de la forma:

$$Y_T = v + \theta_1 Y_{T-1} + \dots + \theta_P Y_{T-P} + V_T,$$

con errores independientes de ruido blanco V_T , esto es, V_T es independiente de V_S para toda $S \neq T$. La esperanza condicional $Y_T(h)$ de Y_{T+h} dado Y_T, Y_{T-1}, \dots , es fácil de determinar. Se denotará por E_T el operador esperanza condicional dado Y_T, Y_{T-1}, \dots ,

$$\begin{aligned} Y_T(h) &= E_T[Y_{T+h}] = v + \theta_1 E[Y_{T+h-1}] + \dots + \theta_P E[Y_{T+h-P}], \\ &= v + \theta_1 Y_T(h-1) + \dots + \theta_P Y_T(h-P). \end{aligned}$$

Donde $Y_T(h-i) = Y_{T+h-i}$ para toda $i \geq h$ y $E_T[V_{T+h}] = 0$. Esta fórmula puede ser aplicada repetidamente por cálculo recursivo de los h pasos de pronóstico para $h = 1, 2, \dots$ Por

ejemplo, si $P=1$, de tal manera que Y_T es un proceso $\text{Var}(1)$, se obtiene:

$$Y_T(1) = v + 0_1 Y_T .$$

$$Y_T(2) = v + 0_1 Y_T(1) = v + 0_1 v + 0_1^2 Y_T .$$

$$Y_T(3) = v + 0_1 Y_T(2) = (I + 0_1 + 0_1^2)v + 0_1^3 Y_T .$$

y así sucesivamente.

Suponiendo que el consumo e ingreso de Estados Unidos del Cuadro No.2 del capítulo anterior son ciertamente generados por un proceso $\text{VAR}(1)$:

$$Y_{1T} = \begin{bmatrix} 6.96 \\ 7.39 \end{bmatrix} + \begin{bmatrix} 0.008 & 0.461 \\ 0.232 & 0.297 \end{bmatrix} \begin{bmatrix} Y_{1,T-1} \\ Y_{2,T-1} \end{bmatrix} + \begin{bmatrix} V_{1T} \\ V_{2T} \end{bmatrix} .$$

Se pueden utilizar estas fórmulas para calcular pronósticos óptimos de Y_{72} , Y_{73} ,... comenzando en un periodo $T=71$. Hay que notar que $Y_{1,71} = 8$ y $Y_{2,71} = 9$, así:

$$\begin{bmatrix} Y_{1,71}(1) \\ Y_{2,71}(1) \end{bmatrix} = \begin{bmatrix} 6.96 \\ 7.39 \end{bmatrix} + \begin{bmatrix} 0.008 & 0.461 \\ 0.232 & 0.297 \end{bmatrix} \begin{bmatrix} Y_{1,T-1} = 8 \\ Y_{2,T-1} = 9 \end{bmatrix} \begin{bmatrix} 11.173 \\ 11.919 \end{bmatrix})$$

$$\begin{bmatrix} Y_{1,71}(2) \\ Y_{2,71}(2) \end{bmatrix} = \begin{bmatrix} 6.96 \\ 7.39 \end{bmatrix} + \begin{bmatrix} 0.008 & 0.461 \\ 0.232 & 0.297 \end{bmatrix} \begin{bmatrix} Y_{1,T-1} = 11.173 \\ Y_{2,T-1} = 11.919 \end{bmatrix} \begin{bmatrix} 12.544 \\ 13.522 \end{bmatrix} .$$

Por supuesto, el modelo es una estimación del proceso, por lo que, los pronósticos son igualmente estimados.

La matriz del error medio cuadrado es a menudo utilizada para medir la certeza del pronóstico. Se denotará esta matriz en el paso h como $\Sigma(h)$, esto es:

$$\Sigma(h) = E\{[Y_{T+h} - Y_T(h)][Y_{T+h} - Y_T(h)]'\} .$$

como el pronóstico $Y_T(h)$ es insesgado, esto es, $E[Y_{T+h} - Y_T(h)] = 0$, $\Sigma(h)$ es la matriz de covarianzas del error del pronóstico.

Esta última aseveración es debido a que:

$$MSE(\theta) = COV(\theta) + [\text{sesgo}(\theta)][\text{sesgo}(\theta)]' ,$$

donde θ es un estimador y $MSE(\theta)$ es la matriz del error medio cuadrado.

Mediante el uso de varios artificios algebraicos se puede mostrar que la matriz del error medio cuadrado de un proceso de vectores autorregresivos de orden P tiene la forma:

$$\begin{aligned} \Sigma(h) &= \Sigma_V + M_1 \Sigma_V M_1' + \dots + M_{n-1} \Sigma_V M_{n-1}' . \\ &= \Sigma(h-1) + M_{n-1} \Sigma_V M_{n-1}' . \end{aligned}$$

Donde M_i puede ser calculada de O_i usando recursiones:

$$M_0 = I \quad \text{y} \quad M_i = \sum_j O M_{i-j} \quad \text{para } i=1, 2, \dots$$

por lo tanto,

$$M_1 = \theta_1.$$

$$M_2 = \theta_1 M_1 + \theta_2 M_0 = \theta_1^2 + \theta_2.$$

$$M_3 = \theta_1 M_2 + \theta_2 M_1 + \theta_3 M_0 = \theta_1^3 + \theta_1 \theta_2 + \theta_2 \theta_1 + \theta_3.$$

y así sucesivamente. Para un proceso VAR(1) es fácil ver que $M_i = \theta_1^i$.

Suponiendo que el proceso que se está estudiando en realidad genera los datos de consumo e ingreso de los Estados Unidos, se tiene que:

$$\hat{\Sigma}(1) = \hat{\Sigma}_V = \begin{bmatrix} 275 & 167 \\ 167 & 390 \end{bmatrix},$$

$$\hat{\Sigma}(2) = \hat{\Sigma}_V + \theta_1 \hat{\Sigma}_V \theta_1',$$

$$= \begin{bmatrix} 275 & 167 \\ 167 & 390 \end{bmatrix} + \begin{bmatrix} 0.008 & 0.461 \\ 0.232 & 0.297 \end{bmatrix} \begin{bmatrix} 275 & 167 \\ 167 & 390 \end{bmatrix} \begin{bmatrix} 0.008 & 0.232 \\ 0.461 & 0.297 \end{bmatrix}.$$

$$= \begin{bmatrix} 359 & 240 \\ 240 & 462 \end{bmatrix}.$$

y así sucesivamente.

Para muestras grandes, $\Sigma(h)$ puede ser utilizada como una aproximación a la matriz del error medio cuadrado de un proceso estimado. Por supuesto, aún cuando sólo se ha visto que las simplificaciones de la matriz del error cuadrado medio son estimaciones de la mismas, es válido debido a que se cuenta con un proceso estimado y no real. Esta es la

razón por la que se pone el 'gorrito' en los cálculos. Sin embargo, estas matrices estimadas son de gran utilidad en lo que respecta a intervalos de confianza, que a continuación se analizarán.

Si se considera que los vectores autorregresivos tienen una distribución conocida, se puede determinar regiones de probabilidad (intervalos de confianza) para el pronóstico de las variables. Por ejemplo, si se considera que los vectores autorregresivos tienen una distribución normal, el pronóstico de los errores son por 'herencia' distribuidos normalmente,

$$Y_{T+h} - Y_T(h) \sim N[0, \Sigma(h)] .$$

Un intervalo de pronóstico para un sólo componente de Y_{T+h} puede ser obtenido de acuerdo a este último resultado

$$\frac{Y_{m,T+h} - Y_{m,T}(h)}{\sigma_m(h)} \sim N[0, 1],$$

donde $Y_{m,T+h}$ y $Y_{m,T}(h)$ son la m -ésima componente de Y_{T+h} y $Y_T(h)$ respectivamente y $\sigma_m(h)$ es la desviación estandar del correspondiente error del pronóstico. Esto es, $\sigma_m(h)$ es la raíz cuadrada del m -ésimo elemento de la diagonal de $\Sigma(h)$. Denotando $Z(\alpha)$ el $(1-\alpha)$ porciento de una distribución Normal Estandar, se tiene que de acuerdo a la última expresión.

$$1-\alpha = \Pr \left[-Z_{(\alpha/2)} \leq \frac{y_{m,T+h} - y_{m,T}(h)}{\sigma_m(h)} \leq Z_{(\alpha/2)} \right]$$

$$= \Pr \left[y_{m,T}(h) - Z_{(\alpha/2)} \sigma_m(h) \leq y_{m,T+h} \leq y_{m,T}(h) + Z_{(\alpha/2)} \sigma_m(h) \right]$$

Por lo tanto, un intervalo de pronóstico al $(1-\alpha)$ por ciento de la m -ésima variable del sistema es

$$y_{m,T}(h) \pm Z_{(\alpha/2)} \sigma_m(h)$$

$$\text{ó } [y_{m,T}(h) - Z_{(\alpha/2)} \sigma_m(h), y_{m,T}(h) + Z_{(\alpha/2)} \sigma_m(h)]$$

En la práctica, $y_{m,T}(h)$ y $\sigma_m(h)$ son cantidades desconocidas, que son remplazadas por valores estimados y así se obtiene un intervalo de pronóstico estimado al $(1-\alpha)$ por ciento. Para el consumo e ingreso de Estados Unidos que se está considerando como ejemplo, y suponiendo que es un proceso de vectores autorregresivos de orden 1, con distribución Normal, se obtiene intervalos de pronóstico al 95% de la siguiente manera:

$$Y_{1,71}(1) \pm 1.96 \sigma_1(1) = 11.2 \pm 32.5 \quad \text{o} \quad [-21.3, 43.6].$$

$$Y_{2,71}(1) \pm 1.96 \sigma_2(1) = 11.9 \pm 38.7 \quad \text{o} \quad [-26.8, 50.6].$$

$$Y_{1,71}(2) \pm 1.96 \sigma_1(2) = 12.5 \pm 37.1 \quad \text{o} \quad [-24.6, 49.7].$$

$$Y_{2,71}(2) \pm 1.96 \sigma_2(2) = 13.5 \pm 42.1 \quad \text{o} \quad [-28.6, 55.7].$$

Hasta este punto es importante enfatizar que en la práctica se tiene que hacer un número de suposiciones con objeto de obtener resultados que se buscan. Por ejemplo, se supuso que los vectores autorregresivos de orden 1, estacionarios y con distribución Normal. Por supuesto, no se sabe si ciertamente estas suposiciones son validas, pero se estimó el modelo y bajo esa estimación se dio indicio de que se pueden considerar como ciertas. Si se revisa los datos del Cuadro No.2 correspondientes a los intervalos de pronóstico estimados, se observa que todos ellos caen dentro de los mismos por lo que se puede considerar que es un buen pronóstico con un 95% de certeza.

5. CONCLUSIONES Y RECOMENDACIONES

En este trabajo se discutió sobre el modelado con procesos autorregresivos. Como se mencionó en la introducción, los vectores autorregresivos son de gran importancia en muchos fenómenos económicos, sin embargo, hay varias hipótesis que se toman por descontado como son: el orden del proceso y la estacionariedad de los vectores autorregresivos. Estos dos supuestos requieren de un análisis más cuidadoso. Existen muchas formas para obtener el orden de los vectores, en este trabajo sólo se presentaron dos que resultan ser los más efectivos por este momento. En cuanto a la estacionariedad hay métodos que se han desarrollado como la aplicación de transformaciones a los datos disponibles. Su estudio es muy complejo y merece un análisis muy profundo, lo que se sale de los objetivos prácticos de este trabajo.

La aplicación económica se presentó para un caso de los Estados Unidos de Norteamérica, sin embargo, durante la elaboración del mismo se procuró hacer una aplicación para un caso de México, desafortunadamente no se encontraron los datos necesarios ya que estos están incompletos porque se tenían para algunos años solamente y no eran constantes, es decir, se podían tener datos de 1963, pero no de 1964 ó 1965 y sí de 1966. Por otro lado, hay discrepancia entre las diferentes fuentes, por lo que para presentar una adecuada

aplicación de la teoría se optó por una base de datos más completa como la utilizada precisamente en el texto.

Finalmente, diremos que los vectores autorregresivos presentan efectivamente una solución para pronósticos de sistemas de ecuaciones simultaneas dinámicas donde se tiene la importante propiedad de que las variables endógenas están en terminos de variables exógenas. El procesos de solución es complicado, pero seguramente a corto plazo se tendrán paquetes que realizen estos cálculos con mayor efectividad.

6. REFERENCIAS

- 1) Judge, Hill, Griffiths, Lutkepohl, Chao Lee
Introduction to the Theory & Practice of
econometrics.
John Wiley & sons. 1988.

- 2) Guerrero, Victor M.
Análisis estadístico de series de tiempo
económicas.
U.A.M. 1987.

- 3) Henderson, James H., Quandt, Richard E.
Microeconomic Theory.
Mc. Graw Hill. 1992.

- 4) Chang, Alpha C.
Fundamental Methods of Mathematical Economics.
Mc. Graw Hill. 1991.

- 5) J. Johnston.
Econometric Methods.
Mc. Graw Hill. 1992.

ARTICULOS:

- 6) Judge, Lutkepohl & Lee (1985).
Theory & practice of econometrics.
2nd. edition, New York, Wiley.

- 7) Lutkepohl (1985)
Comparisons of criteria for estimating the order
of vector autorregressive process.
Journal of time series analysis, 8 (1987), p. 373.

- 8) Lutkepohl (1987)
Forecasting aggregated vector ARMA processes
Econometrica, 41, p.161-164.

- 9) Sims (1981)
An autorregressive index model for the U.S.
1948-1975.
Large-scale macroeconomic models.
North-Holland, p. 283-327