Una Implementación de Orden Cuadrático del Algoritmo de Innovaciones en Series de Tiempo

An Implementation of Quadratic Order of the Algorithm of Innovations in Series of Time

Nadia Yadhira **Martínez Martínez.** Departamento de Estadística y Cálculo. Universidad Autónoma Agraria Antonio Narro. Buenavista, Saltillo, Coah., 25315, México. Correo-e: Nadia mtz@hotmail.com

Resumen

Este trabajo trata sobre el problema de pronóstico para una serie de tiempo estacionaria. El principal resultado es la formulación de un método de orden cuadrático para implementar el algoritmo de innovaciones, cuya codificación usual tiene orden cúbico. El procedimiento propuesto combina el esquema recursivo de Durbin Levinson, con un teorema de transformación de representaciones para un pronóstico.
A Second Order Implementation for the Innovations Algorithm in Time Series Analysis

Abstract

This work is about the problem of prognosis for a stationary time series. The main result is the formulation of a quadratic order method to implement the innovations algorithm, which usual coding has cubic order. The proposed procedure combines the Durbin-Levinson algorithmwith a new theorem of transformation of representations for a forecast, one in terms of the innovations, and the other one in terms of the original observations.

Introducción

Este trabajo se ubica en el área conocida como análisis de una serie de tiempo, y el principal objetivo es presentar una implementación eficiente de un procedimiento de pronóstico conocido como algoritmo de innovaciones. De manera intuitiva, una serie de tiempo es una sucesión $\{X_t\}$ de variables aleatorias definidas sobre un mismo espacio de probabilidad, donde la variable aleatoria X_t se interpreta como la observación que se realiza en el tiempo t, el cual, en el caso considerado en este trabajo, puede variar en un subconjunto de los números enteros. El rasgo fundamental de la sucesión $\{X_t\}$ es que, en contraste con el supuesto comúnmente adoptado en la teoría estadística clásica (Borovkov, 1999, Dudewicz y Mishra, 1988, Shao, 2008, Wackerly *et al*., 2009), no se supone la independencia de las variables X_t, ni que éstas tengan la misma distribución, características que permiten incluir en el estudio una gran variedad de observaciones que surgen en la práctica, como las ventas diarias de una almacén, la asistencia semanal a teatros, la población de un país, o la aparición de manchas solares (Brockwell y Davis, 1991; Shumway y Stoffer, 2006; Shao, 2008). Las series que se analizan son estacionarias, en el sentido de que las propiedades relevantes de un segmento $(X_1,X_2,...,X_n)$ son las mismas que las del segmento trasladado $(X_{1+h},X_{h+2},...,X_{n+h})$ para cada entero h; como no se hace supuesto alguno sobre la distribución de los datos X_t, las propiedades importante de la serie involucran sus momentos. Formalmente, una serie es estacionaria si

(a) $E[X_t]= \mu$ no depende de t;

(b) Para cada s y t, $\mathrm{Cov}(X_s,X_t)$ está bien definida y depende sólo de la diferencia entre s y t

Estas propiedades permiten definir la función de autocovarianza asociada a la serie estacionaria $\{X_t\}$ mediante

$$\gamma(h) = E[X_{t+h}X_t], \quad h = 0, \pm, \pm 2, \ldots. \qquad (1.1)$$

Aunque la idea de serie estacionaria es bastante restrictiva y, en general, una de tales series no es un modelo razonable para los datos que se observan comúnmente en la práctica, las series estacionarias son parte esencial del denominado modelo clásico, el cual si captura una amplia gama de situaciones reales (Shao, 2008).
El problema de pronóstico para una serie estacionaria puede describirse como sigue: Dadas las observaciones $X_1,\ldots,X_n$ registradas hasta el tiempo n, determinar 'la mejor' aproximación para la variable X_{n+h} que se observará en el tiempo futuro $n + h$ en términos de los datos disponibles $X_1,\ldots,X_n$. Un resultado conocido en la teoría estadística establece que la mejor aproximación a X_{n+h} en el sentido de minimizar el error cuadrático esperado es la esperanza condicional $g(X_1,\ldots,X_n)= E[X_{n+h}|X_1,\ldots,X_n]$; sin embargo, al no conocer la distribución de $\{X_t\}$, la función g(·) no se puede calcular, y en la teoría de series de tiempo, la aproximación que se pretende encontrar es 'la mejor aproximación lineal' basada en los datos observados hasta el momento actual n :

$$\hat{g}(X_1,\ldots,X_n)= a_1X_1 + \cdots a_nX_n \qquad (1.2)$$

donde $a_1,\ldots,a_n$ son constantes y

$$\min_{b_1,\ldots,b_n} E[((X_{n+h} - (b_1X_1 + \cdots b_nX_n))^2]= E[((X_{n+h} - (a_1X_1 + \cdots a_nX_n))^2]. \qquad (1.3)$$

tales números ai siempre existen y se expresan solamente en términos de la función de autocovarianza γ (·); de hecho, a1,...,an pueden determinarse por medio de diferenciación (Fulks 1982; Khuri 2003) e invirtiendo una matriz. Sin embargo, se conocen métodos alternativos y más eficientes para determinar pronósticos. En este trabajo el principal interés se centra en el denominado algoritmo de innovaciones, y el problema que se analiza es diseñar una implementación de dicho procedimiento que realice menos operaciones que la codificación usual. Este objetivo se alcanza combinando un algoritmo conocido como método de Durbin-Levinson, con un resultado de transformación entre representaciones de un pronóstico que se establece en este trabajo como Teorema 7.1, el cual representa la principal aportación técnica del artículo.

La presentación ha sido organizada de la siguiente manera. En la Sección 2 se discute la noción de proyección ortogonal en un espacio vectorial con producto interno y su relación con el problema de pronóstico, mientras que estas ideas son usadas en la Sección 3 para formular las ecuaciones de pronóstico y para encontrar los errores correspondientes. A continuación, en la Sección 4 se discute el procedimiento para calcular pronósticos conocidos como método de Durbin-Levinson, y se demuestra que es de segundo orden. Posteriormente, en la Sección 5 se introduce la base ortogonal de innovaciones y en la Sección 6 se formula el algoritmo clásico correspondiente, probando que es de orden cúbico. Finalmente, en la Sección 7 se presenta un método que relaciona las expresiones para un pronóstico en términos de las bases de innovaciones y de observaciones originales, y la exposición concluye en la Sección 8 formulando una implementación del algoritmo de innovaciones que tiene orden cuadrático; el método propuesto combina el algoritmo de Durbin-Levinson con el resultado de transformación previamente obtenido.

Proyecciones

Puesto que este trabajo se centra en la determinación de las constantes $a_1,\ldots,a_n$ en (1.2), las cuales tienen la propiedad de optimalidad en (1.3), las ideas básicas de espacio vectorial y producto interno son de importancia fundamental en el desarrollo subsecuente; para un estudio de estos conceptos en espacios de dimensión finita, vea, por ejemplo, Graybill (2000, 2001), Harville (2008), Hoffman y Kunze (1971) y, particularmente, Strang (2003). La intuición del caso de dimensión finita es de gran importancia en el estudio de series de tiempo, aunque el entorno de trabajo es, naturalmente, el espacio

$$L^2 = L^2(\Omega, \mathcal{F}, P)$$

de todas las variables aleatorias $Y:\Omega \to R$ con segundo momento finito, y donde $(\Omega, \mathcal{F}, P)$ es el espacio de probabilidad donde la serie $\{X_t\}$ está definida. Este espacio es de dimensión infinita y el producto interno esta especificado por

$$\langle Y_1, Y_2\rangle = E[Y_1Y_2], \quad Y_1, Y_2 \in L^2, \qquad (2.1)$$

y la norma correspondiente está dada por

$$\|Y\| = \sqrt{\langle Y, Y\rangle} = E[Y^2]^{1/2}. \qquad (2.2)$$

Con respecto a esta norma, L^2 es un espacio métrico completo, es decir, toda sucesión de Cauchy es convergente (Apostol, 1974; Rudin, 1987; Royden, 1998). Considere ahora, una sucesión finita $W_1,...,W_k$ de variables en L^2, y sea W el espacio vectorial generado por estas variables:

$$\mathcal{W} = \{c_1W_1 + \cdots + c_nW_n \mid c_1, c_2, \ldots, c_n \in \mathbb{R}\}$$

el cual es un subespacio cerrado de L^2 (Rudin, 1987, Royden, 1988). Dado $Y \in L^2$, existe una única variable aleatoria $P_W Y$ que satisface las siguientes propiedades:

$$P_{\mathcal{W}}Y \in \mathcal{W} \quad \text{y} \quad \|Y - P_{\mathcal{W}}Y\| \leq \|Y - \mathbf{w}\| \quad \text{para todo } \mathbf{w} \in \mathcal{W}. \tag{2.3}$$

Estas condiciones son equivalentes a

$$P_{\mathcal{W}}Y \in \mathcal{W} \quad \text{y} \quad \langle P_{\mathcal{W}}Y, \mathbf{w}\rangle = \langle Y, \mathbf{w}\rangle, \quad \mathbf{w} \in \mathcal{W}. \tag{2.4}$$

(Brockwell y Davis, 1991, Strang, 2003); debido a la (bi-)linealidad del producto interno, estas condiciones pueden expresarse como

$$P_{\mathcal{W}}Y \in \mathcal{W} \quad \text{y} \quad \langle P_{\mathcal{W}}Y, W_i\rangle = \langle Y, W_i\rangle, \quad i = 1, 2, \ldots, k. \tag{2.5}$$

La variable $P_W Y$ se llama la proyección (ortogonal) de Y sobre W y es una transformación lineal; la equivalencia de (2.3–(2.5) se conoce teorema de proyección. La siguiente propiedad será útil Si W_0 y W_1 son dos subespacios de L^2, entonces

$$\mathcal{W}_0 \subset \mathcal{W}_1 \Rightarrow P_{\mathcal{W}_0}P_{\mathcal{W}_1} = P_{\mathcal{W}_0} \tag{2.6}$$

A partir de esta discusión se es claro que el mejor pronóstico $g(X_1,...,X_n)$ para X_{n+1} dadas las observaciones $X_1,...,X_n$, es la proyección de X_{n+1} sobre el espacio generado por $X_1,...,X_n$; compare (1.2) y (1.3) con (2.3).

Ecuaciones y Errores de Pronóstico

Dada una serie estacionaria $\{X_t\}$ con función de autocovarianza $\gamma(\cdot)$, en esta sección se establecen las ecuaciones para determinar la proyección ortogonal una variable X_{n+h}, donde $h > 0$, sobre el espacio generado por las variables $X_1,...,X_n$ donde, por el momento, el entero n es fijo, y el $\mathcal{L}_n$ es el subespacio de L^2 generado por las variables X_t con $1 \leq t \leq n$:

$$\mathcal{L}_n(X_1, \ldots, X_n) \equiv \mathcal{L}_n := \{a_1X_1 + \cdots + a_nX_n \mid (a_1, \ldots, a_n) \in \mathbb{R}^n\}. \tag{3.1}$$

Para cualquier variable aleatoria Y con varianza finita, la proyección ortogonal de Y sobre $\mathcal{L}_n$ se denotará mediante $\hat{Y}$:

$$\hat{Y} = P_{\mathcal{L}_n}Y,$$

de manera que, como se estableció en la sección precedente, $\hat{Y}$ está caracterizada por los siguientes requerimientos:

$$\hat{Y} \in \mathcal{L}_n, \quad \text{y} \quad \langle Y, X_k\rangle = \left\langle \hat{Y}, X_k\right\rangle, \quad k = 1, 2, \ldots, n; \tag{3.3}$$

en palabras, $\hat{Y}$ es una combinación lineal de $X_1,...,X_n$, y los productos internos de $\hat{Y}$ y Y con X_i coinciden, i =1, 2,...,n. Ahora se determinará la forma que estas condiciones adoptan para el caso en que Y es la variable X_{n+1} que se observará una unidad de tiempo después de haber registrado el valor de X_n; primero, se estudiará el caso h = 1. Observe que la inclusión $\hat{X}_{n+1} \in \mathcal{L}_n$ significa que $\hat{X}_{n+1}$ es una combinación lineal de $X_1,...,X_n$, de manera que existen constantes $\phi_{n\,1}, \phi_{n\,2},..., \phi_{n\,n}$ tales que

$$\hat{X}_{n+1} = \phi_{n\,1}X_n + \phi_{n\,2}X_{n-1} + \cdots + \phi_{n\,n}X_1,$$

o de forma más compacta,

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{n\,j}X_{n+1-j}, \tag{3.4}$$

A continuación, note que la condición de que X_{n+1} y $\hat{X}_{n+1}$ tengan los mismos productos internos con las variables X_r, r =1, 2,...,n puede escribirse en términos de la función de autocovarianza:

$$\begin{aligned}\langle X_{n+1}, X_r\rangle &= \left\langle \hat{X}_{n+1}, X_r\right\rangle \\ &= \left\langle \sum_{j=1}^{n} \phi_{n\,j}X_{n+1-j}, X_r\right\rangle \\ &= \sum_{j=1}^{n} \phi_{n\,j}\langle X_{n+1-j}, X_r\rangle, \quad r = 1, 2, \ldots, n,\end{aligned} \tag{3.4}$$

igualdades que equivalen a

$$\gamma(n+1-r) = \sum_{j=1}^{n} \phi_{n\,j}\gamma(n+1-j-r), \quad r = 1, 2, \ldots, n.$$

Una forma más conveniente para este sistema se obtiene escribiendo

$$i = n + 1 - r$$

con esta notación, cuando r toma cualquier valor entre 1 y n, i también lo hace y el anterior sistema de ecuaciones se escribe como

$$\gamma(i) = \sum_{j=1}^{n} \phi_{n\,j}\gamma(i-j), \quad i = 1, 2, \ldots, n.$$

Definiendo los vectores $\gamma_n, \varphi_n \in \mathbb{R}^n$ y la matriz Γ_n de orden $n \times n$ mediante

$$\gamma_n = \begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n) \end{bmatrix}, \quad \varphi_n = \begin{bmatrix} \phi_{n\,1} \\ \phi_{n\,2} \\ \vdots \\ \phi_{n,\,n} \end{bmatrix}, \quad \Gamma_n = [\gamma(i-j)]_{i,j=1,2,\ldots,n}; \tag{3.6}$$

con esta notación, las ecuaciones anteriores equivalen a

$$\gamma_n = \Gamma_n \, \varphi_n \tag{3.7}$$

sistema que siempre es consistente, por el teorema de proyección. Después de encontrar una solución φ_n para esta ecuación vectorial, el error cuadrático de pronóstico, denotado por v_n y definido mediante

$$v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2, \tag{3.8}$$

puede determinarse de forma muy simple. En efecto, observando que

$$\left\langle X_{n+1} - \hat{X}_{n+1}, X_r \right\rangle = 0$$

para todo r =1, 2,...,n (vea la primera igualdad en 3.5), como $\hat{X}_{n+1}$ es una combinación lineal de $X_1,\ldots,X_n$ se desprende que

$$\left\langle X_{n+1} - \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle = 0, \tag{3.9}$$

lo que equivale a

$$\left\langle \hat{X}_{n+1}, X_{n+1} \right\rangle = \left\langle \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle = \|\hat{X}_{n+1}\|^2.$$

Con esto en mente,

$$\begin{aligned} \|X_{n+1} - \hat{X}_{n+1}\|^2 &= \left\langle X_{n+1} - \hat{X}_{n+1}, X_{n+1} - \hat{X}_{n+1} \right\rangle \\ &= \left\langle X_{n+1} - \hat{X}_{n+1}, X_{n+1} \right\rangle - \left\langle X_{n+1} - \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle \\ &= \left\langle X_{n+1} - \hat{X}_{n+1}, X_{n+1} \right\rangle \text{ (por la ecuación (3.9))} \\ &= \langle X_{n+1}, X_{n+1} \rangle - \left\langle \hat{X}_{n+1}, X_{n+1} \right\rangle \\ &= \gamma(0) - \left\langle \hat{X}_{n+1}, X_{n+1} \right\rangle \end{aligned}$$

donde, de nueva cuenta, se utilizó la ecuación (3.9) para establecer la ultima igualdad. Por lo tanto,

$$\|X_{n+1} - \hat{X}_{n+1}\|^2 = \gamma(0) - \left\langle \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle \tag{3.10}$$

Ahora se calculará la norma cuadrática de $\hat{X}_{n+1}$. Usando (3.4) se tiene que

$$\begin{aligned} \left\langle \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle &= \left\langle \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j}, \sum_{i=1}^{n} \phi_{n\,i} X_{n+1-i} \right\rangle \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} \phi_{n\,i} \langle X_{n+1-j}, X_{n+1-i} \rangle \phi_{n\,j} \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} \phi_{n\,i} \gamma(i-j) \phi_{n\,j} \\ &= \varphi_n' \Gamma_n \varphi_n \end{aligned}$$

y recordando que $\Gamma_n \varphi_n = \gamma_n$, se obtiene que $\left\langle \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle = \varphi_n' \gamma_n$; combinando esta relación con (3.10), se llega a la expresión

$$v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2 = \gamma(0) - \varphi_n' \gamma_n.$$

De esta manera, al resolver la ecuación (3.7) es posible determinar, tanto la proyección $\hat{X}_{n+1}$ por medio de (3.4), como el error de pronóstico v_n a través de la anterior expresión desplegada. Esta discusión se resume en el siguiente teorema.

Teorema 3.1. Suponga que $\{X_t\}$ es una serie estacionaria con función de autocovarianza $\gamma(\cdot)$.

En este caso, la proyección $\hat{X}_{n+1}$ de X_{n+1} sobre el espacio $\mathcal{L}n$ generado por $X_1,\ldots,X_n$ *está dada por*

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j}$$

donde $\varphi_n = (\phi_{n1}, \ldots, \phi_{nn})$ es cualquier vector que satisface el sistema de ecuaciones

$$\gamma_n = \Gamma_n \varphi_n,$$

y la notación es como en (3.6). Más aún, el error de pronóstico vn esta dado por

$$v_n = \|\hat{X}_{n+1} - X_{n+1}\|^2 = \gamma(0) - \varphi_n' \gamma_n. \tag{3.11}$$

El Algoritmo de Durbin-Levinson

En este sección se formula un procedimiento recursivo para calcular los coeficientes $\phi_{n\,k}$ en la expresión para la proyección de X_{n+1} sobre $\mathcal{L}(X_1, \ldots, X_n)$. Este procedimiento calcula el vector de coeficientes φ_{n-1} el error de pronóstico v_n en términos del vector de coeficientes φ_{n-1} y el correspondiente error de v_{n-1}, evitando el problema de invertir la matriz Γ_n. El esquema de cálculo en el siguiente teorema se conoce como el algoritmo de Durbin-Levinson. En el resto del capítulo, se utiliza la siguiente notación:

$$\hat{X}_1 = 0, \qquad v_0 = \|X_1 - \hat{X}_1\|^2 = \gamma(0) \tag{4.1}$$
$$\hat{X}_n = P_{\mathcal{L}(X_1,\dots,X_{n-1})}X_n, \quad v_{n-1} = \|X_n - \hat{X}_n\|^2, \quad n = 2,3,\dots$$

Por conveniencia, a continuación se reescribe la expresión para $\hat{X}_n$:

$$\hat{X}_n = \sum_{k=1}^{n-1} \phi_{n-1,k} X_{n-k}. \tag{4.2}$$

Teorema 4.1. Suponga que $\{X_t\}$ es un proceso estacionario con función de autocovarianza $\gamma(\cdot)$ tal que $\gamma(0) > 0$ y $\gamma(h) \to 0$ conforme $h \to \infty$, de tal forma que las matrices $\mathbf{\Gamma}_n$ son no singulares. En este caso, los coeficientes $\phi_{n\,k}$ y los errores cuadráticos de pronóstico satisfacen las siguientes relaciones:

(i) $\phi_{11} = \dfrac{\gamma(1)}{\gamma(0)}, \quad v_0 = \gamma(0)$

(ii) $\phi_{n\,n} = \dfrac{1}{v_{n-1}}\left[\gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1\,j}\,\gamma(n-j)\right], \quad v_n = v_{n-1}[1-\phi_{n\,n}^2]$

(iii) $\phi_{n\,k} = \phi_{n-1\,k} - \phi_{n\,n}\phi_{n-1\,n-k}, \quad k = 1,2,\dots,n-1$

Una demostración de este teorema puede encontrarse en Brockwell y Davis (1991), o en Fuller (1996). Este resultado permite encontrar los coeficientes $\phi_{n\,k}$ en

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j},$$

así como el error de pronóstico $v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2$, de manera recursiva, sin invertir la matriz $\mathbf{\Gamma}_n$. El procedimiento se describe como sigue:

Algoritmo de Durbin-Levinson

Entrada*:* Un vector $(\gamma(0), \gamma(1), \dots, \gamma(n))$, donde $\gamma(\cdot)$ es una función de autocovarianza.

Salida*:* El vector $\varphi_n = (\phi_{n\,1}, \dots, \phi_{n\,n})$ de coeficientes de la proyección $\hat{X}_{n+1}$, así como v_n, el correspondiente error cuadrático de pronóstico.

• Inicio: Defina $\phi_{1,1} = \gamma(0)/\gamma(1)$ y $v_1 = (1-\phi_{11}^2)v_0$ y ponga $k = 1$

• Paso Recursivo: En términos de $\phi_{k\,j}, j = 1,2,\dots,k$ y v_k calcule

(a) $\phi_{k+1\,k+1} = v_k^{-1}\left[\gamma(k+1) - \sum_{j=1}^{k} \phi_{k\,j}\,\gamma(k+1-j)\right]$

(b) $v_{k+1} = v_k[1-\phi_{k+1\,k+1}^2]$

(c) $\phi_{k+1\,j} = \phi_{k\,j} - \phi_{k+1\,k+1}\phi_{k\,n-j}, \quad j = 1,2,\dots,k.$

Además

(d) Incremente k en una unidad.

• Paso de Prueba: Si $k = n$ detenga el procedimiento. Los números actuales v_k y $\phi_{k\,j}$, $j = 1,2,\dots,k$ las cantidades deseadas.

Si $k<n$, vaya al paso anterior.

A continuación se analiza el orden del algoritmo, esto es el requerimiento de memoria y la cantidad de operaciones aritméticas necesarias para obtener los coeficientes $\phi_{n\,k}$ se prestará atención al número de multiplicaciones y divisiones, pes estas operaciones consumen un mayor tiempo de cómputo que las sumas y restas.

• En la etapa inicial se realizan una división para evaluar ö11 y dos multiplicaciones para calcular v_1; en total, 3 operaciones

• En la etapa recursiva, para pasar de φ_k y v_k hacia φ_{k+1} y v_{k+1} se llevan a cabo

(a) k multiplicaciones y una división para evaluar $\phi_{k+1\,k+1}$;

(b) dos multiplicaciones para calcular v_{k+1}, y

(c) k multiplicaciones para determinar las restantes componentes $\phi_{k+1,j}, j = 1,2,\dots k$ En total se realizan $2k + 3$ multiplicaciones o divisiones.

Como la etapa iterativa se repite para k =1, 2,..., n-1, todo el algoritmo requiere la realización de $3 + \sum_{k=1}^{n-1}(2k+3) = n(n+2)$ multiplicaciones o divisiones. En este caso, se dice que el algoritmo de Durbin-Levinson es de orden n^2, y se escribe $O(n2)$. En cuanto a los requerimientos de memoria, se necesitan n + 1 lugares para almacenar $\gamma(0), \gamma(1), \dots, \gamma(n), n,$ lugares adicionales para guardar $v_1,\dots,v_n$, y otros n lugares para almacenar $\phi_{n\,k}, k = 1,2,\dots,n,$ totalizando $3n$+1 posiciones de memoria para los datos de entrada y salida. En esta descripción, la palabra 'lugar' se refiere a la cantidad de bytes necesaria para almacenar un número real con precisión determinada, y dicha cantidad depende del entorno de trabajo y del programa que ejecuta la tarea. Para propósitos de referencia futura, esta discusión se resume en el siguiente teorema.

Teorema 4.2. (i) Con respecto a las operaciones aritméticas que realiza, el algoritmo de Durbin-Levinson es de orden $O(n^2)$.

(ii) El requerimiento de memoria del algoritmo de Durbin-Levinson es de orden $O(n)$; si se desea almacenar todos los coeficientes $1 \le j \le s \le n$ el requerimiento de memoria naturalmente aumenta , y es de orden $O(n_2)$.

Un Conjunto Ortogonal

El problema de calcular proyecciones sobre un subespacio $\mathcal{L}$ se facilita al disponer de un conjunto generador que sea ortogonal (Hoffman y Kunze, 1971, Graybill 2001, Harville, 2008). En esta sección se construye uno de tales conjuntos para $\mathcal{L}(X_1, X_2, \ldots, X_n)$ donde, en adelante, n es un entero positivo fijo. Considere las variables aleatorias

$$U_1 = X_1, \quad U_k = X_k - P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k, \quad k = 1,2,\ldots,n. \tag{5.1}$$

En este caso, debido a que

$$U_i \in \mathcal{L}(X_1,\ldots,X_i) \ \text{ y } \ U_j \text{ es ortogonal a } \mathcal{L}(X_1,\ldots,X_{j-1}),$$

se tiene la siguiente relación de ortogonalidad:

$$i < j \Rightarrow \langle U_i, U_j \rangle = 0.$$

Suponga ahora que la matriz de varianza Γ_n del vector $(X_1,\ldots,X_n)$ es no singular para cada n, de manera que $X_1,\ldots,X_n$ son linealmente independientes. Como U_k es una combinación lineal de $X_1,\ldots,X_k$ en la que el coeficiente de X_k es 1, se tiene que $U_k \neq 0$ para cada k. Combinando esta propiedad con la anterior relación de ortogonalidad se desprende que

$U_1,\ldots,U_n$ son linealmente independientes.

y por lo tanto, como $U_1,\ldots,U_n \in \mathcal{L}(X_1,\ldots,X_n)$ un argumento directo de dimensionalidad permite establecer que

$$\mathcal{L}(U_1,\ldots,U_n) = \mathcal{L}(X_1,\ldots,X_n)$$

para cada n. Por otro lado, recuerde que el pronóstico para X_k en términos de $X_1,\ldots,X_{k-1}$ es $\hat{X}_k = P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k$, de manera que $U_k = X_k - \hat{X}_k$, y entonces

$$\|U_k\|^2 = \|X_k - \hat{X}_k\|^2 = v_{k-1}$$

vea (3.11).La discusión precedente se resume como sigue:

Lema 5.1. *Dada una serie estacionaria $\{X_t\}$* defina las variables aleatorias U_k mediante (5.1). En este caso,

(i) Las variables $U_1, U_2, U_3, \ldots$ son ortogonales;

(ii) Se satisface la igualdad $\mathcal{L}(U_1,U_2,\ldots,U_k) = \mathcal{L}(X_1,X_2,\ldots,X_k)$ para cada k;

(iii) Para cada $k = 1, 2, 3, \ldots$, $Uk = vk-1$.

La variable U_k se denomina la innovación en el tiempo k. Para justificar este término, note que X_k se expresa como

$$X_k = P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k + [X_k - P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k]$$

y ubíquese en el momento en que Xk va a ser observada. En ese instante, ya se registraron $X_1,\ldots,X_{k-1}$, y el término $P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k$ es combinación lineal de esas variables; por lo tanto, $P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k$ es conocido para el observador. En contraste, $[X_k - P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_k]$ es ortogonal a $X_1,\ldots,X_{k-1}$, y en ese sentido, representa 'la nueva información' que el observador obtendrá al registrar el valor de X_k. Por otro lado, como el conjunto ortogonal $U_1,\ldots,U_n$ genera el mismo espacio vectorial que $X_1,\ldots,X_n$, se tiene que la proyección

$$\hat{X}_{n+1} = P_{\mathcal{L}(X_1,\ldots,X_n)}X_{n+1} = P_{\mathcal{L}(U_1,\ldots,U_n)}X_{n+1}$$

se expresa como

$$\begin{aligned}\hat{X}_{n+1} &= \theta_{n\,1}U_n + \theta_{n\,2}U_{n-1} + \cdots + \theta_{n\,n-1}U_2 + \theta_{n\,n}U_1 \\ &= \sum_{k=1}^{n} \theta_{n\,k}U_{n+1-k}.\end{aligned} \tag{5.2}$$

Esta es la representación de $\hat{X}_{n+1}$ en términos de la innovaciones $U_1,\ldots,U_n$.

Determinación de los Coeficientes de las Innovaciones

En esta sección se enuncia un resultado clásico sobre la determinación de los coeficientes de la expresión de $\hat{X}_{n+1}$ como combinación lineal de las innovaciones $U_1,\ldots,U_k$. Como punto de partida, suponga que, para cada k =1,...,n - 1, ya se han determinado los coeficientes $\theta_{k\,j}, 1 \leq j \leq k$ en las expresiones

$$\hat{X}_{k+1} = \sum_{j=1}^{k} \theta_{k\,j}U_{k+1-j} \tag{6.1}$$

así como los correspondientes errores cuadráticos de pronóstico

$$v_k = \|U_{k+1}\|^2 = \|X_{k+1} - \hat{X}_{k+1}\|^2. \tag{6.2}$$

El siguiente teorema muestra como pueden determinarse los coeficientes $\phi_{n\,k}$ y el error v_n en términos de estas cantidades.

Teorema 6.1. Suponga que para cada entero k =1, 2,... n-1 se conocen los coeficientes $\theta_{k\,j},\ 1 \le j \le k$ *, así como* v_k*. En este caso,*

(i) Los coeficientes ϕ_{nk} *en la expresión*

$$\hat{X}_{n+1} = \sum_{k=1}^{n} \theta_{n\,k} U_{n+1-k} \tag{6.3}$$

satisfacen $v_0\theta_{n\,n} = \gamma(n),$

$$v_i\theta_{n\,n-i} = \gamma(n-i) - \sum_{s=0}^{i-1} \theta_{i\,i-s}\theta_{n\,n-s}v_s, \quad i = 1,2,\ldots,n-1.$$

(ii) Más aún,

$$v_n = \gamma(0) - \sum_{j=1}^{n} \theta_{n\,j}^2 v_{n-j}$$

Una demostración de este resultado se encuentra, por ejemplo, en Brockwell y Davis (1991), o Fuller (1996). El Teorema 6.1 sugiere una manera natural de encontrar los coeficientes $\theta_{n\,n}, \theta_{n\,n-1}, \ldots, \theta_{n\,1}$, en la expresión (6.3) para el pronóstico $\hat{X}_{n+1}$, así como el correspondiente error cuadrático $v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2$

Suponga que la función de autocovarianza es tal que, para cada n, la matriz $\mathbf{\Gamma}_n$ es invertible, de manera que $v_k > 0$ para todo k. Iniciando con

$$v_0 = \gamma(0),$$

el Teorema 6.1 permite calcular

(i) $\theta_{1\,1} = \gamma(1)/v_0$ y $v_1 = \gamma(0) - \theta_{1\,1}^2 v_0$.

Usando estas cantidades, ahora se calculan los coeficientes $\theta_{2\,2}, \theta_{2\,1}$ y v_1 como sigue:

(ii) $\theta_{2\,2} = \gamma(2)/v_0,\ \theta_{2\,1} = [\gamma(1) - \theta_{1\,1}\theta_{2\,2}v_0]/v_1,$ y $v_2 = \gamma(0) - \theta_{2\,2}^2 v_0 - \theta_{2\,1}^2 v_1$.

A partir de este punto, la fórmulas del Teorema 6.1 permiten calcular $\theta_{3\,3}, \theta_{3\,2}, \theta_{3\,1}$ y v_3, ese orden. En general, dados $\theta_{k\,j},\ 1 \le j \le k$, y $v_o,\ldots,v_k$, es posible evaluar $\theta_{n\,n}, \theta_{n\,n-1}, \ldots, \theta_{n\,1}$, en el orden indicado. Este procedimiento se describe formalmente a continuación:

Algoritmo de Innovaciones

Entrada: Un vector $(\gamma(0), \gamma(1), \ldots, \gamma(n))$, donde $\gamma(\cdot)$ es una función de autocovarianza y $n \ge 2$

Salida: El vector $\theta_n = (\theta_{n\,1}, \ldots, \theta_{n\,n})$ de coeficientes de la proyección $\hat{X}_{n+1}$ en (6.3), así como v_n, el correspondiente error cuadrático de pronóstico.

- Inicio: Defina $v_0 = \gamma(0), \theta_{1\,1} = \gamma(0)/v_0, v_1 = \gamma(0) - \theta_{1\,1}^2 v_0$ y $k = 1$.
- Paso Recursivo: Incremente k en 1 y calcule

$$\theta_{k\,k} = \gamma(k)/v_0,$$

$$\theta_{k\,k-i} = [\gamma(k-i) - \sum_{s=0}^{i-1} \theta_{i\,i-s}\theta_{k\,k-s}v_s]/v_i, \quad i = 1,2,\ldots,k-1,$$

$$v_k = \gamma(0) - \sum_{j=1}^{k} \theta_{k\,j}^2 v_{k-j}$$

- Paso de Prueba: Si $k = n$ detenga el procedimiento. Los números actuales v_k y $\theta_{k\,j},\ j = 1, 2,\ldots,k$ son las cantidades deseadas.

Si $k<n$, vaya al paso anterior.

Un aspecto importante en el algoritmo de innovaciones, es que para evaluar v_k y los números $\theta_{k\,j}$, es necesario tener disponibles todos las cantidades θ_{si} y v_i,$1 \le i \le s < k$. Así, debe reservarse lugar para todos los números en el siguiente arreglo:

$$\begin{array}{cccccc}
v_0 & & & & & \\
\theta_{1\,1} & v_1 & & & & \\
\theta_{2\,2} & \theta_{2\,1} & v_2 & & & \\
\theta_{3\,3} & \theta_{3\,3} & \theta_{3\,1} & v_3 & & \\
\vdots & \vdots & \vdots & \vdots & \ddots & \\
\theta_{n\,n} & \theta_{n\,n-1} & \theta_{n\,n-2} & \cdots & \theta_{n\,1} & v_n
\end{array}$$

totalizando $1 + 2 + \cdots + (n+1) = (n+1)(n+2)/2$ lugares de memoria. Además se requieren $n + 1$ lugares para $(\gamma(0), \gamma(1), \ldots, \gamma(n))$ y uno más para la variable k. Así, el requerimiento de memoria es de $(n+2)(n+3)/2$ lugares numéricos, es decir, $O(n^2)$.

Ahora se discutirá la carga aritmética del algoritmo, contando el número de multiplicaciones y divisiones necesarias para ejecutarlo. En el paso recursivo, aplicado al entero k requiere

(a) Una división para evaluar $\theta_{k\,k}$;

(b) $2(i - 1)$ multiplicaciones y una división para calcular $\theta_{k\,k-i}$, dando un total de $2i$-1 operaciones; esto debe hacerse para i =1, 2,k - 1, arrojando un total de $\sum_{i=1}^{k-1}(2i-1) = k(k-1) - (k-1) = (k-1)^2$ multiplicaciones o divisiones. Como el paso recursivo se realiza con valores k =2,...,n, el total de operaciones en todos los paso, es $\sum_{k=2}^{n}(k-1)^2 = (n-1)(n)(2n-1)/6 \approx n^3/3$.

Esta discusión se resume en el siguiente teorema:

Teorema 6.2. Para determinar los n coeficientes $\theta_{n\,n}, \ldots, \theta_{n\,1}$ y el error cuadrático de pronóstico v_n, el algoritmo de innovaciones requiere *(n + 2)(n + 3)/2* lugares numéricos de memoria, y la realización de $3 + (n-1)(n)(2n-1)/6 \approx n^3/3$ multiplicaciones o divisiones. Así, el algoritmo de innovaciones es de orden $O(n^2)$ en cuanto a requerimiento de memoria, y de orden $O(n^3)$ con respecto a la carga aritmética.

Como ya se mencionó, el principal objetivo de este trabajo es diseñar una implementación del algoritmo de innovaciones que sea más eficiente, en el sentido de que carga aritmética sea de orden menor a n^3. Dicho resultado depende de la relación entre los coeficientes que se utilizan para representar la proyección $\hat{X}_{n+1}$ en términos de las innovaciones $U_1,\ldots,U_n$, y las cantidades que se usan en la representación como combinación lineal de las variables originales $X_1,\ldots,X_n$.

Relación entre dos Expresiones para un Pronóstico

El propósito de esta sección es desarrollar la idea principal en que se sustenta la formulación de un procedimiento de segundo orden para determinar el vector de coeficientes θ_n que aparece al expresar el pronóstico $\hat{X}_{n+1}$ en términos de las innovaciones $U_1,\ldots,U_n$. El resultado principal relaciona dicho vector con la representación de los pronósticos $\hat{X}_{k+1}$ en términos de las variables originales.

Considere las dos representaciones para la proyección $\hat{X}_{n+1}$ de X_{n+1} sobre $\mathcal{L}(X_1,\ldots,X_n)$:

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \theta_{n\,j} U_{n+1-j}$$
$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j} \qquad (7.1)$$

donde, como antes, U_{k+1} es la innovación en el tiempo $k+1$. El siguiente resultado es la principal contribución técnica de este trabajo.

Teorema 7.1. Considere una serie estacionaria $\{X_t\}$ tal que la matriz Γ_k es invertible para cada entero positivo *k; vea (3.6).* Sea n un entero positivo, y suponga que, para algún entero k entre 1 y n, se tiene la siguiente representación:

$$P_{\mathcal{L}(X_1,\ldots,X_k)} X_{n+1} = \sum_{i=1}^{k} c_i X_{k+1-i}, \qquad (7.2)$$

donde $c_1,\ldots,c_k$ *son constantes. En este caso,*
(i) El coeficiente $\theta_{n\;n+1-k}$ *en (7.1) está dado por*

$$\theta_{n\;n+1-k} = c_1.$$

(il) La proyección X_{n+1} *sobre el espacio* $\mathcal{L}(X_1,\ldots,X_{k-1})$ *está dada por*

$$P_{\mathcal{L}(X_1,\ldots,X_{k-1})} X_{n+1} = \sum_{i=1}^{k-1} d_i X_{k-i},$$

donde

$$(d_1, d_2, \ldots, d_{k-1}) = (c_2, c_3, \ldots, c_k) + c_1(\phi_{k-1\,1}, \phi_{k-1\,2}, \ldots, \phi_{k-1\,k-1})$$

Demostración. Primeramente, observe que (7.2) puede escribirse como

$$\begin{aligned} P_{\mathcal{L}(X_1,\ldots,X_k)} X_{n+1} &= c_1 X_k + \sum_{i=2}^{k} c_i X_{k+1-i} \\ &= c_1 X_k + \sum_{j=1}^{k-1} c_{j+1} X_{k-j} \\ &= c_1 [X_k - P_{\mathcal{L}(X_1,\ldots,X_{k-1})} X_k] \\ &\quad + \left[\sum_{j=1}^{k-1} c_{j+1} X_{k-j} + c_1 P_{\mathcal{L}(X_1,\ldots,X_{k-1})} X_k\right]. \end{aligned}$$

El término entre corchetes que multiplica a c_1 es precisamente la innovación U_k, mientras que $P_{\mathcal{L}(X_1,\ldots,X_{k-1})} X_k = \sum_{j=1}^{k-1} \phi_{k-1,j} X_{k-j}$. Por lo tanto,

$$\begin{aligned} P_{\mathcal{L}(X_1,\ldots,X_k)} X_{n+1} &= c_1 U_k + \left[\sum_{j=1}^{k-1} c_{j+1} X_{k-j} + c_1 \sum_{j=1}^{k-1} \phi_{k-1\,j} X_{k-j}\right] \\ &= c_1 U_k + \left[\sum_{j=1}^{k-1} [c_{j+1} + c_1 \phi_{k-1\,j}] X_{k-j}\right] \end{aligned} \qquad (7.3)$$

Por otro lado, la primera igualdad en (7.1) puede escribirse como

$$P_{\mathcal{L}(X_1,\ldots,X_n)} X_{n+1} = \sum_{j=1}^{n} \theta_{n\,j} U_{n+1-j} = \sum_{s=1}^{n} \theta_{n\;n+1-s} U_s$$

Recuerde ahora dos hechos fundamentales:

(a) Si $\mathcal{M}_0$ y $\mathcal{M}_1$ son dos subespacios tales $\mathcal{M}_0 \subset \mathcal{M}_1$, que entonces

$$P_{\mathcal{M}_0} P_{\mathcal{M}_1} = P_{\mathcal{M}_0};$$

vea, por ejemplo, Hoffman y Kunze (1971) o Lipschutz (1996).

(b) $\mathcal{L}(X_1 \ldots, X_s) = \mathcal{L}(U_1, \ldots, U_s)$, igualdad que se estableció en el capítulo precedente. Usando estas dos propiedades, y recordando que k es un entero entre 1 y n, se desprende que

$$\begin{aligned} P_{\mathcal{L}(X_1,\ldots,X_k)} X_{n+1} &= P_{\mathcal{L}(X_1,\ldots,X_k)} P_{\mathcal{L}(X_1,\ldots,X_n)} X_{n+1} \\ &= P_{\mathcal{L}(U_1,\ldots,U_k)} \left[\sum_{j=1}^{n} \theta_{n\,j} U_{n+1-j}\right] \\ &= \sum_{j=1}^{n} \theta_{n\,j} P_{\mathcal{L}(U_1,\ldots,U_k)} U_{n+1-j} \\ &= \sum_{j=n+1-k}^{n} \theta_{n\,j} U_{n+1-j} \end{aligned}$$

donde se usó que $P_{\mathcal{L}(U_1,\ldots,U_k)} U_s = 0$ si $s > k+1$ y $P_{\mathcal{L}(U_1,\ldots,U_k)} U_s = U_s$ cuando $s \le k$. El mismo argumento con k - 1 en vez de k muestra que

$$P_{\mathcal{L}(X_1,\ldots,X_{k-1})} X_{n+1} = \sum_{j=n+2-k}^{n} \theta_{n\,j} U_{n+1-j}.$$

Combinando estas dos últimas relaciones desplegadas se obtiene que

$$\begin{aligned} P_{\mathcal{L}(X_1,\ldots,X_k)} X_{n+1} &= \sum_{j=n+1-k}^{n} \theta_{n\,j} U_{n+1-j} \\ &= \theta_{n\,n+1-k} U_k + \sum_{j=n+2-k}^{n} \theta_{n\,j} U_{n+1-j} \end{aligned}$$

y entonces

$$P_{\mathcal{L}(X_1,\ldots,X_k)} X_{n+1} = \theta_{n\,n+1-k} U_k + P_{\mathcal{L}(X_1,\ldots,X_{k-1})} X_{n+1}$$

Comparando esta expresión con (7.3), la ortogonalidad y no nulidad de las innovaciones implica que $\theta_{n\;n+1-k} = c_1$ como se establece en la parte (i), mientras que

$$P_{\mathcal{L}(X_1,\ldots,X_{k-1})}X_{n+1}=\sum_{j=1}^{k-1}[c_{j+1}+c_1\phi_{k-1j}]X_{k-j}=\sum_{j=1}^{k-1}d_jX_{k-j},$$

donde $d_j=c_{j+1}+c_1\phi_{k-1j}$ para j =1, 2,...,k - 1, como se establece en la parte (*ii*).

Conociendo los vectores $\varphi_k=(\phi_{k1},\ldots,\phi_{kk}), k=1,2,\ldots,n,$ este teorema permite calcular el vector $\theta_n=(\theta_{n1},\theta_{n2},\ldots,\theta_{nn})$, cuyas componentes son los coeficientes en la representación de $\hat{X}_{n+1}$ en términos de las innovaciones $U_1,\ldots,U_n$, Iniciando con

$$P_{\mathcal{L}(X_1,\ldots,X_n)}X_{n+1}=\sum_{i=1}^{n}\phi_{n\,i}X_{n+1-i}$$

el Teorema 7.1 con $k=n$ y $c=(c_1,\ldots,c_n)=\varphi_n$ establece que $\theta_{n\,n+1-n}=\theta_{n\,1}=c_1=\phi_{n\,1}$, y permite calcular el vector $d=(d_1,d_2,\ldots,d_{n-1})$ tal que

$$P_{\mathcal{L}(X_1,\ldots,X_{n-1})}X_{n+1}=\sum_{i=1}^{n-1}d_jX_{n-j}$$

Otra aplicación del teorema precedente con $k=n-1$ y con el vector d en lugar de c arroja que $\theta_{n\,n+1-(n-1)}=\theta_{n2}=d_1$, y un vector $\tilde{d}=(\tilde{d}_1,\tilde{d}_2,\ldots,\tilde{d}_{n-2})$ que satisface

$$P_{\mathcal{L}(X_1,\ldots,X_{n-2})}X_{n+1}=\sum_{i=1}^{n-2}\tilde{d}_jX_{n-1-j}$$

A partir de aquí se obtiene $\theta_{n\,3}=\tilde{d}_1$ y otro vector $\hat{\tilde{d}}$ tal que

$$P_{\mathcal{L}(X_1,\ldots,X_{n-3})}X_{n+1}=\sum_{i=1}^{n-3}\tilde{d}_jX_{n-2-j}$$

Continuando con este proceso, se llega finalmente a determinar todos los coeficientes $\theta_{n\,1},\theta_{n\,2},\ldots,\theta_{n\,n}$. Estas ideas se presentan en el siguiente algoritmo:

Algoritmo de Transformación

Entrada: Un entero positivo n y los vectores φ_k tales que

$$\hat{X}_{k+1}=(X_1,\ldots,X_k)\varphi_k$$

para cada k =1,...,n.

Salida: El vector $\theta_n=(\theta_{n\,1},\ldots,\theta_{n\,n})$ tal que $\hat{X}_{k+1}=(U_1,\ldots,U_k)\theta_n$.

Inicio: Ponga k = 1 y defina el vector $\alpha=\varphi_n$

Etapa de cálculo: (a) Si $k=n$ ponga $\theta_{n\,n}=\alpha_1$

(b) Si $k<n$, ponga $\theta_{n,n+1-k}=\alpha_1$, y sustituya α por el vector

$$(\alpha_2,\ldots,\alpha_{n+1-k})+\alpha_1\varphi_{n-k}$$

e incremente k por una unidad.

Etapa de salida: (a) Si $k>n$ salga del procedimiento mostrando el vector θ_n.

(b) Si $k\le n$ vaya a la etapa de cálculo.

Para concluir la sección, se analiza el orden del procedimiento de transformación para obtener el vector $\boldsymbol{\theta}_n$.

Teorema 7.2. El procedimiento de transformación descrito anteriormente es de orden cuadrático, tanto en el requerimiento de memoria, como en la carga aritmética.

Demostración. Debido a que el procedimiento requiere utiliza los vectores $\boldsymbol{\varphi}_1,\ldots,\boldsymbol{\varphi}_n$ cuyas longitudes suman 1+2+··· $n=n(n+1)/2$, es claro que el requerimiento de memoria tiene orden $O(n^2)$. Por otro lado, en la etapa de cálculo se realizan $n-k$ multiplicaciones al realizar el producto $\alpha_1\boldsymbol{\varphi}_{n-k}$, y esto se hace para k =1, 2,...,n - 1. Por lo tanto se realizan

$$(n-1)+(n-1)+\cdots+1=n(n-1)/2$$

multiplicaciones, así que el procedimiento también es de orden cuadrático respecto a la carga aritmética.

Implementación Cuadrática del Algoritmo de Innovaciones

En esta sección se describe un procedimiento para determinar el vector

$$\boldsymbol{\theta}_n=(\theta_{n\,1},\theta_{n\,2},\ldots,\theta_{n\,n},)$$

en la representación

$$\hat{X}_{n+1}=\sum_{k=1}^{n}\theta_{n\,k}X_{n+1-k}$$

así como el error de pronóstico

$$v_n=\|X_{n+1}-\hat{X}_{n+1}\|^2$$

El siguiente procedimiento combina el algoritmo de Durbin-Levinson el esquema de transformación de la sección precedente para resolver el problema, cuya solución es el principal objetivo del algoritmo de innovaciones.

Algoritmo combinado para determinar $\boldsymbol{\theta}_n$ y v_n

Paso 1: Utilice el método de Durbin-Levinson para determinar vn y los vectores φ_k, $k=1,2,\ldots,n,$ en las representaciones $\hat{X}_{k+1}=(X_1,\ldots,X_k)\varphi_k$.

Paso 2: Aplique el método de transformación descrito en la sección precedente para determinar θ_n a partir de φ_k, $k=1,2,\ldots,n$.

El siguiente teorema condensa el trabajo realizado previamente, y se enuncia para enfatizar la importancia del resultado.

Teorema 8.1. El algoritmo combinado descrito anteriormente es de orden cuadrático, tanto en los requerimientos de memoria como en la carga aritmética.

Demostración. Debido a que tanto el algoritmo de Durbin-Levinson como el método de transformación son de orden cuadrático en requerimientos de memoria y carga aritmética, se desprende que el esquema combinado también es de orden $O(n^2)$.

Este resultado indica claramente la ventaja del algoritmo combinado respecto a la formulación original del algoritmo de innovaciones. Ambos procedimientos encuentran θ_n y v_n, pero el algoritmo combinado es de orden $O(n^2)$ y por lo tanto requiere realizar menos operaciones aritméticas que el algoritmo de innovaciones original, el cual tiene orden $O(n^3)$.

Literatura Citada

Apostol, T.M. 1974. Mathematical Analysis: A Modern Approach to Advanced Calculus. 2a. ed. Addison-Wesley. Reading, Massachusetts, USA. 492 p.

Borovkov, A.A.1999.Mathematical Statistics.Taylor & Francis. Amsterdam, The Netherlands. 592 p.

Brockwell, P.J. and R.A. Davis. 1991. Time Series: Theory and Methods. 2a. ed. Springer-Verlag, New York, USA. 577 p.

Dudewicz, E.J. and S.N. Mishra. 1988. Mathematical Statistics. 2a. ed. Wiley, New York, USA. 864 p.

Fuller, W.A. 1996. Introduction to Statistical Time Series. Wiley. New York, USA.

Fulks, W. 1982. Calculo Avanzado. Limusa. México, D.F. 551 p.

Graybill, F.A. 2000. Theory and Application of the Linear Model. Duxbury. New York, USA. 704 p.

Graybill, F.A. 2001. Matrices with Applications in Statistics. Duxbury. New York, USA. 461 p.

Harville, D.A. 2008. Matrix Algebra from a Statistician's Perspective. Springer-Verlag. New York, USA. 629 p.

Hoffman, K. and R. Kunze. 1971. Linear Algebra. Prentice-Hall. New York, USA. 407 p.

Khuri, A.I. 2003. Advanced Calculus with Applications in Statistics. 2a. ed. Wiley. New York, USA. 677 p.

Lipschutz, S. 1996. Linear Algebra. McGraw-Hill. New York, USA. 473 p.

Rudin, W. 1987. Real and Complex Analysis. McGraw-Hill. New York, USA. 416 p.

Royden, H.L. 1988. Real Analysis. MacMillan. New York, USA. 444 p.

Shao, J. 2008. Mathematical Statistics. Springer. New York, USA. 591 p.

Shumway, R.H. and D.S. Stoffer. 2006. Time Series Analysis and Its Applications. With R

Examples. 2a. ed. Springer-Verlag, New York, USA. 575 p.

Strang, G. 2003. Introduction to Linear Algebra. 3a. ed. Wellesley-Cambridge Press, New York. USA. 571 p.

Wackerly, D., W. Mendenhall and R.L. Scheaffer. 2009. Mathematical Statistics with Applications. Thomson Brooks/Cole, Belmont, CA. USA. 914 p.