

45  
29

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

---

Facultad de Ciencias

" SIACASET: SISTEMA PARA EL ANALISIS CANONICO  
DE SERIES DE TIEMPO "

T E S I S

que para obtener el título de

ACTUARIO

presenta

MARIA ANTONIETA DE MUCHA MENDIOLA

TESIS CON  
FALLA DE ORIGEN

México D.F. Julio de 1991.



## **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

Introducción.	2
Introducción a las series de tiempo multivariadas.	
I. Procesos estocásticos.	
I.1 Definición.	3
I.2 Función de distribución y momentos.	4
I.3 Estacionariedad y Ergodicidad.	6
I.4 Representación espectral.	8
I.5 Filtros.	12
II. Series de tiempo.	
II.1 Definición.	15
II.2 Estimación de momentos.	15
II.3 Estimación de la densidad espectral.	16
Introducción a las componentes principales y correlación canónica.	
I. Datos no correlacionados.	
I.1 Componentes principales.	23
I.2 Correlación canónica.	26
II. Series de tiempo.	
I.1 Componentes principales.	30
I.2 Correlación canónica.	32
Ejemplo.	35
Conclusiones y Recomendaciones.	62
Bibliografía.	63

## INTRODUCCION

En los últimos 25 años el análisis de series de tiempo ha tenido un gran desarrollo, en particular, el enfoque espectral referente a la descomposición de series de tiempo en componentes principales, se ha convertido en una herramienta muy importante para diversas aplicaciones. El presente trabajo, presenta una alternativa para el análisis canónico de series de tiempo utilizando el enfoque.

En el primer capítulo se desarrolla la teoría general de series de tiempo multivariadas, partiendo de las definiciones de un proceso estocástico. Se define la densidad espectral de un proceso estocástico y se discuten dos alternativas para su estimación.

En el segundo capítulo se presentan la teoría general para la estimación de componentes principales y variables canónicas. Finalmente es desarrollada para vectores aleatorios de datos no correlacionados y para finalizar se utiliza el enfoque espectral para obtener resultados similares en el caso de series de tiempo correlacionadas.

En el tercer capítulo ejemplifica la metodología desarrollada en los capítulos anteriores, para esto se creó un programa escrito en FORTRAN 77 apoyado por rutinas estadísticas de IMSL. Finalmente se presentan las conclusiones del trabajo desarrollado.

I. PROCESOS ESTOCASTICOS

1. Definición.

Sea  $(\Omega, \mathcal{A}, P)$  un espacio de probabilidad y  $T$  un conjunto de índices. Un proceso estocástico se define como una función  $X$  tal que:

$$\begin{aligned} X : T \times \Omega &\longrightarrow \mathbb{R}^n \\ (t, \omega) &\longmapsto X(t, \omega) \end{aligned}$$

Generalmente el índice  $t$  está asociado al tiempo, por lo que podemos referirnos a un proceso estocástico como un fenómeno estadístico que evoluciona en el tiempo de acuerdo a ciertas leyes de probabilidad.

Diremos que un proceso estocástico es *continuo* cuando  $t$  pertenezca a un conjunto de valores continuos y *discreto* cuando  $t$  pertenezca a un conjunto de valores discretos, los cuales generalmente están igualmente espaciados. A lo largo del presente trabajo consideraremos únicamente procesos estocásticos discretos.

Como ejemplos de procesos estocásticos podemos mencionar: la temperatura diaria de una región, la inflación mensual de un país, el precio de cierre diario de una acción bursátil en una bolsa de valores, etc. Para ejemplificar un proceso múltiple podemos pensar en la temperatura diaria de las cinco ciudades más importantes de México, o bien en el precio de cierre diario de todas las acciones en una bolsa de valores.

## 2. Función de distribución y momentos.

Sabemos que el comportamiento de una variable aleatoria queda completamente descrito por su función de distribución; en el caso de un proceso estocástico es muy difícil lograr una descripción probabilística completa, pues tendríamos que definir la función de distribución conjunta del proceso para todo valor de  $t$  in  $T$ . Esto en la práctica resulta imposible; de hecho sólo podemos definir distribuciones marginales monovariadas, bivariadas, trivariadas, etc, como por ejemplo:

$$F_{t_1}(\underline{x}) = F_{t_1}(x_1, \dots, x_n) = \text{Prob} \{ X_1(t_1, \omega) \leq x_1, \dots, X_n(t_1, \omega) \leq x_n \},$$

$$\begin{aligned} F_{t_1, t_2}(\underline{x}, \underline{y}) &= F_{t_1, t_2}(x_1, \dots, x_n, y_1, \dots, y_n) = \\ &= \text{Prob} \{ X_1(t_1, \omega) \leq x_1, \dots, X_n(t_1, \omega) \leq x_n, \\ &\quad X_1(t_2, \omega) \leq y_1, \dots, X_n(t_2, \omega) \leq y_n \}, \end{aligned}$$

etc.,

y a través de ellas obtener información de la distribución conjunta de infinitas variables aleatorias.

Una manera alterna de aproximarse al comportamiento de un proceso estocástico es a través de sus momentos y aún cuando éstos estén definidos en términos de la función de distribución, es posible obtener estimadores confiables que nos permiten lograr nuestro propósito. Más aún, conociendo los primeros dos momentos de un proceso, se obtiene la información suficiente para lograr una buena aproximación de manera lineal.

Definimos la *media de un proceso estocástico* como:

$$E[X(t, \omega)] = C_X(t) = \begin{bmatrix} C_1(t) \\ \vdots \\ C_n(t) \end{bmatrix} \quad (1.2.1)$$

donde

$$C_j(t) = E[X_j(t, \omega)] = \int_{\mathbb{R}^n} u_j dF_t(u).$$

$C_X(t)$  existirá si y sólo si  $E|X_j(t, \omega)| < \infty$ ,  $j=1, \dots, n$ .

La función de covarianza de un proceso estocástico queda definida como:

$$C_{XX}(t_1, t_2) = \begin{bmatrix} C_{11}(t_1, t_2) & C_{12}(t_1, t_2) & \dots & C_{1n}(t_1, t_2) \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1}(t_1, t_2) & C_{n2}(t_1, t_2) & \dots & C_{nn}(t_1, t_2) \end{bmatrix} \quad (1.2.2)$$

donde

$$\begin{aligned} C_{jk}(t_1, t_2) &= E\{[X_j - C_j(t_1)][X_k - C_k(t_2)]\} = \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} [u_j - C_j(t_1)][u_k - C_k(t_2)] dF_{t_1, t_2}(u, v), \end{aligned}$$

si  $j=k$ , diremos que  $C_{jj}(t_1, t_2)$  es la función de autocovarianza de la componente  $X_j(t, \omega)$  y si  $j \neq k$ ,  $C_{jk}(t_1, t_2)$  será la función de covarianza cruzada de  $X_j(t, \omega)$  con  $X_k(t, \omega)$ .

$C_{XX}(t_1, t_2)$  existirá si  $C_{jj}(t_1, t_2)$  y  $C_{kk}(t_1, t_2)$  son finitos, pues por la desigualdad de Schwartz sabemos que:

$$|C_{jk}(t_1, t_2)|^2 \leq C_{jj}(t_1, t_2)C_{kk}(t_1, t_2), \quad j, k = 1, \dots, n$$

Finalmente definiremos la función de autocorrelación la cual explica la relación lineal existente entre las observaciones de un proceso.

Sean

$$-1 \leq \rho_{jj}(t_1, t_2) = \frac{C_{jj}(t_1, t_2)}{[C_{jj}(t_1, t_1)C_{jj}(t_2, t_2)]^{1/2}} \leq 1$$

y

$$-1 \leq \rho_{jk}(t_1, t_2) = \frac{C_{jk}(t_1, t_2)}{[C_{jj}(t_1, t_1)C_{kk}(t_2, t_2)]^{1/2}} \leq 1$$

entonces,  $\rho_{jj}(t_1, t_2)$  será la función de autocorrelación de  $X_j(t, \omega)$ , la cual mide la relación lineal entre los datos de un mismo proceso a través del tiempo y  $\rho_{jk}(t_1, t_2)$ , será la función de correlación cruzada de  $X_j(t, \omega)$  con  $X_k(t, \omega)$ .

### 3. Estacionariedad y ergodicidad.

Hasta ahora hemos definido a un proceso estocástico y sus características dentro de un contexto muy general; en esta situación es difícil llevar a cabo las tareas de análisis de interés para nosotros, como por ejemplo, la inferencia estadística. Esto nos lleva a imponer ciertas restricciones sobre los procesos a considerar, de tal manera que logremos obtener variables con las que podamos trabajar.

Pediremos entonces que los procesos estocásticos a ser considerados cumplan los principios de estacionariedad y ergodicidad.

Un proceso estocástico es llamado *estrictamente conjuntamente estacionario* si:

$$F_{t_1}(\underline{x}) = F_{t_1+k}(\underline{x})$$
$$F_{t_1, t_2}(\underline{x}, \underline{y}) = F_{t_1+\tau, t_2+\tau}(\underline{x}, \underline{y})$$

y en general

$$F_{t_1, \dots, t_k}(\underline{x}, \underline{y}, \dots, \underline{z}) = F_{t_1+\tau, \dots, t_k+\tau}(\underline{x}, \underline{y}, \dots, \underline{z})$$

para todo  $k, t_1, t_2, \dots, t_k, \tau \in T$  y  $\underline{x}, \underline{y}, \dots, \underline{z} \in \mathbb{R}^n$ .

Es decir, cuando la función de distribución conjunta de  $X(t_1+\tau, \omega), X(t_2+\tau, \omega), \dots, X(t_k+\tau, \omega)$  no depende del origen que se haya elegido, sino de las discrepancias entre  $t_1, t_2, t_3, \dots$ , etc. Cuando  $n=1$  se dice simplemente que el proceso es *estacionario*.

Cuando un conjunto de procesos univariados es conjuntamente estacionario los procesos individuales son estacionarios. Lo opuesto no es necesariamente cierto.

Un proceso estocástico es llamado *estacionario en segundo orden o débilmente estacionario* si:

$$-\infty < C_X(t) = E\{X(t, \omega)\} = C_X < \infty,$$

y

$$-\infty < C_{XX}(t+\tau, t) = \text{cov}\{X(t+\tau, \omega), X(t, \omega)\} = C_{XX}(\tau) < \infty,$$

es decir, cuando los primeros dos momentos del proceso existen, son finitos e independientes del origen seleccionado. Observamos además que:

$$F_{t_1}(\underline{x}) = F_{t_1+k}(\underline{x}) \implies E\{X(t, \omega)\} = C_X$$

$$F_{t_1, t_2}(\underline{x}, \underline{y}) = F_{t_1+\tau, t_2+\tau}(\underline{x}, \underline{y}) \implies \text{cov}\{X(t+\tau, \omega), X(t, \omega)\} = C_{XX}(\tau)$$

De aquí en adelante usaremos el término "estacionariedad" para referirnos a estacionariedad en segundo orden.

Denotaremos a los momentos de un proceso estacionario por:

$$E\{X(t, \omega)\} = C_X$$

y

$$E\{(X(t, \omega) - C_X)(X(t+\tau, \omega) - C_X)\} = C_{XX}(\tau).$$

Respecto al principio de ergodicidad, podemos mencionar que cuando un proceso estocástico estacionario es ergódico, es posible garantizar la existencia de estimadores consistentes de los primeros dos momentos de éste, entre otros.

Más adelante nos interesará obtener estimadores de los primeros momentos de un proceso, por lo que la aplicación de ambos principios será indispensable.

#### 4. Representación espectral.

Consideremos un proceso  $X(t, \omega)$  tal que:

$$E\{X(t, \omega)\} = 0$$

y

$$X(t+s, \omega) = X(t, \omega) \quad \forall t \in T;$$

es decir, un proceso estrictamente periódico con periodo igual a  $s$ . Debido a dicha periodicidad la función de autocovarianza del proceso cumple con:

$$C_{XX}(\tau) = C_{XX}(s-\tau)$$

Este proceso puede ser escrito como una suma de cosenoidales; es decir, de la forma:

$$X(t, \omega) = \sum_j \exp(i\lambda_j t) Z_X(j)$$

$$\lambda_j = 2\pi j/s, \quad j=0, \dots, s.$$

donde  $Z_X(j)$  son vectores no correlacionados para distintos valores de  $j$ .

A tal representación se le conoce como *representación espectral de un proceso* o *representación de Cramér*.

Al escribir un proceso en la forma anterior, logramos descomponerlo como una combinación lineal de variables aleatorias no correlacionadas, cuyos coeficientes están relacionados con frecuencias en el intervalo  $[-\pi, \pi]$ .

Resulta interesante observar lo que sucede con los segundos momentos del proceso en relación a las nuevas variables aleatorias, pues la función de autocovarianza se comportará de la siguiente manera:

$$C_{XX}(0) = \sum_j \text{Var } Z_X(j)$$

$$C_{XX}(\tau) = \sum_j \text{Var } Z_{XX}(j) e^{-i\omega j \tau}, \quad \omega = 2\pi/s;$$

es decir, la varianza del proceso se descompone también en la suma de la varianza de variables no correlacionadas.

Podemos generalizar a procesos no periódicos la descomposición de Cramér, para ello definamos las siguientes funciones:

$$Z_X(\lambda) = \sum_{\frac{2\pi j}{s} \leq \lambda} Z_X(j).$$

$$\Phi_{XX}(\lambda) = \sum_{\frac{2\pi j}{s} \leq \lambda} \text{Var } Z_X(j).$$

Observamos que  $\Phi_{XX}(\lambda)$  es una función escalonada, monótona no decreciente y continua por la derecha tal que:

$$\lim_{\lambda \rightarrow -\infty} \Phi_{XX}(\lambda) = 0$$

$$\lim_{\lambda \rightarrow \infty} \Phi_{XX}(\lambda) = C_{XX}(0) < \infty,$$

es decir,  $\Phi_{XX}(\lambda)$  es una función de distribución con un número finito de discontinuidades en  $[-\pi, \pi]$ .

Si denotamos a:

$$dZ_X(\lambda) = \lim_{h \rightarrow 0^+} [Z(\lambda+h) - Z(\lambda)]$$

y

$$d\Phi_{XX}(\lambda) = \lim_{h \rightarrow 0^+} [\Phi_{XX}(\lambda+h) - \Phi_{XX}(\lambda)]$$

tenemos entonces que:

$$d\Phi_{XX}(\lambda) = E[dZ_X(\lambda), \overline{dZ_X(\nu)}] = \begin{cases} 0, & \lambda \neq \nu \\ d\Phi_{XX}(\lambda), & \lambda = \nu \end{cases}$$

y podemos reescribir a  $X(t, \omega)$  y  $C_{XX}(u)$  como:

$$X(t, \omega) = \int_{-\pi}^{\pi} \exp(i\lambda t) dZ_X(\lambda),$$

$$C_{XX} = \int_{-\pi}^{\pi} \exp(i\lambda t) d\Phi_{XX}(\lambda)$$

las cuales serán válidas independientemente de  $s$ . En este caso tenemos además que:

$$\Phi_{XX}(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_{XX}(\tau) \exp(i\lambda\tau) d\lambda,$$

Sabemos que cualquier función de distribución puede expresarse como

$$F_{XX}(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda),$$

donde  $F_1(\lambda)$  es absolutamente continua,  $F_2(\lambda)$  es una distribución escalonada y  $F_3(\lambda)$  es una función continua con derivada cero en casi todas partes.

Supondremos en adelante que  $\phi_2(\lambda)=0$  y en consecuencia usaremos

$$\phi_{XX}(\lambda) = \frac{d\phi_{XX}(\lambda)}{d\lambda}$$

para denotar a la densidad espectral de  $X(t, \omega)$  en  $[-\pi, \pi]$ .

Hemos introducido de manera poco rigurosa el concepto de densidad espectral, por lo que a continuación formalizaremos.

Sea  $X(t, \omega)$  un proceso estacionario tal que:

$$\sum_{\tau=-\infty}^{\infty} |C_{jk}(\tau)| < \infty, \quad j, k=1, \dots, n \quad (1.4.1)$$

Se define el espectro de segundo orden de  $X_j(t, \omega)$  con  $X_k(t, \omega)$  como:

$$\phi_{jk}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} C_{jk}(\tau) e^{-i\lambda\tau},$$

$$-\infty < \lambda < \infty, \quad j, k = 1, \dots, r.$$

Bajo la condición (1.4.1),  $\phi_{jk}(\lambda)$  es acotado y uniformemente continuo y el hecho de que las componentes de  $X(t, \omega)$  sean valores reales, implica que:

$$\phi_{jk}(\lambda) = \bar{\phi}_{jk}(-\lambda) = \phi_{kj}(-\lambda) = \bar{\phi}_{kj}(\lambda)$$

es decir, es una función no negativa, par y periódica con periodo  $2\pi$  respecto a  $\lambda$ .

El parámetro  $\lambda$  es la frecuencia del espectro. Si  $j=k$ ,  $\phi_{jj}(\lambda)$  será el espectro de  $X_j(t, \omega)$  a la frecuencia  $\lambda$  y si  $j \neq k$ ,  $\phi_{jk}(\lambda)$  será el espectro cruzado de  $X_j(t, \omega)$  con  $X_k(t, \omega)$  a la frecuencia  $\lambda$ .

Nótese que  $\phi_{XX}(\lambda)$  es una matriz Hermitiana definida positiva, es decir,

$$\overline{\phi_{XX}(\lambda)} = \phi_{XX}(\lambda)^t$$

y

$$\alpha^t \phi_{XX}(\lambda) \bar{\alpha} \geq 0$$

para todo vector  $\alpha$  tal que  $|\alpha| > 0$ .

## 5. Filtros.

Definimos un *filtro lineal* como un operador lineal e invariante bajo el tiempo que dado un proceso  $X(t, \omega)$  de dimensión  $r$  lo transforma en otro  $Y(t, \omega)$  de dimensión  $s$ .

Una clase importante de filtros es aquella que toma la forma:

$$Y(t, \omega) = \sum_{\tau=-\infty}^{\infty} a(t-\tau)X(\tau, \omega) = \sum_{\tau=-\infty}^{\infty} a(\tau)X(t-\tau, \omega),$$

donde  $a(\tau)$ , es una matriz de  $n \times n$ . A estos filtros los llamamos *filtros sumables* y los denotamos por  $\{a(\tau)\}$ . Notemos que  $a(\tau)$  no depende de  $t$ ; se dice entonces que el filtro es invariante en el tiempo.

Consideremos el proceso  $Y(t, \omega)$  dado por:

$$Y(t, \omega) = \sum_{\tau=-\infty}^{\infty} a(\tau)X(t-\tau, \omega) = A(L)X_t$$

donde

$$L^j X_t = X_{t-j},$$

$$A(z) = \sum_{\tau=-\infty}^{\infty} a(\tau) z^\tau$$

para  $z$  en alguna región del plano complejo. Además  $X(t, \omega)$  es un proceso estacionario con función de covarianza  $C_{XX}(\tau)$  y matriz de densidad espectral dada por  $\Phi_{XX}(\lambda)$ .

Nos podemos preguntar cuál es la función de covarianza de  $Y(t, \omega)$  y si además existe alguna relación entre los segundos momentos de  $X(t, \omega)$  y  $Y(t, \omega)$ . Si

$$\sum_{\tau=-\infty}^{\infty} |a_{ij}(\tau)|^2 < \infty \quad i, j = 1, \dots, n$$

$Y(t, \omega)$  será también un proceso estacionario, es decir, sus segundos momentos existen. Tratar de obtener la función de covarianza de  $Y(t, \omega)$  de manera directa puede resultar no muy sencillo; sin embargo usando su densidad espectral la situación cambia, tal como se muestra a continuación.

Sin pérdida de generalidad supondremos que  $C_Y = 0$ .

$$\begin{aligned} \int_{-\pi}^{\pi} e^{i\lambda\tau} d\Phi_{YY}(\lambda) &= C_{YY}(\tau) = E[ \underline{Y}(t, \omega) \overline{\underline{Y}(t+\tau, \omega)} ] = \\ &= E \left[ \left[ \sum_{\tau} a(\tau) X(t-\tau, \omega) \right] \overline{\left[ \sum_k a(k) X(t+\tau-k, \omega) \right]} \right] = \\ &= \sum_{\tau} \sum_k a(\tau) E [ (X(t-\tau, \omega)) (\overline{X(t+\tau-k, \omega)}) ] \overline{a(k)}^t = \\ &= \sum_{\tau} \sum_k a(\tau) E \left[ \left[ \int_{-\pi}^{\pi} e^{i\lambda(t-\tau)} dZ_X(\lambda) \right] \overline{\left[ \int_{-\pi}^{\pi} e^{i\nu(t+\tau-k)} dZ_X(\nu) \right]} \right] \overline{a(k)}^t = \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{i(\lambda t - \nu(t-\tau))} \sum_{\tau} a(\tau) e^{-i\lambda\tau} E [ dZ_X(\lambda) \overline{dZ_X(\nu)} ] \sum_k \overline{a(k)}^t e^{i\nu k} = \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{-i\lambda\tau} \sum_{\tau} a(\tau) e^{-i\lambda\tau} d\Phi_{XX}(\lambda) \sum_k \overline{a(k)}^t e^{i\nu k} \end{aligned}$$

En consecuencia la densidad espectral de  $\underline{Y}(t, \omega)$  está dada por:

$$\phi_{YY} = A(\lambda) \phi_{XX}(\lambda) \overline{A(\lambda)}^t$$

A la función

$$A(\lambda) = \int_{\tau} a(\tau) e^{-j\lambda\tau}$$

se le conoce como *función de respuesta frecuencial* .

## II. SERIES DE TIEMPO

### 1. Definición

Sea  $X(t, \omega)$ ,  $t \in T$  y  $\omega \in \Omega$  un proceso estocástico. Para cada valor de  $\omega$ , llamaremos al recorrido de  $X(t, \omega)$  sobre todos los valores del índice  $t$  una *trayectoria*.

Una *serie de tiempo* está constituida por el registro histórico de una trayectoria de  $X$  entre  $t_0$  y  $t_1$ . Es decir, una serie de tiempo es parte de una realización particular de un proceso estocástico, cuyas observaciones están hechas en un cierto intervalo de tamaño finito.

Por ejemplo el registro durante tres años de las temperaturas diarias de una región, corresponde a una parte de la realización de un proceso estocástico.

Una serie de tiempo será *discreta* o *continua* si el proceso que observamos es discreto o continuo. Denotaremos a una serie como:

$$X(t), \quad t \in \{1, 2, \dots, T\}.$$

Las series que consideraremos en el presente trabajo serán series de tiempo *discretas* y *multivariadas*. Supondremos además que éstas provienen de procesos que cumplen con los supuestos de estacionariedad y ergodicidad.

### 2. Estimación de momentos.

Easándonos en las definiciones (1.2.1) y (1.2.2) de los momentos de un proceso y bajo los supuestos anteriormente mencionados, podemos obtener estimadores consistentes de los primeros dos momentos de una serie de tiempo.

Definimos el estimador de  $C_X$  como:

$$\hat{C}_X = \frac{1}{T} \sum_{t=1}^T X(t)$$

En cuanto a  $C_{XX}(\tau)$ , existen básicamente dos maneras de estimarlo:

$$\tilde{C}_{XX}(\tau) = \frac{1}{T-|\tau|} \sum_{t=1}^{T-|\tau|} (X(t+\tau) - \hat{C}_X)(X(t) - \hat{C}_X)$$

$$\hat{C}_{XX}(\tau) = \frac{1}{T} \sum_{t=1}^{T-|\tau|} (X(t+\tau) - \hat{C}_X)(X(t) - \hat{C}_X)$$

Es claro que  $E[\hat{C}_X] = C_X$ , es decir,  $\hat{C}_X$  es un estimador insesgado. En el caso de  $C_{XX}$  cuando  $C_X$  es desconocida, ambos estimadores serán asintóticamente insesgados y en el caso de que  $C_X = 0$ ,  $\tilde{C}_{XX}$  será un estimador insesgado.<sup>1</sup>

### 3. Estimación de la densidad espectral.

Sea  $X(t)$  una serie de tiempo de dimensión  $n$  tal que:

$$E[X(t)] = C_X$$

y

$$E[(X(t) - C_X)(X(t-\tau) - C_X)] = C_{XX}(\tau)$$

definimos el *periodograma* de  $X(t)$  como la matriz  $I_{XX}(\lambda)$  de  $n \times n$  tal

<sup>1</sup> Una discusión, mas amplia sobre el comportamiento de estos estimadores, se encuentra en el libro de Anderson, 1971, pp. 448-449, 463-464.

que:

$$I_{XX}(\lambda) = \frac{1}{2\pi T} \left[ \sum_{s=1}^T X(s) e^{-i\lambda s} \right] \left[ \sum_{s=1}^T X(s) e^{i\lambda s} \right]^t$$

la cual corresponde al producto de la transformada de Fourier de la serie por el conjugado transpuesto de la misma, salvo por el factor  $1/2\pi T$ .

El periodograma de una serie mide la asociación de tipo lineal entre la función estocástica  $X(t)$  y las funciones  $\text{sen}(\lambda t)$  y  $\text{cos}(\lambda t)$  ( $0 \leq \lambda \leq \pi$ ).

Sabemos que el espectro de una serie queda definido como:

$$\Phi_{XX}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} C_{XX}(\tau) \exp(-i\lambda\tau)$$

donde  $C_{XX}(\tau)$  es la función de autocovarianza. Esto último sugiere un camino para tratar de estimar  $\Phi_{XX}(\lambda)$ , para ello debemos considerar un estimador de  $C_{XX}(\tau)$ . Sin pérdida de generalidad supondremos que  $C_X=0$ .

Sea

$$\hat{C}_{XX}(\tau) = \frac{1}{T} \sum_{t=1}^{T-|\tau|} X(t+\tau)X(t)$$

entonces

$$\hat{\Phi}_{XX}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-(T-1)}^{T-1} \hat{C}_{XX}(\tau) \exp(-i\lambda\tau).$$

Si sustituimos el valor de  $\hat{C}_{XX}(\tau)$  en la expresión anterior tenemos que:

$$\begin{aligned}
\Phi_{XX}(\lambda) &= \frac{1}{2\pi T} \sum_{\tau=-(T-1)}^{T-1} \exp(-i\lambda\tau - i\lambda\tau) \sum_{t=0}^{T-|\tau|} X(t+\tau)X(t) = \\
&= \frac{1}{2\pi T} \sum_{s=1}^T \sum_{t=1}^T \exp(-i\lambda(t-s)) X(t)X(s) = \\
&= \frac{1}{2\pi T} \left[ \sum_{t=1}^T X(t) \exp(i\lambda t) \right] \left[ \sum_{s=1}^T X(s) \exp(-i\lambda s) \right] = I_{XX}(\lambda)
\end{aligned}$$

Es decir, un primer estimador de  $\Phi_{XX}(\lambda)$  es el periodograma de la serie, el cual tiene las propiedades de simetría y no negatividad de la matriz de densidad espectral. Este estimador es asintóticamente insesgado, sin embargo es posible probar que independientemente del tamaño de  $T$ , no es un estimador consistente de  $\Phi_{XX}(\lambda)$ .

El problema al estimar  $\Phi_{XX}(\lambda)$  a través del periodograma consiste en la falta de precisión con que se estiman las autocovarianzas asociadas a rezagos grandes, pues mientras que para  $\tau=1$  el estimador de  $C_{XX}(\tau)$  es un promedio de  $t-1$  productos; para  $\tau=T-1$ ,  $\hat{C}_{XX}(\tau)$  resulta un promedio de un sólo producto y el hecho de que en  $I_{XX}(\lambda)$  todos los estimadores de las autocovarianzas reciban igual ponderación, incrementa la varianza de los estimadores de  $\Phi_{XX}(\lambda)$  para las frecuencias elegidas.

Una manera de resolver este problema es aplicando una función ponderadora  $w(m, \tau)$  a la función de autocovarianza que asigne ponderaciones nulas a  $\hat{C}_{XX}(\tau)$  a partir del rezago  $m$  en  $I_{XX}(\lambda)$ .

Tenemos entonces que si

$$\hat{\Phi}_{XX}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-(T-1)}^{T-1} \hat{C}_{XX}(\tau) \exp(-i\lambda\tau)$$

donde

$$\hat{C}_{XX}(\tau) = \frac{1}{T} \sum_{t=1}^{T-|\tau|} X(t+\tau)X(t),$$

aplicando la función  $w(m, \tau)$  a  $\hat{C}_{XX}(\tau)$  podemos reescribir a  $\hat{\Phi}_{XX}(\lambda)$  como:

$$\bar{\Phi}_{XX}(\lambda) = \frac{1}{2\pi} \int_{\tau=-\infty}^{\infty} w(m, \tau) \hat{C}_{XX}(\tau) \exp(-i\lambda\tau) d\tau$$

donde  $m$  es una constante arbitraria y  $w(m, \tau)$  es tal que:

$$w(m, \tau) = w(m, -\tau),$$

y además su valor será cero cuando  $\tau \geq m$ .

$\bar{\Phi}_{XX}(\lambda)$  puede ser reescrita a su vez como:

$$\begin{aligned} \bar{\Phi}_{XX}(\omega) &= \frac{1}{2\pi} \int_{\tau=-\infty}^{\infty} \left[ w(m, \tau) \int_{-\pi}^{\pi} \hat{\Phi}_{XX}(\omega) \exp(i\omega\tau) d\omega \right] \exp(-i\lambda\tau) d\tau = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\Phi}_{XX}(\omega) \left[ \int_{\tau=-\infty}^{\infty} w(m, \tau) \exp(-i\tau(\lambda-\omega)) d\tau \right] d\omega \end{aligned}$$

si hacemos

$$W_m(\lambda) = \frac{1}{2\pi} \int_{\tau=-\infty}^{\infty} w(m, \tau) \exp(-i\lambda\tau) d\tau$$

tenemos que:

$$\bar{\Phi}_{XX}(\lambda) = \int_{-\pi}^{\pi} \hat{\Phi}_{XX}(\omega) W_m(\lambda-\omega) d\omega$$

A la función  $W_m(\lambda)$  se le conoce como *ventana espectral* y al valor de  $m$  como la *anchura de la ventana*. Las ventanas más usuales son:

1. Ventana de *Bartlett modificada*:

$$W_m(\lambda) = \frac{1}{2\pi m} \left[ \frac{\text{sen}(m\lambda/2)}{\text{sen}(\lambda/2)} \right]^2 = F_m(\lambda)$$

2. Ventana de Daniell:

$$W_n(\lambda) = \begin{cases} n/2\pi, & -\pi/n \leq \lambda \leq \pi/n \\ 0, & \text{en otro caso.} \end{cases}$$

3. Ventana de Tukey:

$$W_n(\lambda) = aD_n(\lambda - \pi/m) + (1-2a)D_n(\lambda) + aD_n(\lambda + \pi/m),$$

con  $0 < a \leq 0.25$  y donde

$$D_n(\lambda) = \frac{\text{sen}(n+1/2)\lambda}{\text{sen}(\lambda/2)}$$

Cuando  $a=0.23$  tenemos la ventana de Tukey-Hamming y cuando  $a=0.25$ , tenemos la ventana de Tukey-Hanning.

4. Ventana de Parzen:

$$W_m(\lambda) = \frac{6\pi}{m} [F_{m/2}(\lambda)]^2 \left[ 1 - \frac{2}{3} \text{sen}^2(\lambda/2) \right]$$

donde  $m$  es par; y en caso de ser non, entonces se usarán  $m+1$  valores en lugar de  $m$ .

5. Ventana de Bartlett-Priestley:

$$W_n(\lambda) = \begin{cases} \frac{3m}{4\pi} \left\{ 1 - (m\lambda/\pi)^2 \right\}, & |\lambda| \leq \pi/m \\ 0, & |\lambda| > \pi/m. \end{cases}$$

Si pedimos que la función  $W_m(\lambda)$  cumpla la siguiente restricción:

$$\int_{-\infty}^{\infty} W(\alpha) \delta\alpha = 1$$

lograremos entonces obtener un nuevo estimador correspondiente a un promedio de valores del periodograma alrededor de una frecuencia determinada. Es decir, logramos suavizar el periodograma.

De aquí que el estimador de la matriz de densidad espectral estará dado por:

$$\bar{\phi}_{XX}(\lambda) = \frac{1}{2\pi} \sum_{s=0}^{T-1} W_m \left[ \lambda - \frac{2\pi s}{T} \right] I_{XX} \left[ \frac{2\pi s}{T} \right]$$

el cual será un estimador consistente que conserva las propiedades de simetría y periodicidad de  $\phi_{XX}(\lambda)$ , resultando una matriz definida positiva.

Para toda ventana, el valor de  $m$  influirá en el comportamiento de éste estimador. Para valores de  $m$  muy grandes el sesgo se reduce, pero la varianza aumenta; para valores de  $m$  menores, el sesgo aumenta y la varianza disminuye.

Una alternativa para estimar la densidad espectral de una serie, es usar el enfoque de Blackman-Tukey llamado "pre-blanqueo" de la serie. Este consiste en filtrar las observaciones de modo que el espectro de la serie filtrada resulte casi constante. El espectro de esta serie puede ser estimado usando alguna de las ventanas ya mencionadas; esta estimación es corregida por el efecto del filtro original, para obtener así el espectro de la serie de interés. Es decir, sea  $X(t)$  la serie cuyo espectro queremos estimar, calculamos entonces:

$$Y(t, \omega) = \int_{\tau} a(\tau) X(t-\tau)$$

procurando escoger  $\{a(\tau)\}$  de modo que  $\phi_{YY}(\lambda) = cte$ . A partir de  $Y(t)$  estimamos  $\phi_{YY}(\lambda)$  y entonces

$$\phi_{XX} = \overline{A(\lambda)}^t \phi_{YY}(\lambda) A(\lambda)$$

A esta última operación se le da el nombre de "recolorrear".

Y CORRELACION CANONICA

I. DATOS NO CORRELACIONADOS

1. Componentes principales.

Sea  $\underline{X}$  un vector aleatorio n-dimensional con matriz de covarianza  $\Sigma$ . En esta sección nos concentraremos en encontrar una transformación lineal de  $\underline{X}$  que resuma toda la información relativa a la variabilidad de sus componentes. Es decir, buscaremos una matriz  $A=(a_1, \dots, a_n)$  que transformando a  $\underline{X}$  de la siguiente manera:

$$\underline{Y} = A \underline{X}$$

resuelva los siguientes problemas:

1.- Encontrar  $Y_1 = a_1'X$ , con  $a_1'a_1=1$ , tal que su varianza sea la mayor de entre todas las combinaciones lineales de las componentes de  $X$

2.- Encontrar  $Y_2 = a_2'X$ , con  $a_2'a_2=1$ , tal que su varianza sea la mayor de entre todas las combinaciones lineales de las componentes de  $\underline{X}$  y que no esté correlacionada con la combinación lineal solución al problema 1.

n.- Encontrar  $Y_n = a_n'X$ , con  $a_n'a_n=1$ , tal que su varianza sea la mayor de entre todas las combinaciones lineales de las componentes de  $\underline{X}$  under que no estén correlacionadas con las combinaciones lineales que solucionan los problemas  $1 \leq k < n$

Una manera de solucionar estos problemas la encontraremos aplicando multiplicadores de Lagrange a cada uno de ellos. Por ejemplo, para el problema 1 tenemos la siguiente situación:

$$\begin{aligned} & \text{Max } \underline{a}'_1 \Sigma \underline{a}_1 \\ \text{s. a.} & \\ & \underline{a}'_1 \underline{a}_1 = 1, \end{aligned}$$

y la solución la encontramos de la siguiente manera:

Sea

$$\theta = \underline{a}'_1 \Sigma \underline{a}_1 - \nu (\underline{a}'_1 \underline{a}_1 - 1)$$

donde  $\nu$  es un multiplicador de Lagrange. El vector de derivadas parciales de  $\theta$  está dado por:

$$\frac{\partial \theta}{\partial \underline{a}_1} = 2 \Sigma \underline{a}'_1 - 2 \nu \underline{a}_1 = 0$$

el cual debe satisfacer:

$$(\Sigma - \nu I) \underline{a}_1 = 0$$

y de aquí que:

$$|\Sigma - \nu I| = 0$$

La función  $|\Sigma - \nu I|$  es un polinomio en  $\nu$  de orden  $n$ , que tendrá  $n$  raíces a las que denotamos  $\nu_1 \geq \nu_2 \geq \dots \geq \nu_n$ .  $\underline{a}_1$  corresponderá al vector propio asociado al valor de  $\nu_1$ , ya que  $\text{Var}(\underline{a}'_1 X) = \underline{a}'_1 \Sigma \underline{a}_1 = \nu_1$ .

Podemos seguir este procedimiento para todo  $\underline{a}_j$ ,  $j=2, \dots, n$  hasta construir el siguiente teorema:

**Teorema:**

Sea  $X$  un vector aleatorio con matriz de covarianza  $\Sigma$ . Existe una transformación lineal ortogonal

$$\underline{Y} = A \underline{X}$$

tal que la matriz de covarianza de  $\underline{Y}$  es:

$$N = \begin{bmatrix} \nu_1 & 0 & \dots & 0 \\ 0 & \nu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \nu_n \end{bmatrix}$$

donde  $\nu_1 \geq \nu_2 \geq \dots \geq \nu_n \geq 0$  son los valores propios de la matriz  $\Sigma$ . La  $r$ -ésima columna de  $A$ ,  $\underline{a}_r$ , satisface  $(\Sigma - \nu_r I)\underline{a}_r = 0$ . La  $r$ -ésima componente de  $\underline{Y}$ ,  $Y_r = \underline{a}_r' \underline{X}$ , tiene la máxima varianza de todas las combinaciones lineales de  $\underline{X}$  y no está correlacionada con  $Y_1, \dots, Y_{r-1}$ .

Esto nos dice que la matriz  $A$  que buscábamos, no es otra que la matriz de vectores propios asociados a los respectivos valores propios ordenados de  $\Sigma$ .

Al vector  $\underline{Y}$  se le conoce como el *vector de componentes principales* de  $\underline{X}$ .

El hecho de que las componentes de  $\underline{Y}$  estén no correlacionadas implica que el conocer el valor de una de ellas no nos ayuda a explicar, al menos linealmente, el comportamiento de las otras. Es decir, se reducen las influencias lineales mutuas, volviéndose más sencilla la matriz de covarianzas de la transformación.

Ya que  $\nu_1 \geq \nu_2 \geq \dots \geq \nu_n \geq 0$ , es posible que a partir de algún valor de  $j$ ,  $\nu_j$  sea casi cero; la información que las correspondientes componentes de  $\underline{Y}$  nos proporcionan será muy pequeña, así si consideramos en  $\underline{Y}$  únicamente aquellas componentes para las cuales  $\nu_j > 0$ , éstas resumirán casi toda la información del problema original pero en dimensión más pequeña. Siempre que en  $\underline{X}$  haya información redundante, es decir, información altamente correlacionada, estos valores propios muy pequeños existirán.

## 2. Correlación canónica.

Sea  $\underline{X}$  un vector aleatorio  $n$ -dimensional y  $\underline{Y}$  un vector aleatorio  $m$ -dimensional ( $n \leq m$ ), con matriz de covarianza

$$\begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

En esta sección buscaremos las transformaciones que maximicen la correlación entre combinaciones lineales de las componentes de  $\underline{X}$  y  $\underline{Y}$ . Es decir, buscaremos las matrices  $A=(a_1, \dots, a_n)$  y  $B=(b_1, \dots, b_n)$  que transformando a los vectores  $\underline{X}$  y  $\underline{Y}$  de la siguiente manera:

$$\underline{U} = A \underline{X},$$

$$\underline{V} = B \underline{Y}$$

resuelvan los siguientes problemas:

1.- Encontrar aquellas combinaciones lineales de  $\underline{X}$  y  $\underline{Y}$  cuya correlación sea máxima.

2.- Encontrar aquellas combinaciones lineales de  $\underline{X}$  y  $\underline{Y}$  cuya correlación sea máxima y que no estén correlacionadas con la solución del problema 1.

n.- Encontrar aquellas combinaciones lineales de  $\underline{X}$  y  $\underline{Y}$  cuya correlación sea máxima y que no estén correlacionadas con combinaciones lineales anteriores.

Al igual que en la sección anterior, encontraremos la solución a estos problemas, aplicando multiplicadores de Lagrange a cada uno de ellos. Por ejemplo para el primer problema tenemos lo siguiente:

$$\text{Max } \underline{a}'_1 \Sigma_{12} \underline{b}_1$$

s. a.

$$\underline{a}'_1 \Sigma_{11} \underline{a}_1 = 1$$

$$\underline{b}'_1 \Sigma_{22} \underline{b}_1 = 1$$

y la solución la encontramos de la siguiente manera:

Sea

$$\theta = \underline{a}'_1 \Sigma_{12} \underline{b}_1 - \frac{1}{2} \nu (\underline{a}'_1 \Sigma_{11} \underline{a}_1 - 1) - \frac{1}{2} \mu (\underline{b}'_1 \Sigma_{22} \underline{b}_1 - 1)$$

donde  $\nu$  y  $\mu$  son multiplicadores de lagrange. Los vectores de derivadas parciales de  $\theta$  están dados por:

$$\frac{\partial \theta}{\partial \underline{a}_1} = \Sigma_{12} \underline{b}_1 - \nu \Sigma_{11} \underline{a}_1 = 0$$

y

$$\frac{\partial \theta}{\partial \underline{b}_1} = \Sigma_{12} \underline{a}_1 - \mu \Sigma_{22} \underline{b}_1 = 0$$

los cuales satisfacen:

$$\underline{a}'_1 \Sigma_{12} \underline{b}_1 - \nu \underline{a}'_1 \Sigma_{11} \underline{a}_1 = 0$$

$$\underline{b}'_1 \Sigma_{12} \underline{a}_1 - \mu \underline{b}'_1 \Sigma_{22} \underline{b}_1 = 0.$$

Como  $\underline{a}'_1 \Sigma_{11} \underline{a}_1 = 1$  y  $\underline{b}'_1 \Sigma_{22} \underline{b}_1 = 1$ , se tiene que  $\nu = \mu = \underline{a}'_1 \Sigma_{12} \underline{b}_1$ . La solución a todos los problemas consistirá entonces en encontrar aquellas  $\nu$ 's tales que:

$$\left| \nu^2 \Sigma_{XX} - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \right| = 0$$

ó

$$\left| \nu^2 \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \right| = 0$$

y para cada  $\nu \neq 0$ , resolver

$$(\nu_j^2 \Sigma_{XX} - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}) a_j = 0$$

$$(\nu_j^2 \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}) b_j = 0 \quad j=1, \dots, n$$

restringidas a  $a_i' \Sigma_{11} a_j = b_i' \Sigma_{22} b_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

De aquí que la matriz  $A$  que buscábamos corresponde a la matriz de vectores propios asociados a los respectivos  $n$  valores propios ordenados de la matriz:

$$\Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$$

en la métrica definida por  $\Sigma_{XX}$ . La matriz  $B$  corresponderá a la matriz de vectores propios asociados a los respectivos  $m$  valores propios ordenados de la matriz:

$$\Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$$

en la métrica definida por  $\Sigma_{YY}$ ; en este caso  $\nu_j = 0$  para  $j > n$ .

La matriz de covarianza de la transformación estará dada por:

$$\text{Cov} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} I & N \\ N & I \end{bmatrix}$$

donde

$$N = \begin{bmatrix} \nu_1 & 0 & \dots & 0 \\ 0 & \nu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \nu_n \end{bmatrix}$$

A las variables  $u_j$  y  $v_j$  se les denomina *variables canónicas*.

Debido a las propiedades de las transformaciones que resuelven los problemas antes mencionados,  $\underline{U}$  y  $\underline{V}$  serán una nueva descripción del fenómeno estudiado, pero con una estructura de covarianza más simple. Por otro lado las relaciones que con frecuencia se busca establecer entre combinaciones lineales de variables de un conjunto con combinaciones de variables en el otro, adquieren ahora una representación más sencilla y se caracteriza a aquellas que guardan una relación más estrecha, es decir, se gana en simplicidad en el planteamiento de las relaciones, imponiendo restricciones estadísticamente razonables. Más aún, es de esperarse que si  $\underline{X}$  y  $\underline{Y}$  explican el mismo fenómeno de manera similar, el vector de dimensión mayor contenga información redundante, misma que al ser transformada, se detecta y elimina con una pérdida despreciable de información.

## II. SERIES DE TIEMPO

En esta sección generalizaremos los conceptos de componentes principales y variables canónicas al caso de series de tiempo. Aplicaremos los métodos descritos en las secciones anteriores tomando en cuenta la correlación existente en las observaciones, para lo cual será necesario utilizar la representación espectral de la series.

### 1. Componentes principales.

Sea  $X(t)$  una serie de tiempo estacionaria de dimensión  $n$  tal que:

$$C_{XX}(\tau) = E [(X(t) - C_X) (X(t+\tau) - C_X)]$$

y

$$\phi_{XX}(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_{XX}(\tau) \exp(-i\lambda\tau)$$

Queremos encontrar una transformación de  $X(t)$  que resuma la información sobre la variabilidad de sus componentes, eliminando las influencias lineales mutuas. A diferencia del caso anterior, en el caso de series de tiempo las observaciones sucesivas están autocorrelacionadas. Lo anterior nos obliga a modificar el procedimiento descrito en la sección 1.1.

Consideremos la siguiente transformación de  $X(t)$ :

$$Y(t) = \int_{\tau} a(t-\tau) X(\tau)$$

Para que las influencias lineales mutuas sean eliminadas,  $Y(t)$  debe ser tal que:

$$\text{Cov } [Y(t)] = \begin{bmatrix} \Delta_0 & \Delta_1 & \dots & \Delta_{n-1} \\ \Delta_{-1} & \Delta_0 & \dots & \Delta_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{-n+1} & \Delta_{-n+2} & \dots & \Delta_0 \end{bmatrix}$$

donde cada una de las matrices  $\Delta_j$  es diagonal. Es decir, la transformación que buscamos debe ser tal que diagonalice no solamente a  $\Delta_0$  sino a todas las  $\Delta_j$ . Tratar de hacer esto de manera directa es prácticamente imposible; sin embargo utilizando la representación espectral de la serie es posible dar solución a este problema, tal como se muestra a continuación.

Consideremos la representación espectral de  $X(t)$ , la cual está dada por:

$$X(t) = \int_{-\pi}^{\pi} \exp(i\lambda t) dZ_X(\lambda),$$

donde  $dZ_X(\lambda)$  es tal que:

$$\text{Cov } [dZ_X(\lambda), \overline{dZ_X(\nu)}] = \begin{cases} 0, & \lambda \neq \nu \\ \phi(\lambda), & \lambda = \nu. \end{cases}$$

Entonces si formamos el vector  $Z = (dZ(\lambda_1), \dots, dZ(\lambda_n))$ , se tiene que:

$$\text{Cov } (\underline{Z}) = \begin{bmatrix} \phi_{XX}(\lambda_1) & 0 & \dots & 0 \\ 0 & \phi_{XX}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{XX}(\lambda_n) \end{bmatrix}$$

El nuevo proceso tiene una estructura de covarianzas más simple y además para cada valor de  $\lambda$  es posible encontrar una matriz  $B(\lambda)$  tal que:

$$\Delta(\lambda) = B(\lambda) \phi_{XX}(\lambda) \bar{B}^t(\lambda) = \text{Cov} [(B(\lambda) dZ_X(\lambda))]$$

donde  $\Delta(\lambda)$  es una matriz diagonal. La matriz  $B(\lambda)$  corresponderá a la matriz de vectores propios asociados a los respectivos valores propios de la matriz  $\phi_{XX}(\lambda)$ .

Definimos entonces a:

$$Y(t) = \int_{-\pi}^{\pi} \exp(i\lambda t) B(\lambda) dZ_X(\lambda),$$

como la series de *componentes principales* de  $X(t)$ ; ya que su matriz de densidad espectral es  $\Delta(\lambda)$ , se tiene que:

$$C_{YY}(\tau) = \int_{-\pi}^{\pi} \Delta(\lambda) \exp(-i\lambda\tau) d\lambda,$$

la cual es una matriz diagonal.

En vista de que entre las componentes de  $Y(t)$  no existen influencias lineales mutuas y autocorrelación, es posible ajustar modelos univariados a cada una de ellas con el fin de extrapolarias.

## 2. Correlación canónica.

Sea

$$\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}$$

un vector de series de tiempo estacionarias de dimensión  $n+m$ , con función de autocovarianza

$$\begin{bmatrix} C_{XX}(t) \\ C_{YY}(t) \end{bmatrix}$$

y matriz de densidad espectral:

$$\begin{bmatrix} \phi_{XX}(\lambda) & \phi_{XY}(\lambda) \\ \phi_{YX}(\lambda) & \phi_{YY}(\lambda) \end{bmatrix}$$

Por último definiremos el vector de variables canónicas de  $(X'(t), Y'(t))$ , utilizando la representación espectral conjunta del vector, la cual está dada por:

$$X(t) = \int_{-\pi}^{\pi} \exp(i\lambda t) dZ_X(\lambda),$$

$$Y(t) = \int_{-\pi}^{\pi} \exp(i\lambda t) dZ_Y(\lambda)$$

donde  $dZ_X$  y  $dZ_Y$  tienen incrementos no correlacionados; es decir,

$$\text{Cov} [ dZ_X, dZ_Y ] = \delta(\lambda - \nu) \phi(\lambda)$$

donde  $\delta(0)=1$  y cero en otro caso.

En consecuencia, la estructura de correlación de los incrementos es más simple y está dada por:

$$\text{Cov} \begin{bmatrix} dZ_X(\lambda) \\ dZ_Y(\lambda) \end{bmatrix} = \begin{bmatrix} \phi_{XX}(\lambda) & \phi_{XY}(\lambda) \\ \phi_{YX}(\lambda) & \phi_{YY}(\lambda) \end{bmatrix}.$$

Para cada valor de  $\lambda$  existen  $A(\lambda)$  y  $B(\lambda)$  tales que:

$$dZ_U(\lambda) = A(\lambda) dZ_X$$

$$dZ_V(\lambda) = B(\lambda) dZ_Y$$

las cuales tienen una matriz de covarianza dada por:

$$\text{Cov} \begin{bmatrix} dZ_U(\lambda) \\ dZ_V(\lambda) \end{bmatrix} = \begin{bmatrix} I & \phi_{UV}(\lambda) \\ \phi_{VU}(\lambda) & I \end{bmatrix}$$

donde las matrices  $\phi(\lambda)$  son diagonales y deberán ser determinadas para cada valor de  $\lambda$ .

Las series  $U(t)$  y  $V(t)$ , definidas como:

$$U(t) = \int_{-\pi}^{\pi} \exp(i\lambda t) dZ_U(\lambda),$$

$$V(t) = \int_{-\pi}^{\pi} \exp(i\lambda t) dZ_V(\lambda)$$

corresponden a las series canónicas de  $X(t)$  y  $Y(t)$ , respectivamente.

## EJEMPLO

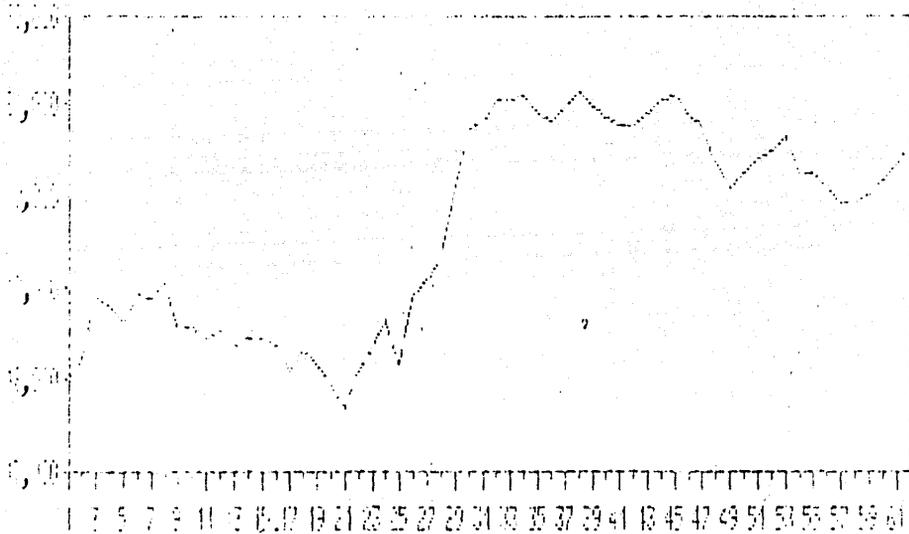
Para ejemplificar la metodología anteriormente desarrollada, utilizaremos una serie correspondiente al registro del precio de cierre diario de 12 acciones de la Bolsa Mexicana de Valores durante el primer trimestre de 1990. La selección de estas acciones se hizo de manera aleatoria lo cual puede causar heterogeneidad en su comportamiento.

Los datos fueron analizados en un programa escrito en fortran 77 apoyado por rutinas estadísticas de IMSL. Los resultados que se presentan corresponden a las series de componentes principales de estas acciones, además de una tabla comparativa de la varianza de las series originales y las componentes principales.

A continuación presentamos el nombre de las acciones utilizadas y sus gráficas, así como las gráficas de las componentes principales.

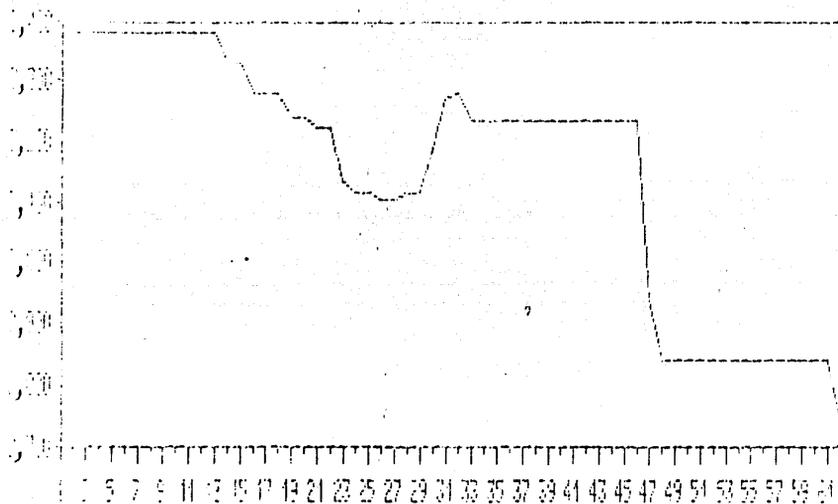
- Acción 1 : Alfa-A.
- Acción 2 : Apasco-AF.
- Acción 3 : Carbide-BICPR.
- Acción 4 : Celanes-A.
- Acción 5 : Cemex-A.
- Acción 6 : Cifra-ACP.
- Acción 7 : Desc-D.
- Acción 8 : Ericson-B.
- Acción 9 : Frisco-I.
- Acción 10 : San Luis-AI.
- Acción 11 : Visa.
- Acción 12 : Vitro.

SECRET



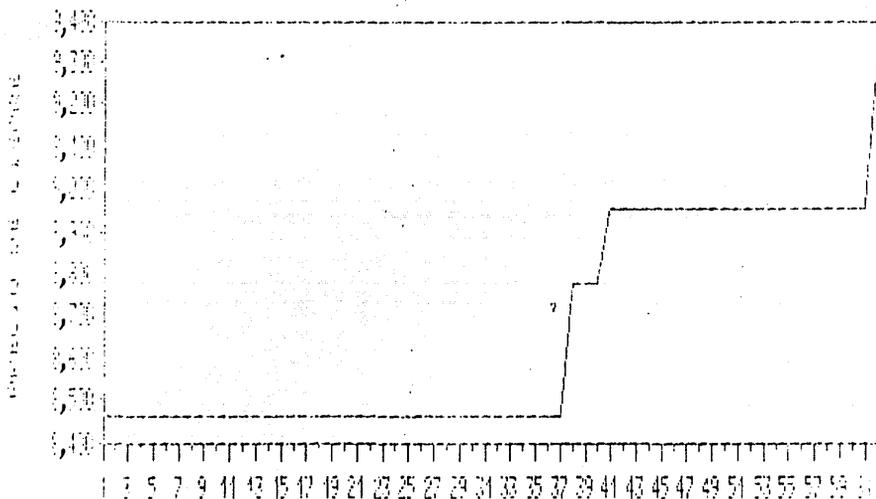
0185

SITE OF CIPOL 2



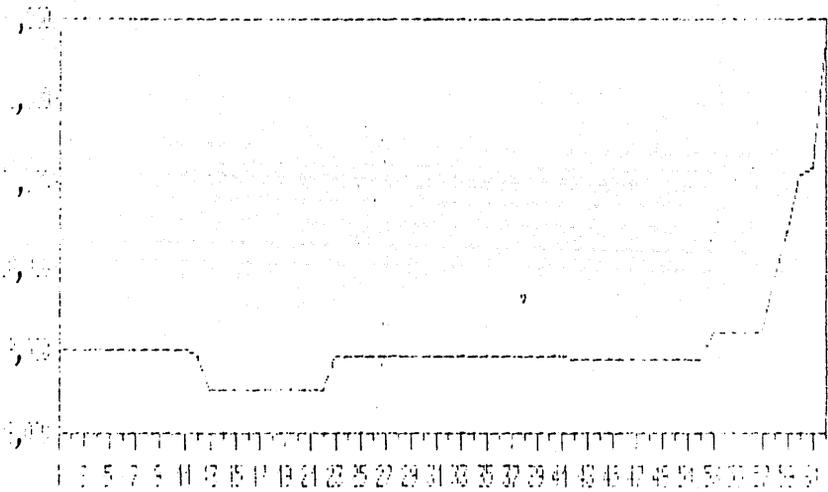
DAYS

SERIE DE ORIGINAL 3



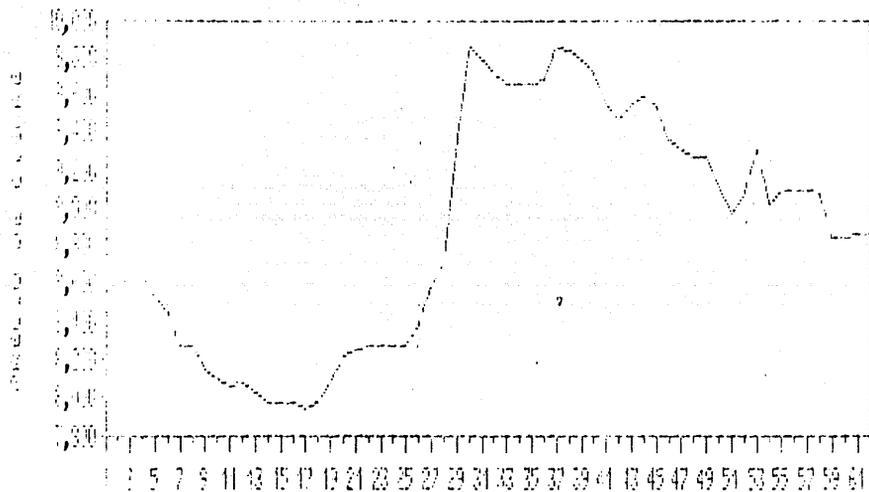
D I A S

SERIE ORIGINAL 4



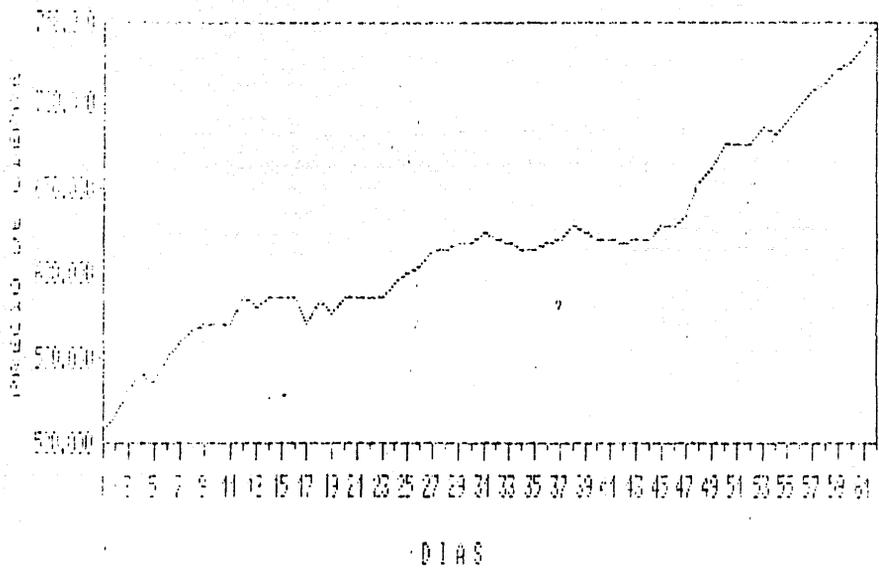
D I H S

SERIE ORIGINAL 5

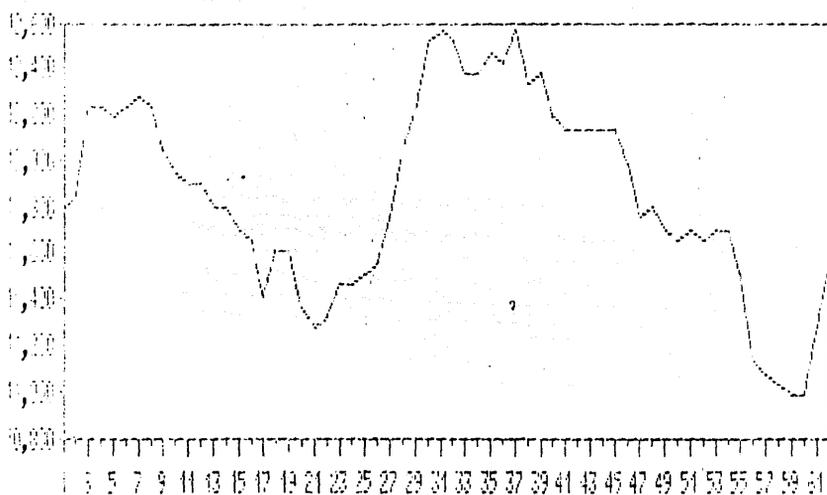


DIAS

SERIE ORIGINAL 6

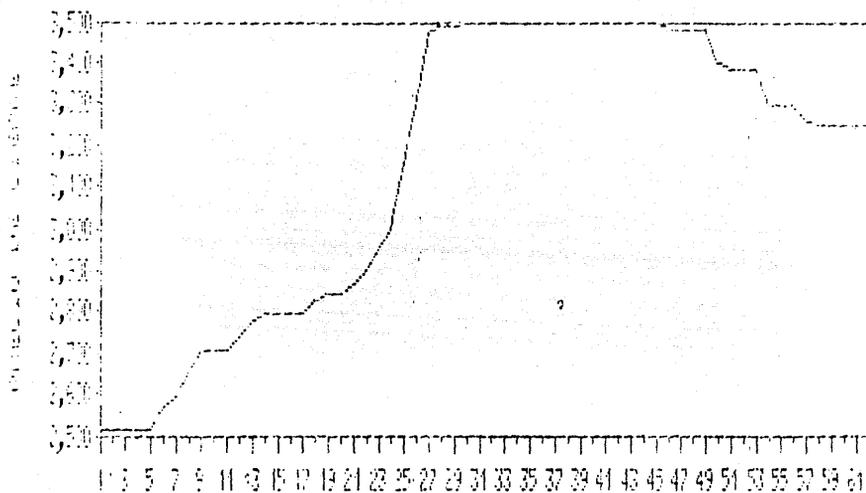


## SERIE ORIGINAL 7



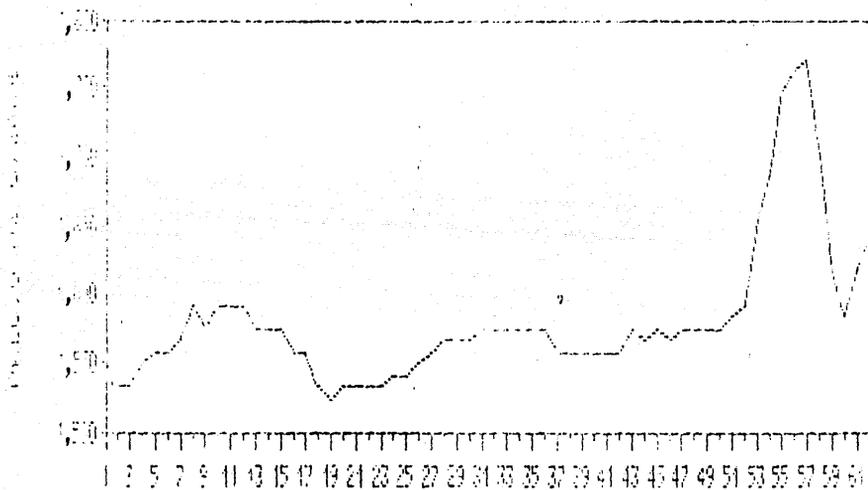
DÍAS

SERIE ORIGINAL 8



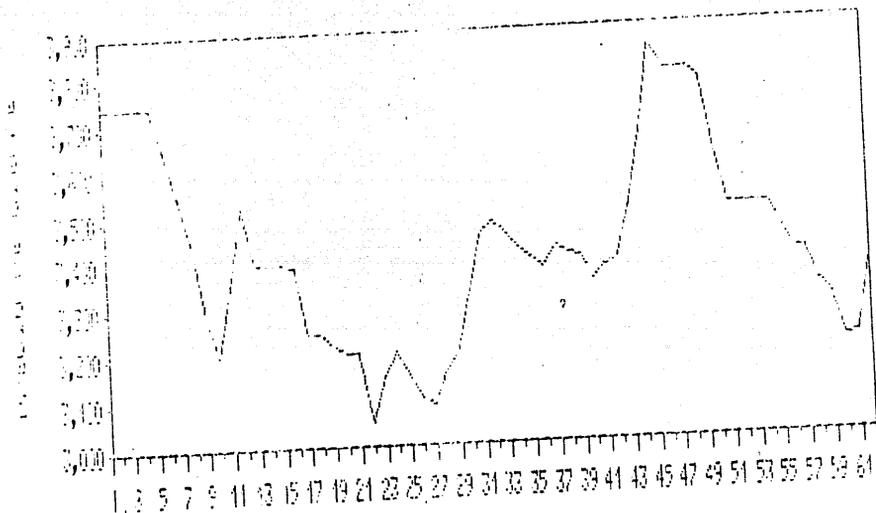
Dias

SERIE ORIGINAL 3



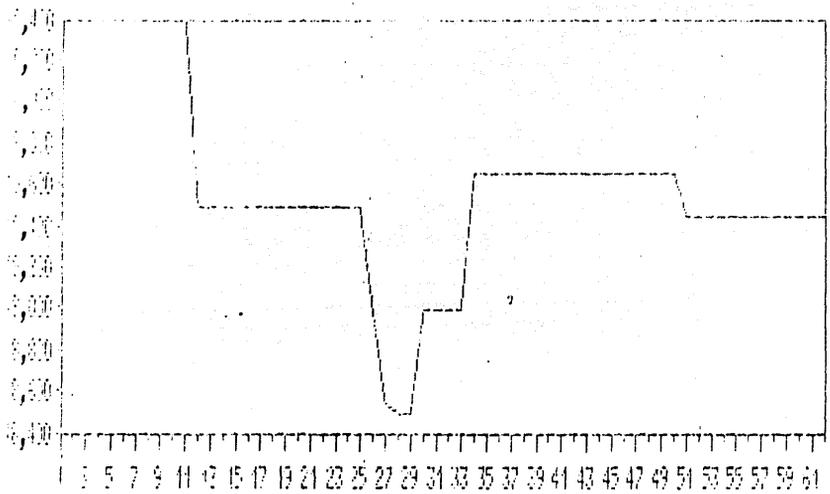
D I A S

SERIE ORIGINAL 10



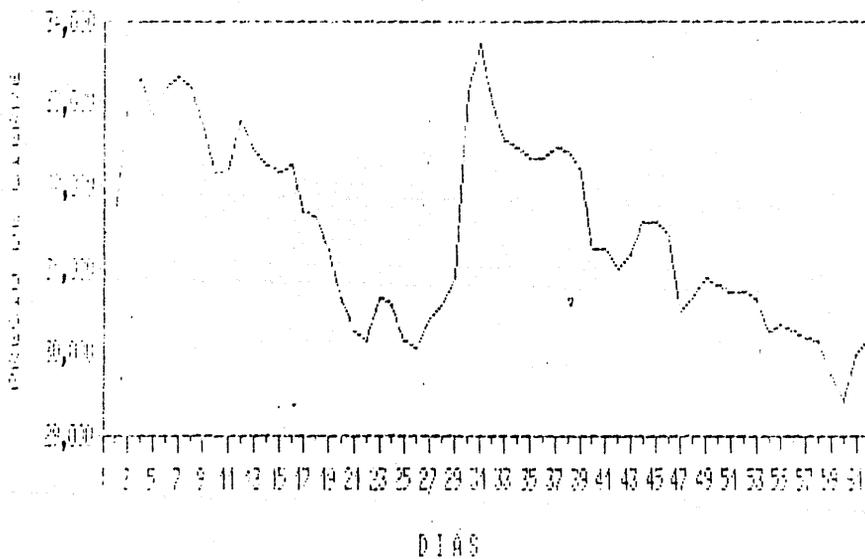
D I A S

SCORE ORIGINAL 11

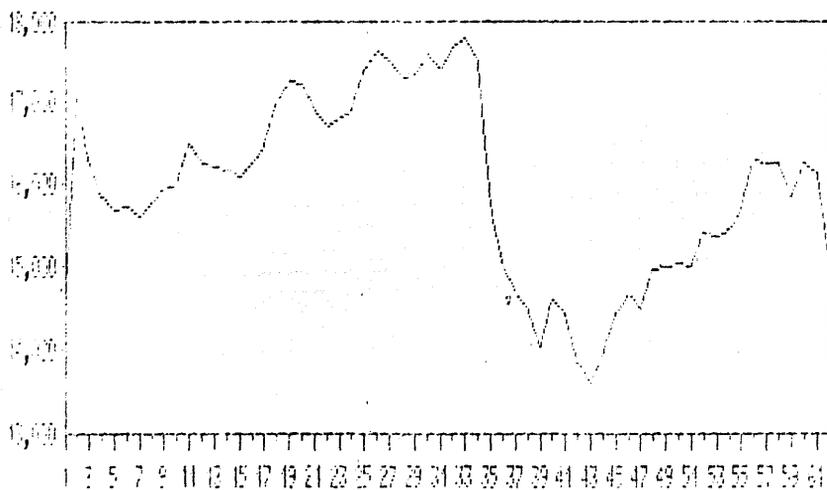


DIAS

SERIE ORIGINAL 12

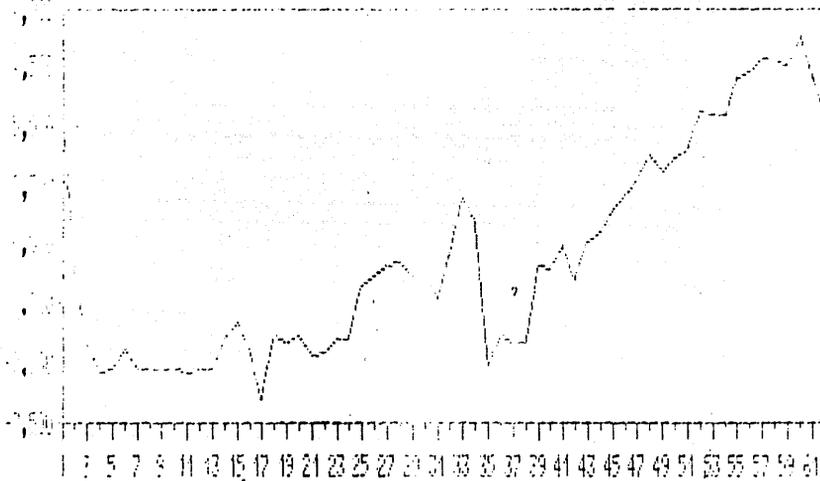


PRIMERA CONDICIÓN PRINCIPAL



DÍAS

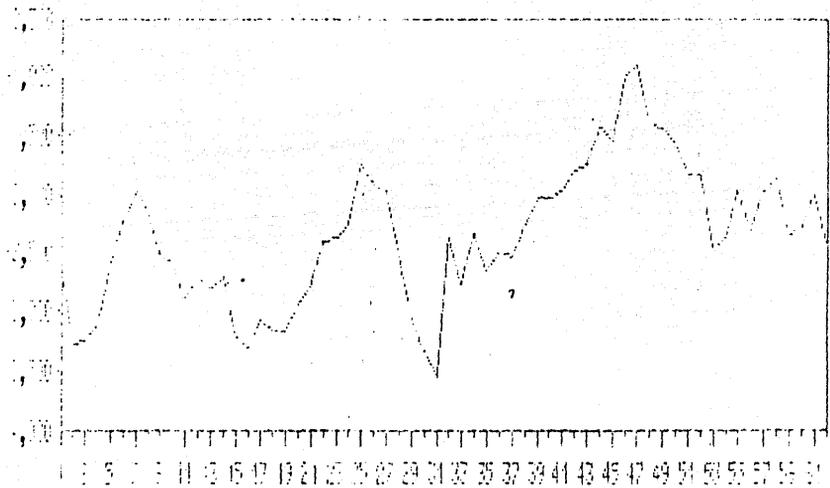
SEGUNDA COMPONENTE TRIMESTRAL



DÍAS

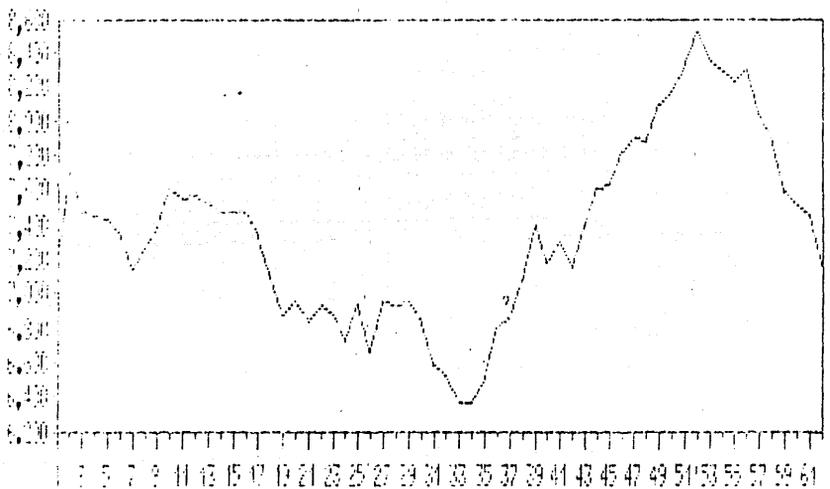
ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

# TERCEIRAS ORDENADAS PRINCIPAL



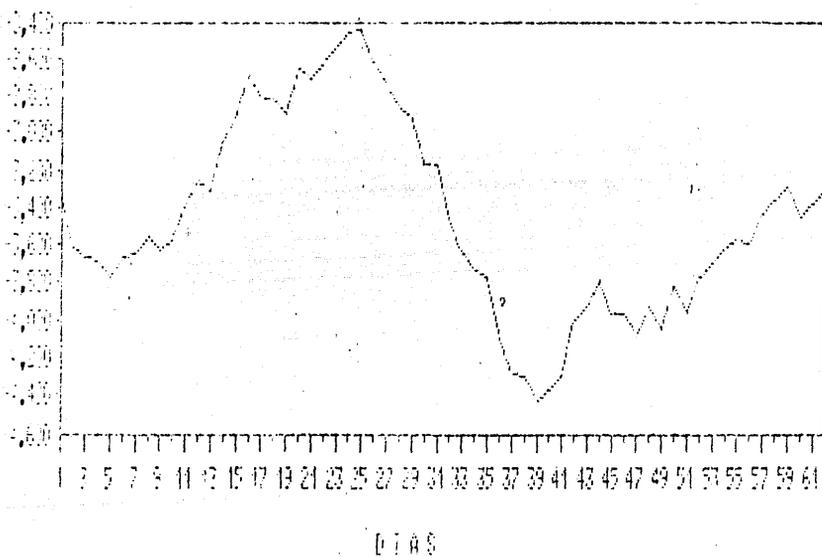
0 1 1 8

### CURVA COMERCIAL PRINCIPAL

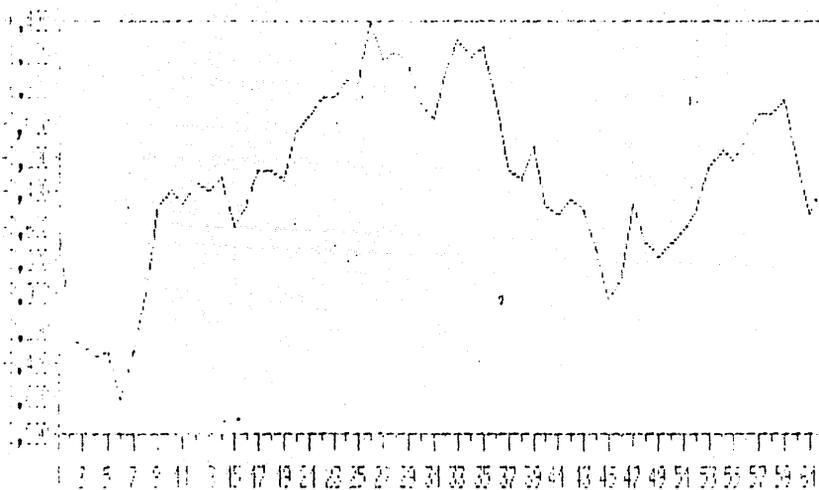


PLAS

QUINTA COMPONENTE PRINCIPAL

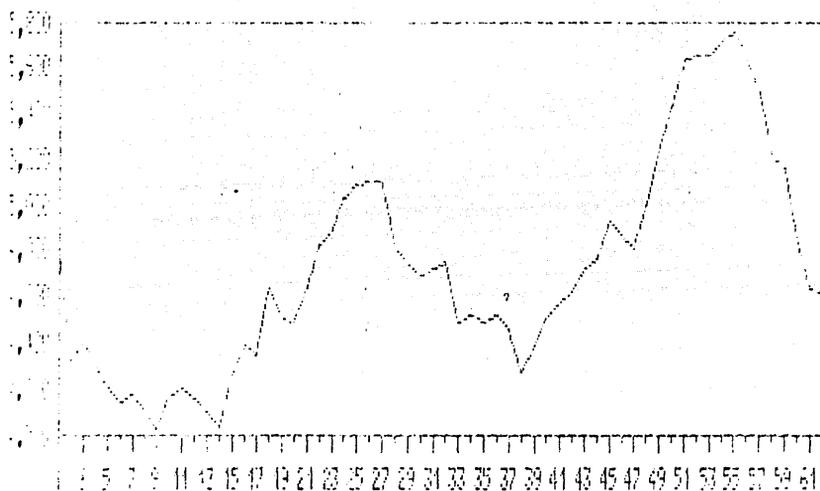


### SOLO COMPONENTE PRINCIPAL



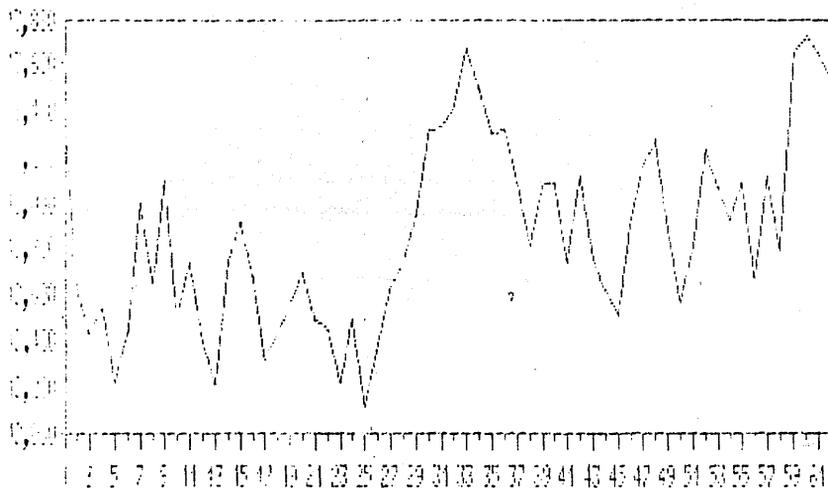
Dias

### SERIAL CORRELATION FUNCTION



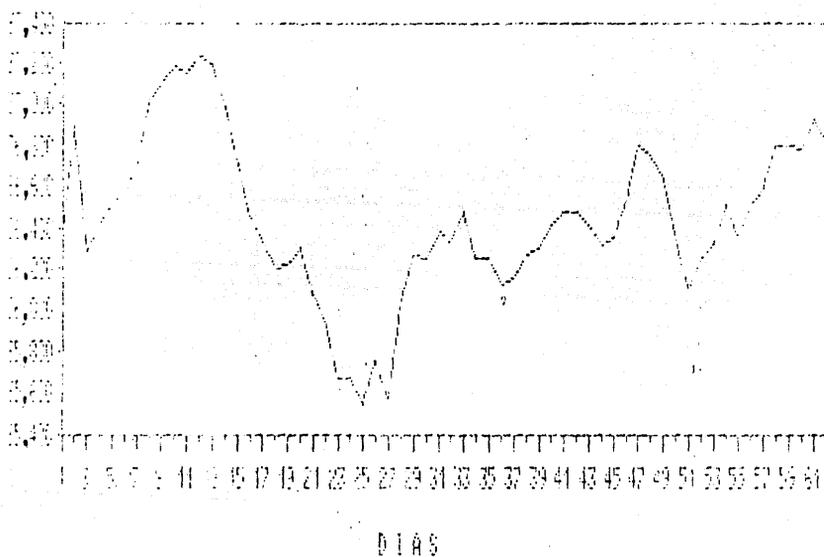
0 1 0 8

OCTAVA COMPONENTE PRINCIPAL

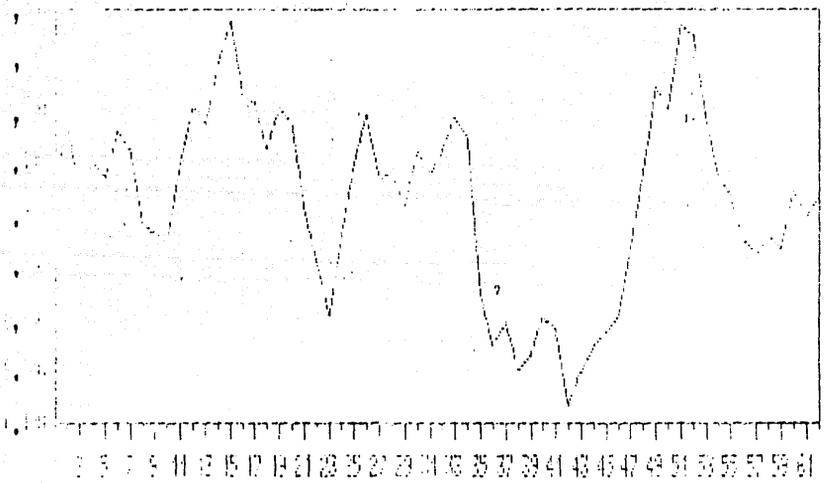


DIAS

### NOVEDA COMPONENTE PRINCIPAL

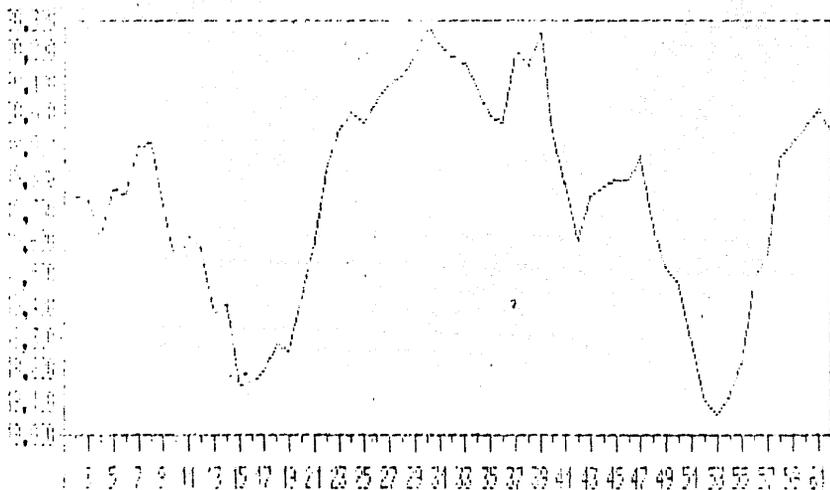


DECIMA COMPONENTE PRINCIPAL



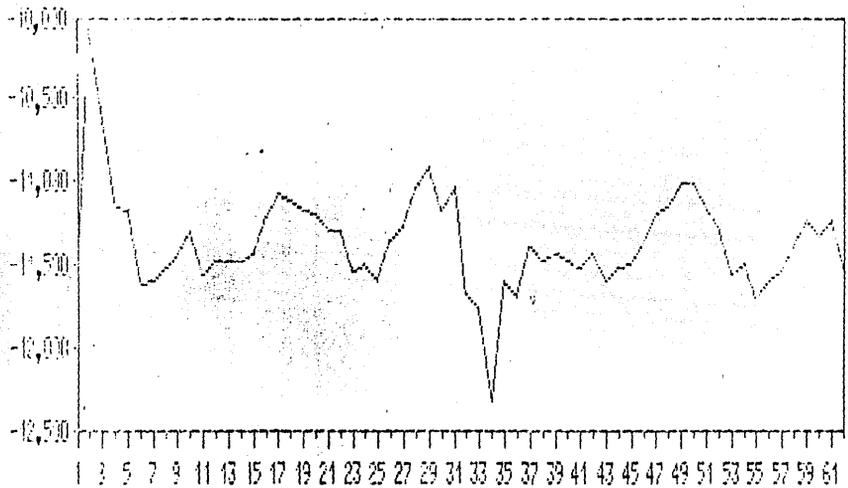
D I A S

CONTINUIDADE COMPLETA INTEGRAL



DIAS

DECIMASEGUNDA COMPONENTE PRINCIPAL



DIAS

Varianzas de las series originales y de las componentes principales.

	Serie original	%	componente pral.	-%
1	1,256,019.12	31.64%	1,196,103.69	30.13%
2	40,602.23	1.02%	784,101.63	19.75%
3	59,737.69	1.50%	341,339.37	8.60%
4	452,859.10	11.41%	272,222.75	6.86%
5	383,252.56	9.65%	270,098.00	6.80%
6	2,939.25	0.07%	257,860.74	6.50%
7	179,768.15	4.53%	213,717.98	5.38%
8	126,570.03	3.19%	164,289.64	4.14%
9	2,701.26	0.07%	150,243.84	3.78%
10	41,492.72	1.05%	123,345.67	3.11%
11	206,683.14	5.21%	107,617.67	2.71%
12	1,217,245.77	30.66%	88,940.15	2.24%
TOTAL	3,969,881.13	100.00%	3,969,881.13	100.00%

A continuación presentamos una lista de conclusiones resultantes del ejemplo utilizado:

- Como primer punto quisiéramos resaltar que el total de la varianza de la serie original es preservada en la serie de componentes principales.

- Observamos que la primera componente principal resume la mayor varianza de todas las series, seguida por la segunda, etc.

- Como resultado de la aleatoriedad con que se seleccionaron

lo cual implica que en las series originales haya poca información redundante, es decir, aún las de varianza muy pequeña aportan información en el comportamiento global de la serie.

- Los resultados en las componentes principales, aún cuando acumulen menos varianza que las originales, presentan información libre de redundancia, pues obedecen a las restricciones del planteamiento del problema.

- Para un análisis global, podríamos considerar unicamente las primeras siete componentes principales, pues éstas resumen mas del 80% de la varianza del comportamiento de las series; sin embargo, para poder recuperar información de las series originales a partir de las componentes principales, tendríamos que considerar todas para no perder información (ninguna componente principal es despreciable).

- Observando la información de las componentes principales, tenemos que:

- 1) Aproximadamente en la mitad del periodo, se presenta un cambio en la tendencia del comportamiento de las series.
- 2) La primera componente principal explica de manera inversa el comportamiento de la primera, quinta, séptima y decimosegunda series originales, es decir mantienen una estrecha correlación inversa.
- 3) La segunda componente principal presenta una tendencia inversa similar a la serie original número dos y otra similar a la de la serie original seis.

## CONCLUSIONES Y RECOMENDACIONES

1. Podemos concluir que las técnicas presentadas a lo largo del trabajo, funcionan y generan resultados satisfactorios.
2. El enfoque espectral es en efecto una alternativa para lograr los objetivos dentro del análisis canónico de series de tiempo.
3. Al igual que el análisis de componentes principales, se presentó el análisis para la obtención de variables canónicas y aún cuando éste no fué ejemplificado, se esperaría obtener resultados satisfactorios.
4. En el planteamiento original de este trabajo, se esperaba realizar un análisis sobre el comportamiento de todas las acciones bursátiles de la Bolsa Mexicana de Valores con el fin de obtener índices que lo describieran de manera mas sencilla y eliminando toda la información no necesaria. Desgraciadamente por problemas de almacenamiento en computadoras, no fué posible realizar este proyecto, quedando como un problema en el que miraremos a futuro.
5. Una aplicación directa de las técnicas aquí presentadas es la recuperación de información a través de componentes principales y la posibilidad de pronósticos certeros de la información que se maneja.

## BIBLIOGRAFIA

1. T.W. Anderson  
An Introduction to Multivariate Statistical Analysis  
John Wiley and Sons, Inc. 1971.
2. T.W. Anderson  
The Statistical Analysis of Time Series  
John Wiley and Sons, Inc. 1971.
3. Brillinger  
Time Series Data Analysis and Theory  
Expanded edition, Holden-day 1981.
4. A. Bustos y de la Tijera  
Análisis de Series de Tiempo  
Notas de Clase, 1979.
5. J.L. Doob  
Stochastic Processes  
John Wiley and Sons, Inc. 1952.
6. Fishman  
Spectral methods in Econometrics  
Harvard University Press, 1969.
7. Granger and Hatanaka  
Spectral Analysis of Economic Time Series.  
Princeton University Press, 1964.
8. E.J. Hannan  
Multiple Time Series  
John Wiley and Sons, Inc. 1970.