



03061  
4  
24

**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO**

**UACP y P del CCH**

**EL METODO X-II ARIMA Y EL FILTRO DE KALMAN  
EN EL AJUSTE ESTACIONAL DE SERIES DE TIEMPO**

**FALLA DE ORIGEN**

**TESIS**  
**QUE PARA OBTENER EL TITULO DE**  
**MAESTRO EN CIENCIAS**  
**PRESENTA:**  
**MARTIN ROMERO MARTINEZ**

**México, D. F.**

**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.

# I N D I C E

CAPITULO.	Pag.
0. - OBJETIVO	0
1. - INTRODUCCION	2
1.1 COMPONENTES DE UNA SERIE DE TIEMPO	
1.2 ASPECTOS DEL AJUSTE ESTACIONAL	
1.3 SISTEMAS LINEALES	
2. - METODO X-11-ARIMA	17
2.1 DESCRIPCION DEL METODO	
2.2 EJEMPLO	
3. - EL FILTRO DE KALMAN	40
3.1 INTRODUCCION	
3.2 FILTRO DE KALMAN	
3.3 SERIES DE TIEMPO	
3.4 EJEMPLO	
4. - SIMULACIONES	80
4.1 CONDICIONES DE SIMULACION	
4.2 ESTIMACION DEL COMPONENTE TENDENCIA	
4.2.1 METODO X-11-ARIMA	
4.2.2 FILTRO DE KALMAN	
4.2.3 COMPARACION	
4.3 INCLUSION DE UNA VARIABLE AJUSTADA ESTACIONALMENTE EN UN MODELO DE REGRESION	
5. - CONCLUSIONES	120
APENDICE A PROGRAMA X-11-ARIMA	122
APENDICE B PROGRAMA K	124
BIBLIOGRAFIA	131

EL METODO X-11 ARIMA Y EL FILTRO DE KALMAN  
EN EL AJUSTE ESTACIONAL DE SERIES DE TIEMPO

Martín Romero Martínez

Se estudian dos métodos de ajuste estacional : el X-11-ARIMA (Dagum,1983) y el uso del filtro de Kalman (Harvey,1989), los capítulos de la tesis son : Introducción, Método X-11-ARIMA, Filtro de Kalman y Simulaciones.


En los primeros capítulos se presenta con cierto detalle a los métodos en comparación, y en la parte final se realizaron simulaciones, para medir la bondad de los métodos en la estimación del componente tendencia de una serie de tiempo, y analizar el modelo de regresion lineal simple, cuando la variable regresora es el componente tendencia de una serie de tiempo.

Por último, las referencias básicas de la tesis son :

Dagum, E.B. (1983): The X-11-ARIMA Seasonal Adjustment Method, Statistics Canada, Time Series Research and Analysis Staff, Catalogue 12-564E.

Harvey, A.C.(1989): Forecasting, Structural Time Series Models and the Kalman Filter, Cambrdige University Press.

Vo. Bo.

  
Dr. Víctor Alfredo Bustos  
y de la Tijera.

## OBJETIVO.

El ajuste estacional de series de tiempo es una técnica estadística importante, porque muchas series de tiempo, en su mayor parte económicas, son ajustadas estacionalmente para su publicación o inclusión en un modelo.

El propósito del ajuste estacional es la eliminación de las fluctuaciones estacionales e irregulares, que contribuyen poco en la comprensión del comportamiento básico o tendencia de la serie de tiempo observada, porque según Dagum (1978,p.10) :

" Las principales causas de estacionalidad son factores climáticos o institucionales, que son exógenos al sistema económico y no pueden ser controladas por los tomadores de decisiones, en un período corto ",

por lo tanto, el ajuste estacional de una serie de tiempo, consiste de manera muy general, en la estimación de los componentes tendencia, estacional e irregular.

Las diversas soluciones propuestas al ajuste estacional podemos agruparlas en función de los supuestos hechos sobre los componentes de la serie, y con éste criterio, se presenta la clasificación siguiente :

- a) métodos que no modelan a las componentes,  
X-11, X-11-ARIMA (Dagum 1983),
- b) métodos que modelan los datos observados, y deducen modelos para los componentes,  
Burman (1980), Hillmer y Tiao (1982),
- c) métodos que definen modelos para los componentes, y entonces implican un modelo para los datos observados,  
Akaike (1980), Gersh y Kitagawa(1984).

En éste trabajo se estudiará el método X-11-ARIMA de Dagum (1983) y el desarrollado por Gersh y Kitagawa(1984); las razones para escoger estos métodos se explican a continuación.

El método X-11-ARIMA es el más usado en oficinas de gobiernos y en industrias, tiene como atractivo ser casi automático, y no supone modelo alguno para los componentes de la serie; su origen se remonta al método de Shiskin y Eisenpress(1957), uno de los primeros métodos implantados en una computadora.

El método de Gersh y Kitagawa(1984) es relativamente nuevo, se escogió por estar fundamentado por todo un desarrollo teórico, basado en el filtro de Kalman, y es una generalización del trabajo de Akaike(1980).

En éste trabajo se pretende comparar los métodos en estudio en dos aspectos, la estimación del componente tendencia por sí sola, y su estimación para su inclusión en un modelo lineal. Para lograr el objetivo mencionado, éste trabajo se dividió en tres partes, descritas en el parrafo siguiente.

En una primera parte, se discuten de manera breve los conceptos más usados, a continuación, se presentan los fundamentos de los métodos en consideración, y en la parte final, se realizaron simulaciones para comparar a los métodos.

En las simulaciones, se mide la eficiencia de los métodos en la estimación de los componentes tendencia y estacional, y se estudia el modelo de regresion lineal simple, cuando la variable regresora es el componente tendencia de una serie de tiempo.

Al momento de iniciar el trabajo, no había en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la U.N.A.M (I.I.M.A.S), implantaciones de los métodos, por lo tanto, como parte del trabajo se instaló el programa X-11-ARIMA proporcionado por la Dra. E.B. Dagum, y se programó el método de Gersh y Kitagawa (1984).

En los apéndices se muestran los aspectos de cómputo relevantes para un uso posterior de los programas.

## CAPITULO I

### INTRODUCCION

En este capítulo se resumen conceptos y aspectos importantes del ajuste estacional de series de tiempo.

Ejemplos de series de tiempo en las que se manifiesta la presencia de un componente estacional son :

a) el precio de alguna verdura, generalmente en los meses de producción el precio baja,

b) la temperatura máxima observada en un día, en invierno se observan temperaturas bajas, pero en primavera y verano las temperaturas son altas,

c) el número de animales presentes en algún lugar, cuando la población a la que pertenecen es migrante, estas series tienen como característica común, la existencia de movimientos dentro de un intervalo de tiempo que se repiten periódicamente, y en algunos casos, esos movimientos contribuyen significativamente en el comportamiento global de la serie.

El primer supuesto del ajuste estacional es modelar a la serie de tiempo observada como la suma o producto de componentes independientes : tendencia  $N_t$ , estacional  $S_t$ , determinista  $D_t$ , e irregular  $I_t$ . Cuando la relación entre la serie y los componentes es multiplicativa, la transformación logaritmo la reduce a una relación aditiva.

La idea de que una serie de tiempo esta compuesta de varios componentes no observables, es una manera de interpretar a la serie. Según Nerlove, Grether y Carvalho (1979), la idea comenzó a ser popular en economía durante los años 1825-1875, y también hacen referencia a un trabajo del meteorólogo Buys Ballot dado a conocer en 1847, como uno de los primeros intentos de ajustar estacionalmente a una serie de datos.

Desde los trabajos iniciales hasta nuestros días, los conceptos de los componentes tendencia y estacional no han sido bien definidos, esto ha ocasionado que diferentes métodos estimen diferentes cosas, por esta razón, en la sección 1.2 se presentan varias descripciones, mas que una definición de los componentes.

### 1.1 COMPONENTES DE UNA SERIE DE TIEMPO.

Comenzaremos por hablar del componente tendencia, pues resulta ser el más importante.

#### TENDENCIA.

Desde un punto de vista matemático, el concepto de tendencia no es un concepto bien definido, y hablando vagamente, uno puede entender a la tendencia de una serie de tiempo, como aquella función que representa el comportamiento básico o fundamental.

Los primeros intentos para modelar a la tendencia suponían que ésta era una función del tiempo, o de otras cantidades observables, si la función resultaba ser lineal en los parámetros se manejaba un modelo de regresión.

Pero la mayoría de las veces, la tendencia es una función desconocida del tiempo o de otras variables, que se mueve suavemente en el tiempo y no puede ser bien representada en todo su dominio por un polinomio de bajo grado, aunque la supuesta suavidad, nos permitiría representar a la tendencia cuando menos localmente por polinomios, ya que el supuesto de suavidad es una propiedad local.

En el dominio de las frecuencias, la densidad espectral del componente tendencia se caracteriza por tomar valores altos en las frecuencias pequeñas y valores bajos en las frecuencias altas.



Cuando no es posible modelar a la tendencia con funciones conocidas del tiempo o de otras variables, porque depende de factores no observables o la relación que explica a la tendencia es muy compleja, es común suponer que el componente tendencia fue generado por un modelo Autoregresivo Integrado y de Promedios Móviles (ARIMA), con la estructura siguiente,

$$(I - B)^d N_t = \eta_t(B)c_t, \quad (1.1)$$

donde  $\eta_t$  es un polinomio en B,  $B(X)_t = X_{t-1}$ , y  $\{c_t\}$  es una sucesión de v.a.i.i.d.  $(0, \sigma_c^2)$ , es decir, variables aleatorias independientes e idénticamente distribuidas con media cero y varianza  $\sigma_c^2$ .

Si  $N_t$  es generado por el modelo (1.1), la función de densidad espectral de  $N_t$  no existe, porque el proceso no es estacionario, pero podemos definir a la función de densidad pseudo-espectral de (1.1) por

(1.2)

$$f_N(w) = (\sigma_c^2 / (2\pi)) \eta_t(e^{iw}) \eta_t(e^{-iw}) / ((1 - e^{iw})^d (1 - e^{-iw})^d)$$

con  $w \in [0, \pi]$ , función que no es acotada cuando  $w$  es cero.

## ESTACIONAL.

Algunas descripciones del componente estacional que han sido dadas son :

"Por variación estacional de una serie de tiempo económica, entenderemos a los movimientos intra-anales sistemáticos, no necesariamente regulares, que algunas veces son causados por fenómenos no económicos, tales como cambios climáticos y la ocurrencia regular de festividades religiosas " (Thomas y Wallis, 1978 ), mientras que Dagum (1978, p. 10) afirma,

" Las principales causas de estacionalidad son factores climáticos o institucionales, que son exógenos al sistema económico y no pueden ser controlados por los tomadores de decisiones, en un período corto ". La primera caracterización del componente estacional fue dada por Nerlove (1964), para quien,

" La estacionalidad es aquella característica de una serie de tiempo, que da lugar a picos en la función de densidad espectral

en las frecuencias estacionales ".

Usando el modelo de regresión, un componente estacional determinista  $S_t$  de período  $s$  (con  $s$  estaciones), tiene como características,

- repetir sus valores cada  $s$  períodos,
- la suma de cualesquiera  $s$  valores consecutivos debe ser cero,

y puede ser modelado con la inclusión de  $s-1$  variables indicadoras, donde la  $i$ -ésima variable indica cuando una observación corresponde a la estación  $i$  del año.

Un modelo ARIMA propuesto para  $S_t$  es

$$UC(B) S_t = \psi(B)b_t, \quad (1.3)$$

donde  $UC(B) = 1 + B^1 + \dots + B^{s-1}$ ,  $\langle b_t \rangle$  son v.a.i.i.d  $(0, \sigma_b^2)$ , y  $\psi$  es un polinomio en  $B$ .

Cuando  $S_t$  es generado por (1.3),  $S_t$  es capaz de evolucionar en el tiempo, pero localmente presenta un comportamiento parecido al de un componente estacional determinista.

La función de densidad pseudo-espectral de (1.3) es

$$f_g(w) = (\sigma_b^2 / (2\pi)) \psi(e^{iw}) \psi(e^{-iw}) / (UC(e^{iw}) UC(e^{-iw})) \quad (1.4)$$

y esta función tiene como características

- ser infinita en las frecuencias estacionales  $w = 2\pi k/s$ , con  $k = 1, 2, \dots, [(s/2)]$ ,
- tener mínimos relativos cerca de las frecuencias  $w = \pi(2k-1)/s$  para  $k = 1, 2, \dots, [(s-1)/2]$ .

## DETERMINISTA.

El componente determinista es modelado por

$$D_t = \sum_{j=1}^p x_{tj} \beta_j, \quad (1.5)$$

donde los valores  $x_{tj}$  son fijos y conocidos.

En series económicas que representan cantidades acumuladas, como puede ser la producción mensual de una máquina, algunas veces es necesario introducir un componente determinista, debido al llamado "efecto calendario", que se presenta porque los meses no son de igual tamaño, o la ocurrencia de un día particular de la semana no es constante en todos los meses.

Con el componente determinista se modela una intervención exógena conocida, como puede ser, el cambio de una política gubernamental, la aparición de un nuevo competidor, o la introducción de una campaña de publicidad.

## IRREGULAR.

El componente irregular puede ser modelado con un proceso  $ARMA(p, q)$ , pero la mayoría de las veces se supone que  $I_t$  no depende de su pasado ni de su futuro y es modelado por una sucesión de v.a.i.i.d.  $(0, \sigma_I^2)$ .

Por último, es necesario manejar observaciones faltantes y observaciones extremas, como en todo análisis estadístico.

Los modelos para los componentes tendencia y estacional han sido arbitrarios, porque son conceptos no totalmente definidos, dependientes del usuario y modelan a cantidades no observables, por lo tanto, cualquier modelo explícito para éstas componentes será arbitrario.

Una revisión del ajuste estacional es hecha por Bell y Hillmer (1984), y una introducción al tema se puede leer en los libros de Box y Jenkins (1976), Anderson (1971) y Kendall, Stuart y Ord (1983).

EL análisis de intervención es tratado por Box y Tiao (1976), y el efecto calendario por Cleveland y Devlin (1980).

## 1.2 ASPECTOS DEL AJUSTE ESTACIONAL.

Dos preguntas naturales son, si el usuario puede modelar a la serie de tiempo satisfactoriamente, cuales son los razones para ajustar estacionalmente a la serie, y la segunda, cuales son las características de un buen método de ajuste estacional.

Posibles respuestas a las preguntas planteadas, se dan en los párrafos siguientes.

### JUSTIFICACION DEL AJUSTE ESTACIONAL.

Conviene observar, que en los métodos de ajuste estacional donde se modelan los datos y deducen modelos para los componentes, el ajuste estacional se realiza después de modelar, y entonces tiene sentido la primer pregunta del párrafo anterior, mientras que, en los métodos que definen modelos para los componentes e implican un modelo para la serie, el ajuste estacional se efectúa al modelar.

Algunas razones para ajustar estacionalmente a una serie, independientemente del método a usar son las siguientes.

#### a) Simplificación de la Información.

Bell y Hillmer (1984, p.301) dicen : "El ajuste estacional es hecho para simplificar los datos, de manera que pueda ser más fácil su interpretación, por usuarios con no muchos conocimientos

estadísticos, sin una pérdida de información", ésta es una buena justificación cuando el porcentaje de variabilidad explicada por el componente estacional no es muy grande.

#### b) Pronósticos.

Cuando se intenta predecir el comportamiento global y no el valor puntual de la serie, es más fácil detectar cambios de dirección en una serie ajustada que en la serie original, anticipar por ejemplo, cuando una serie que mostraba un comportamiento creciente deja de hacerlo.

#### c) Relación de la serie con otros datos.

La presencia de un efecto estacional puede oscurecer las relaciones entre la serie de tiempo y otras series o eventos.

#### d) Estudio de ciclos económicos.

Al suponer que el componente no estacional, puede ser modificado en alguna medida, la estimación del componente tendencia sirve para tomar decisiones en el control de algunas actividades económicas y determinar el estado de la economía dentro del ciclo económico.

Cuando ninguna de las razones dadas es suficiente para ajustar estacionalmente, debe recordarse que la descomposición de la serie es una manera de ver a los datos, y como tal, proporciona información, pero la decisión de sustituir a la serie original por la tendencia en análisis posteriores, debe ser tomada en el contexto de cada problema.

### CRITERIOS PARA BUEN AJUSTE ESTACIONAL

La idea general para identificar un buen método de ajuste estacional es : que el método elimine unicamente al componente estacional, y ésta idea se ha traducido en los criterios mencionados a continuación.

a) Cualquier comportamiento intra-anual repetitivo debe ser eliminado.

b) Remover los picos en las frecuencias estacionales de la densidad espectral de la serie original, sin producir baches en esas frecuencias o afectar grandemente la densidad espectral en otras frecuencias.

c) El componente irregular debe comportarse como un ruido blanco.

d) Los movimientos cíclicos globales no deben ser distorsionados.

e) Debe existir alta coherencia entre las series original y ajustada.

f) El espectro cruzado de las series original y ajustada debe tener un desplazamiento mínimo de fase.

Estos criterios son buenos pero no suficientes, para determinar con certeza el mejor ajuste estacional de una serie, por las razones siguientes.

Los criterios b), e), y f) fueron propuestos por Nerlove (1964), pero Nerlove, Grether y Carvallo (1979) muestran que :

1.- la coherencia entre la serie ajustada y observada es uno, si la serie ajustada se obtiene como una combinación lineal (constante en el tiempo) de valores pasados y posiblemente futuros,

2.- los únicos filtros lineales e invariantes con desfaseamiento igual a cero son los filtros simétricos.

También Nerlove y Grether (1970) investigaron el estimador lineal con mínimo error cuadrático medio a través de simulaciones, y su conclusión fue que no satisfacía totalmente los criterios de Nerlove (1964).

Bell y Hillmer (1984, p. 345) y Pierce (1979) muestran que si el componente irregular es ruido blanco, el componente irregular estimado con el criterio de mínimo error cuadrático medio no es

un ruido blanco.

Debe recordarse, que cuando el ajuste estacional esta basado en la definición de un modelo para los componentes, un criterio general para seleccionar entre dos alternativas es maximizar la verosimilitud o una función de ésta, como puede ser minimizar el criterio  $AIC = -2 * \ln(L(\theta)) + 2p$ , propuesto por Akaike, donde  $L(\theta)$  es la función de verosimilitud y  $p$  el número de parámetros.

De esta pequeña discusión puede concluirse, que los criterios propuestos son deseables pero no necesarios, y nuevamente, la decisión final debe ser tomada por el usuario, conjuntando su conocimiento del problema, los criterios expuestos y de ser posible el objetivo del ajuste estacional.

### 1.3 SISTEMAS LINEALES.

El modelo más simple de la Estadística es el modelo lineal, y una generalización de éste son los sistemas lineales, pero el uso de los sistemas lineales en la Estadística, comenzó de manera muy discreta en Duncan y Horn (1972) y Harvey y Phillips (1976). La aplicación de éstos modelos en las series de tiempo se revisa de manera breve en el capítulo 3.

Comenzaremos por describir un sistema lineal.

Un sistema lineal invariante y discreto en la forma espacio-estado, es especificado por las matrices  $(\Phi, C, D, F)$  y las ecuaciones

$$\begin{aligned}x(t+1) &= \Phi x(t) + D \xi(t), \\ y(t) &= C x(t) + F \xi(t),\end{aligned}\tag{1.6}$$

donde  $\Phi \in \mathbb{R}^{n \times n}$ ,  $C \in \mathbb{R}^{s \times n}$ ,  $D \in \mathbb{R}^{n \times r}$ , y  $F \in \mathbb{R}^{s \times r}$

y  $\xi$  : proceso de entradas no observables al sistema,

$\xi(t)$  son v.a.i.i.d. con vector de medias cero y matriz de covarianzas  $I$ ,

$y$  : proceso observable o de salida,

$x$  : el estado del sistema a tiempo  $t$ , es una variable no observable,

$C$  : la matriz de observación, que sirve como ventana para  $x$ ,

$\Phi$  : la matriz de transición para el sistema, define la manera en que evolucionaría el sistema en ausencia de ruido,

$F, D$ : Matrices para correlacionar a  $\xi$ .

Existen sistemas lineales en los que el vector de estados evoluciona de manera continua y el proceso se observa de manera continua o discreta, pero no serán discutidos en este trabajo.

El sistema (1.6) no es el más general, pues carece de variables de control, que pueden ser variables exógenas al sistema. Un sistema más general se presenta en la sección 3.2.



Los sistemas lineales con orientación a la Ingeniería pueden verse en Kailath(1980), y el tratamiento estadístico es hecho por Caines(1988) y Hannan(1988) .

Ahora se definen los conceptos de sistema observable y controlable, que como se verá son muy importantes.

Observabilidad.

El sistema (1.6) es observable, si la matriz  $\Phi'$ ,

$$\Phi' = [ C', \Phi \cdot C', \Phi^2 C', \dots, \Phi^{n-1} C' ], \quad (1.7)$$

es de rango completo.

El significado de la definición de observabilidad es el siguiente, si  $\Phi'$  es de rango completo, es posible resolver las n ecuaciones

$$y(t) = C \Phi^t x(0), \quad t = 0, \dots, n-1, \quad (1.8)$$

es decir, en ausencia de  $\xi$  es posible determinar (observar)  $x(0)$  (y en consecuencia a  $x(n)$ ) a partir de  $y(t)$ ,  $t=0, \dots, n-1$  .

Cuando el sistema no es observable,  $\Phi'$  tiene rango  $n_1$  ( $n_1 < n$ ), existe una matriz T tal que  $(T')^{-1} \Phi'$  tiene  $n-n_1$  renglones de ceros. los primeros  $n_1$  renglones de T son linealmente independientes y T satisface

$$T \Phi T^{-1} = \begin{bmatrix} \Phi_{11} & 0 \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \quad CT^{-1} = [ C_1 \quad 0 ], \quad (1.9)$$

al multiplicar por T al sistema (1.6)

$$\begin{aligned} T x(t+1) &= T \Phi T^{-1} T x(t) + T D \xi_t, \\ y(t) &= C T^{-1} T x(t) + F \xi_t, \end{aligned} \quad (1.10)$$

y renombrando adecuadamente en el sistema, este se transforma en

$$\begin{aligned} x_1(t+1) &= \Phi_{11} x_1(t) + D_1 \xi_t, \\ x_2(t+1) &= \Phi_{21} x_1(t) + \Phi_{22} x_2(t) + D_2 \xi_t, \\ y(t) &= C_1 x_1(t) + F \xi_t, \end{aligned} \quad (1.11)$$

donde  $(\Phi_{11}, C_1)$  es observable. De la nueva representación del

sistema, se deduce que puede omitirse a  $x_2(t)$ , pues no aparece en la ecuación para  $y(t)$ , es decir, el efecto de  $\xi$  sólo se observa a través de  $x_1(t)$ .

Ejemplificaremos el concepto de observabilidad con el sistema

$$\begin{bmatrix} N_t \\ S_t \\ S_{t-1} \\ S_{t-2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} N_{t-1} \\ S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{bmatrix} + \begin{bmatrix} e_1(t) \\ e_2(t) \\ 0 \\ 0 \end{bmatrix} \quad (1.12)$$

$$y(t) = [1, 1, 0, 0] [N_t, S_t, S_{t-1}, S_{t-2}]' + e_3(t)$$

La matrices  $\Phi'$  y  $(T')^{-1}$  son (1.13)

$$\Phi' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (T')^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & -1 & -1 & -1 \end{bmatrix}$$

además,

$$T\Phi T^{-1} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \quad (1.14)$$

y el nuevo vector de estados es (1.15)

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} N_t \\ S_t \\ S_{t-1} \\ S_{t-2} \end{bmatrix} = \begin{bmatrix} N_t + S_{t-2} \\ S_t - S_{t-2} \\ S_{t-1} - S_{t-2} \\ -S_{t-2} \end{bmatrix}$$

ahora tenemos un sistema que nos permite modelar a  $y_t$  de igual manera que (1.15), con menos variables, mostrando que el concepto de observabilidad, corresponde a la existencia de rango completo en el modelo lineal.

Controlabilidad.

El sistema (1.6) es controlable si la matriz  $C$ ,

$$C = [ D, \Phi D, \dots, \Phi^{n-1} D ], \quad (1.16)$$

es de rango completo.

El sentido de la definición de controlabilidad, se obtiene de la igualdad

$$x(t+1) = \sum_{j=1}^{t+1} \Phi^{j-1} D \xi(t+1-j) + \Phi^{t+1} x(0), \quad (1.17)$$

si pudieramos manejar a  $\xi(t), \dots, \xi(0)$ , podríamos hacer que el sistema a tiempo  $t+1$  tome cualquier valor, mientras  $x(0)$  sea finito y conocido.

Cuando el sistema no es controlable, podemos encontrar una matriz no singular  $T$  de manera que el sistema transformado sea

$$\begin{aligned} x_1(t+1) &= \Phi_{11} x_1(t) + \Phi_{12} x_2(t) + D_1 \xi_t, \\ x_2(t+1) &= \Phi_{22} x_2(t), \end{aligned} \quad (1.18)$$

donde el sub-vector de estados  $x_2(t)$  no es controlable, porque su valor no puede ser modificado a través de  $\xi$ .

Otras definiciones importantes se dan a continuación.

Construibilidad.

El sistema (1.6) es construible si

$$\text{rango}(O') \geq \text{rango}(\Phi^n), \quad (1.19)$$

y significa que el conocimiento de  $\xi_k$ ,  $0 \leq k \leq t$ ,  $(t) = n-1$ , nos permite conocer a  $x(t+1)$ .

Detectabilidad.

El sistema (1.6) es detectable si

a) existe una matriz  $T$  tal que,

$$T \Phi T^{-1} = \begin{bmatrix} \Phi_{11} & 0 \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \quad CT^{-1} = [ C_1, 0 ] \quad (1.20)$$

con  $(\Phi_{11}, C_1)$  construible y

b) los eigenvalores de  $\Phi_{22}$  son menores que 1 en valor absoluto, es decir  $|\lambda_i(\Phi_{22})| < 1$ .

La propiedad de detectabilidad se interpreta del modo siguiente, la parte del vector que no observamos, no dependerá después de mucho tiempo de  $x_2(0)$ , porque los eigenvalores de  $\Phi_{22}$  son menores que 1 en valor absoluto.

Controlabilidad al origen.

El sistema (1.6) es controlable al origen si  $(\Phi', D')$  es construible, y se interpreta así, para sistemas controlables con un valor conocido de  $x(0)$ , existen  $\xi_i$ ,  $0 \leq t \leq k$ ,  $(k \geq n-1)$ , tal que  $x(k+1) = 0$ .

Y finalmente, un sistema es estabilizable si  $(\Phi', D')$  es detectable.

En seguida se presentan tres teoremas, que muestran la importancia de los conceptos anteriores.

El primer teorema muestra que un sistema lineal e invariante, puede ser expresado en las formas entrada-salida o espacio-estado (Hannan, 1988, p.16).

Teorema 1.1

Cualquier función de transferencia racional en  $z$

$$Z(z) = \sum_{j=0}^{\infty} L_j z^j, \quad L_0 < \infty, \quad (1.21)$$

puede ser expresada como  $C(zI - \Phi)^{-1}D + F$ ,

donde  $F = L_0$ ,  $L_j = C\Phi^j D$ ,  $F \in \mathbb{R}^{a \times r}$ ,  $\Phi \in \mathbb{R}^{n \times n}$ ,  $C \in \mathbb{R}^{a \times n}$  y  $D \in \mathbb{R}^{n \times r}$

y se dice que  $(\Phi, C, D, F)$  es una realización algebraica de  $Z$ .

Un concepto importante es la minimalidad de un sistema lineal.

Definición.

El sistema lineal  $(\Phi, C, D, F)$  es minimal, si  $\Phi$  es la matriz de menor dimensión posible entre todas las realizaciones posibles de la función de transferencia  $Z(z)$ .

Y dos teoremas, que nos dicen que es suficiente trabajar con los sistemas observables y controlables.

Teorema 1.2 .

Un sistema lineal  $(\Phi, C, D, F)$  es minimal si y sólo si es observable y controlable.

Teorema 1.3.

Si los sistemas lineales minimales  $(\Phi, C, D, F)$  y  $(\Phi^*, C^*, D^*, F^*)$  son realizaciones de la misma función de transferencia, entonces existe un matriz no singular  $T$  tal que

$$\begin{aligned} \Phi^* &= T \Phi T^{-1} & C^* &= C T^{-1} \\ D^* &= D & F &= F^* \end{aligned} \quad (1.22)$$

## CAPITULO II.

### METODO X-11-ARIMA.

El método X-11-ARIMA tiene como antecedentes a los métodos I y II de la Oficina del Censo de los E.U.A. (Shiskin y Eissenpres, 1957), que fueron dos de los primeros métodos en ser implantados en una computadora.

El método X-11-ARIMA consta basicamente de tres fases (Dagum, 1983) :

- a) modelar a la serie de tiempo observada con un modelo ARIMA,
- b) extender la serie al principio y final del periodo de observación, con pronósticos para un año obtenidos del modelo seleccionado en a),
- c) ajustar estacionalmente a la serie extendida, haciendo uso del Método II-X-11, tal y como fue propuesto por Shiskin, Young, y Musgrave (1967).

En la sección 2.1 se explica con algún detalle cada una de las fases del método X-11-ARIMA.

#### 2.1.- DESCRIPCION DEL METODO.

##### MODELADO Y EXTRAPOLACION.

Existen dos maneras de obtener el modelo ARIMA que servirá para extender a la serie :

- el usuario lo proporciona,
- el programa lo seleccione de manera automática.

El programa maneja tres tipos de relación entre los componentes, y el tipo de relación determina los posibles modelos que ajustaría el programa. Las relaciones y los modelos se muestran en la tabla 2.1

Tabla 2.1.

Relación	Modelos
Multiplicativa	(0,1,1)(0,1,1) log
$Z_t = N_t * S_t * I_t$	(0,2,2)(0,1,1) log
o Log-Aditiva	(2,1,2)(0,1,1)
$\text{Log}(Z_t) = \text{Log}(N_t) + \text{Log}(S_t) + \text{Log}(I_t)$	
Aditiva	(0,1,1)(0,1,1)
$Z_t = N_t + S_t + I_t$	(0,2,2)(0,1,1)
	(2,1,2)(0,1,1)

Los modelos que aparecen en la tabla fueron seleccionados de un conjunto más grande (12 modelos), los cuales habían demostrado su efectividad para modelar una gran variedad de series observadas.

La manera en que se seleccionaron los modelos que ajusta el programa X-11-ARIMAC(Dagum,1983), fue la siguiente.

Se obtuvieron 174 series de tiempo de diversos sectores económicos del Canada, cada serie abarcaba 15 años de observaciones mensuales y trimestrales, para cada una de las series se ajustaron los 12 modelos, y por cada modelo, se contó el número de series en las que se cumplieran los criterios siguientes :

- a) se acepta la hipótesis de aleatoriedad de los residuos con la prueba  $\chi^2$  de Ljung y Box (1978) al 10%.
- b) que el promedio de los valores absolutos de los errores relativos de predicción para los últimos tres años sea menor al 12%.

Con los criterios a) y b) los mejores modelos fueron

Tabla 2.2

Modelo	Porcentaje de las series en el que se cumplieron los 2 criterios para el modelo i, y no para los modelos 1 .. i-1.
$(2,1,2)(1,1,1)_g$	73 %
$(0,2,2)(0,1,1)_g \log$	5 %
$(0,1,1)(0,1,1)_g \log$	2 %

Por último, se incluyeron los modelos  $(0,1,1)(0,1,1)_g$  y  $(0,2,2)(0,1,1)_g$  para series modeladas con el supuesto de aditividad de los componentes, porque después de experimentar con series que seguían una relación aditiva, la transformación logarítmica afectó negativamente el promedio de los errores de predicción y la estadística  $\chi^2$ .

Para seleccionar el modelo, se estiman los dos primeros modelos propuestos, si alguno satisface los criterios a) y b), éste es el modelo seleccionado, si no, se ajusta el tercer modelo. Cuando ninguno de los dos primeros modelos es rechazado, se escoge el de menor error de pronóstico.

Para cada modelo ajustado el programa imprime.

a) El promedio de los valores absolutos de los errores relativos de pronóstico para cada uno de los tres últimos años, y para los últimos tres años. Si el promedio de los errores para los tres últimos años es mayor que el 12% el modelo es rechazado.

b) La prueba  $\chi^2$  de aleatoriedad de los residuales (Box y Ljung, 1987), cuando se rechace la hipótesis al 10% el modelo es rechazado.

c) El coeficiente de determinación  $R^2$ .

d) Verifica que no exista sobrediferenciación en el modelo estimado, revisando que las sumas de los parámetros del promedio móvil y del promedio móvil estacional sean menores que 0.9.

Cuando fue posible seleccionar un modelo, éste es usado para hacer pronósticos hacia atrás (backast); al hacer pronósticos



hacia atrás a) se sustituye por

a'). el porcentaje de error debe ser menor que el 18%, se impuso esta condición, porque para series de tiempo grandes, es pequeña la influencia de los pronósticos hacia atrás en los últimos factores estacionales.

El usuario puede proponer un modelo o hacer un ajuste sin extender la serie, cuando ninguno de los modelos fue aceptado.

#### METODO II-X-11.

Este es un procedimiento que descompone a la serie en la suma o producto de cinco componentes :

- tendencia (ND),
- estacional (S),
- irregular (ID),
- "Trading-day Variation" o efecto del número de días (TD).
- factores iniciales (PI):
  - para cada observación,
  - del Trading-day Variation,

y los principales pasos del método son:

- a) ajuste a priori,
- b) estimación preliminar de "Trading-day Variation" y pesos para los irregulares,
- c) estimación final de "Trading-day Variation" y pesos para los irregulares,
- d) estimación final de factores estacionales, tendencia, irregulares y serie ajustada estacionalmente.

Se describe el programa X-11-ARIMA en el ajuste de series mensuales, y las diferencias con X-11-ARIMA trimestral se señalan. Las palabras encerradas en corchetes "[ ]", son las adecuadas cuando la relación entre los componentes es aditiva o log-aditiva.

En la parte de ajuste a priori,

a) se extiende a la serie con pronósticos cuando se seleccionó algún modelo.

b) se divide [resta] la serie entre los factores iniciales de cada observación,

c) si se tienen los pesos iniciales para las variables que miden el número de veces que aparece un día en un período, también se hace un ajuste a priori.

En las tres últimas partes, se estima iterativamente a los factores estacionales y pondera a las observaciones irregulares.

Los pasos para la estimación de los componentes tendencia y estacional son.

1.- Dividir [restar a] la serie original por un promedio móvil  $2 \times 12$ , para tener una estimación inicial de  $S \times I$  [ $S+I$ ].

2.- Aplicar un promedio móvil  $3 \times 3$  a las estimaciones de  $S \times I$  [ $S+I$ ] obtenidas en (1), para cada mes por separado, para tener una estimación inicial de los factores estacionales  $S$ .

3.-Ajustar los factores estacionales, para que sumen aproximadamente 12 [0] sobre cualquier conjunto de 12 observaciones consecutivas, por medio de dividir [restarle] a cada factor  $S$  un promedio móvil  $2 \times 12$ . Para tener los seis valores faltantes al principio y final de la serie, repetir el primer y último factor disponible.

4. Divida [ reste a ] las estimaciones de  $S \times I$  del paso 1 por la estimación de los factores estacionales del paso 3, para tener una estimación preliminar de  $I$ .

5. Calcular una desviación estandar móvil ( $\sigma$ ) que incluya a irregulares estimados correspondientes a 5 años, y pruebe los componentes irregulares del año central contra  $2.5\sigma$ , remover los valores más grandes que  $2.5\sigma$  y repita la evaluación de  $\sigma$ .

Asigne un peso de cero a los irregulares mayores que  $2.5\sigma$ , un peso de 1 a los irregulares menores que  $1.5\sigma$ , y asignar un peso de manera lineal entre 0 y 1 a los irregulares entre  $1.5\sigma$  y

2.5o.

Para los primeros dos años usar la  $\sigma$  calculada para el tercer año, y para los dos últimos, usar la última  $\sigma$  calculada.

Para reemplazar un valor extremo en los dos primeros y últimos años, se toma un promedio de

- el valor por su peso, y

- los primeros ó últimos irregulares del mismo mes a los que se les asigne peso completo.

7.- A las estimaciones de  $S \times I$  con los valores extremos de  $I$  reemplazados, aplicar un promedio móvil  $3 \times 3$  para cada mes por separado, y obtener otra estimación de los factores estacionales.

8.- Repita el paso 3 a los factores encontrados en 7.

9.- Para tener una estimación preliminar de la tendencia, divida [reste a] la serie original por la serie de factores encontrados en 8.

10.- Para estimar el componente tendencia-ciclo, aplicar un promedio móvil de Henderson(Kenny y Durbin,1982) de 9,13 o 23 términos a la serie obtenida en 9, y divida [reste a] la serie original por la estimación de tendencia-ciclo, para tener una segunda estimación de  $S \times I$  [ $S+I$ ].

11.- A las estimaciones de  $S \times I$  [ $S+I$ ] aplique un promedio móvil  $3 \times 5$  por cada mes, para tener una segunda estimación del componente estacional.

12.- Repita el paso 3.

13.- Divida la serie original por las estimaciones de  $S \times I$  [ $S+I$ ] del paso 11, para obtener la serie ajustada estacionalmente.

En las fases 2,3 y 4 del método X-11, se aplican casi todos los pasos descritos, y el resultado de cada paso es una tabla denominada  $\alpha\beta$ , donde  $\alpha \in \langle A,B,C,D \rangle$  representa a la fase del método y  $\beta \in \langle 1, \dots, 13 \rangle$  es el paso.

En el programa X-11, el único componente determinista que es posible modelar es el efecto del calendario ( "Calendar Effect" o "Trading-day Variation" ), y se añade a los 13 pasos el siguiente procedimiento, propuesto por Eisenpress(1956).

Si  $ID_t$  representa a la serie de irregulares que se obtienen del paso 13, se tienen dos modelos para  $ID_t$ .

$$ID_t = \sum_{i=1}^7 x_{ti} \beta_i, \quad \text{relación aditiva,} \quad (2.1)$$

$ID_t - 1 = \left( \sum_{i=1}^7 x_{ti} \beta_i \right) / n_t$ , relación multiplicativa,  
 donde  $x_{ti}$  es el número de veces que el día  $i$  ocurre en el mes o trimestre,

$n_t$  es el número de días que tiene el mes o trimestre  $t$ , y  
 $\beta_i$  es el coeficiente asociado al efecto del día  $i$ .

Con los residuos del paso 13 estimar a  $\beta$  por mínimos cuadrados ordinarios OLS y construir

$$M_t = \sum_{i=1}^7 x_{ti} \hat{\beta}_i \quad (\text{relación aditiva}), \quad (2.2)$$

$$M_t = \sum_{i=1}^7 x_{ti} (1 + \hat{\beta}_i) / n_t \quad (\text{relación multiplicativa}),$$

los irregulares estimados son

$$I_t = ID_t - M_t, \quad \text{o} \quad I_t = ID_t / M_t,$$

dependiendo del tipo de relación entre los componentes, y por último, aplicar el paso 5 a los nuevos irregulares.

#### Consideraciones del método.

1.- No es posible usar un filtro simétrico para ajustar a las últimas observaciones, lo que hace necesario, usar los filtros asimétricos dados en Dagum(1983).

2.- Si la serie es trimestral, se usan los filtros de 5 y 7 términos de Henderson y un promedio móvil 2x4, en sustitución del promedio 2x12 y de los filtros de 9, 13 y 23 términos. La elección del filtro de Henderson la hace el usuario o el programa, en función de la variabilidad de los componentes.

3.- Los filtros usados por el método fueron seleccionados con base en la experiencia.

4.- Aunque el método no supone modelo alguno, Cleveland y Tiao(1979) han probado que X-11-ARIMA es óptimo en la estimación de  $N_t$  para el modelo

$$Z_t = N_t + S_t + I_t, \quad (2.3)$$

$$(I - B)^2 N_t = (I - \eta_1 B - \eta_2 B^2) b_t,$$

$$(I - B^{12}) S_t = (I - \psi_{12} B^{12} - \psi_{24} B^{24}) c_t,$$

con

$$\eta_1 = -0.49 \quad \eta_2 = 0.49 \quad \psi_{12} = -0.64 \quad \psi_{24} = -0.83$$

$$\sigma_c^2 = 1.3\sigma_b^2 \quad \sigma_e^2 = 14.4\sigma_b^2$$

$$I_t = NCO, \sigma_e^2).$$

por lo tanto, para series cercanas de éste modelo, el ajuste será bueno en términos de error cuadrático.

5.- Es un método poco flexible por el limitado número de filtros disponibles, pues ya en 1957, Shiskin y Eisenpress (1957) al presentar su Método II decían :

" Debe ser claro a los estudiosos de éste problema que no existe una fórmula de graduación, la cual pueda delinear los movimientos cíclicos básicos de todas las series ",

y mencionaron la posibilidad de desarrollar criterios para lograr una buena selección del filtro, dependiendo de las fluctuaciones estacionales e irregulares particulares de cada serie. El método X-11-ARIMA ya incluye criterios para seleccionar al filtro de Henderson y otros filtros, pero aún con éstas opciones, el número de filtros es limitado.

Como se ve, desde sus orígenes ya se conocía la principal limitación del método, que es su falta de flexibilidad.

6.- Los filtros de Henderson tienen la propiedad de minimizar la suma cuadrados de las terceras diferencias de la serie ajustada, la demostración puede verse en Kenny y Durbin(1982).

7.- Una de las características que hizo difícil el estudio del método X-11, son los elementos no lineales que contiene, entre ellos, la estimación móvil de  $\sigma$  para sustituir valores. Una aproximación lineal al método fue dada por Young(1968).

8.- Cuando la longitud de la serie es mayor que 29 años, el programa X-11-ARIMA no ajusta modelo.

El método X-11-ARIMA representó un avance respecto al método II-X-11, al hacer que el ajuste estacional dependa de un modelo ajustado a la serie, y resolver parcialmente su principal limitación, sus ventajas según Dagum(1983) son.

La existencia de un modelo que ajuste a la serie original, da una medida de la calidad de los datos, aunque el modelo no satisfaga los requisitos para la extrapolación. Cualquier descomposición de una serie es de dudosa validez, si no es posible identificar un modelo ARIMA de la serie.

El valor esperado y la varianza de la serie original pueden ser calculados si la serie tiene un modelo ARIMA, y entonces, pueden ser construidos intervalos de confianza para las observaciones, e identificar posibles valores extremos.

La estimación del componente tendencia para las últimas observaciones combina a los filtros simétricos de Henderson y los pesos que el modelo ARIMA asigna al extrapolar, de ésta manera, los últimos pesos cambian con el modelo ajustado, y el filtro final refleja los movimientos más recientes de la serie.

Al añadir 1 o 2 años de datos extrapolados, se tiene una mejor estimación de la varianza de los irregulares, y significa una mejora en la identificación y reemplazo de las observaciones extremas. Se encontró que un año de pronósticos hacia atrás y adelante es la mejor elección para muchas series.

Si los últimos factores estacionales son aplicados para ajustar estacionalmente a las observaciones actuales, no hay necesidad de revisar el ajuste más de 2 veces (evidencia empírica), y para muchas series una revisión es suficiente,

Es un método que en la practica ha funcionado aceptablemente, es casi automático, y maneja automaticamente a posibles observaciones extremas.

En la aplicación de todo método estadístico es necesario probar o cuando menos cuantificar, la medida en que son satisfechos los supuestos del método, y para X-11-ARIMA se han propuesto algunas estadísticas de control. Como no existe un modelo para la serie y los componentes, las pruebas de hipótesis que se proponen son aproximadas.

#### ESTADÍSTICAS DE CONTROL Y PRUEBAS DE HIPÓTESIS.

P1.- Prueba de presencia de estacionalidad.

El propósito de ésta prueba es verificar la significancia del componente estacional estimado, es decir

$H_0$  : no existe factor estacional en la serie, o

$H_0$  : los factores estacionales son iguales para todos los meses.

La prueba se obtiene de un Análisis de Varianza (ANOVA) con una entrada, para la estimación del componente estacional del paso 1 fase 2 (Tabla B1), si  $w_{ij}$  es el j-ésimo factor estacional correspondiente al i-ésimo año, se definen las sumas de cuadrados siguientes ,

$$S.C.M. = \sum_j (\bar{w} - \bar{w}_{+j})^2 \quad \text{entre meses,} \quad (2.4)$$

$$S.C.T. = \sum_{ij} (\bar{w} - w_{ij})^2 \quad \text{total,}$$

$$S.C.R. = S.C.T. - S.C.M. \quad \text{residual,}$$

y bajo el supuesto de que los factores estacionales son constantes en el tiempo, la estadística de prueba es

$$F_s = \frac{S.C.M. / (s-1)}{S.C.R. / (T - 2*1 - s)} \quad (2.5)$$

donde s es el periodo de estacionalidad,  $1 = 0$  para observaciones mensuales y  $1 = 2$  para observaciones trimestrales. Dado que se viola el supuesto de normalidad, se prueba la hipótesis con un nivel de significancia del 0.1% y una distribución F de Fisher, con  $(s-1, T-2*1-s)$  grados de libertad, siendo deseable un valor alto de la estadística de prueba.

## P2.- Prueba de Estacionalidad Móvil.

La hipótesis a probar es

$H_0$  : El factor estacional es estable,

y se trabaja con la serie  $w_{ij} = \text{abs}(W_{ij})$  [ $w_{ij} = \text{abs}(100 - W_{ij})$ ], donde  $W_{ij}$  es la estimación del componente  $S \times I$  [ $S+I$ ] correspondiente al año  $i$  mes  $j$  de la tabla DS.

La prueba se obtiene de una tabla de ANOVA de dos entradas para la serie  $w_{ij}$ , y trata de medir el cambio del componente estacional en el tiempo.

Se definen las sumas de cuadrados entre años (S.C.A) y de los residuos (S.C.R),

$$S.C.A = \sum ( \bar{w} - \bar{w}_{i+} )^2, \quad (2.6)$$

$$S.C.R. = S.C.T. - S.C.A. - S.C.M.,$$

la estadística de prueba es

$$F_m = \frac{S.C.A. / (n - 1)}{S.C.R. / (T - 1 - (n-1) - (s-1))} \quad (2.7)$$

donde  $n$  es el número de años. La prueba de hipótesis se hace con un nivel de significancia del 5% y una distribución  $F(n-1, T-1-(n-1)-(s-1))$ . En ésta prueba es deseable un valor bajo de la estadística.

P3.- También se hace la prueba de Kruskal-Wallis (Conover, 1980, p. 229) al 5%, para probar estacionalidad estable.

P4.- Prueba para la presencia de un componente estacional identificable. En esta prueba existen tres posibles decisiones

$H_1$  : Probable no presencia de estacionalidad identificable,

$H_2$  : Presencia de Estacionalidad Identificable,

$H_3$  : No presencia de Estacionalidad Identificable.

Si  $F_s^*$ ,  $F_m^*$  y  $K^*$  son los niveles de significancia descriptiva para  $F_s$ ,  $F_m$  y la prueba de Kruskal-Wallis, se evalúan

$$T1 = 7 / (F_m - F_s), \quad T2 = (3F_m) / F_s, \quad (2.8)$$

$$T = (T1 + T2) / 2,$$



y la regla de decisión es :

a) si  $F_g^* > 0.1\%$  entonces  $H_3$ .

b) si  $F_m^* \leq 5\%$  y  $T > 1$  entonces  $H_3$ .

c) si  $F_m^* > 5\%$  y  $T_1 > 1$  entonces  $H_1$ .

d) si  $F_m^* > 5\%$  y  $K^* > 1\%$  entonces  $H_1$ .

e) si  $F_m^* > 5\%$  y  $T_2 < 1$  entonces  $H_2$ .

f) si no a), b), c), d) y e) en ese orden, entonces  $H_1$ .

El programa evalúa 11 estadísticas para medir la calidad del ajuste estacional, las estadísticas tienen como característica que son forzadas a tomar valores en  $[0,3]$ , y su región de rechazo es el intervalo  $[1,3]$ .

Las estadísticas de calidad del ajuste estacional son.

1.- La contribución relativa de los irregulares sobre periodos de tres meses.

Si es  $I(k) = \sum_j |I_j - I_{j+k}|$  cuando la relación entre los componentes es aditiva e  $I(k) = \sum |I_j - I_{j+k}| / |I_j|$  cuando la relación entre los componentes es multiplicativa, y se define de manera análoga a  $N(k)$ ,  $S(k)$ ,  $P(k)$  y  $TD(k)$ . La estadística reportada depende de

$$\frac{I^2(L)}{I^2(L) + N^2(L) + S^2(L) + TD^2(L)} \quad (2.9)$$

donde  $L = 4$  para series mensuales y  $L = 1$  en series trimestrales.

2.- La contribución relativa del componente irregular en la parte estacionaria de la varianza.

a) Relación aditiva. Se ajusta la línea recta  $m(t) = a + b*t$  a la serie  $Z$ , y la estadística reportada depende de la cantidad

$$\frac{\text{Varianza}(I(t))}{\text{Varianza}(O(t) - m(t))}$$

b) Relación multiplicativa. Se ajusta la línea recta a  $\ln(Z)$ , y el divisor se sustituye por  $\text{Varianza}(O(t) / \exp(m(t)))$ .

3.- El total de cambio mes a mes (trimestre a trimestre) en el componente irregular, comparado con el total de cambio mes a mes (trimestre a trimestre) del componente tendencia.

La estadística reportada depende de  $I(1)/N(1)$ .

4.- El total de correlación en los irregulares, medido por el promedio de la duración de una corrida.

5.- El primer mes o trimestre para el cual  $I(k)/N(k) < 1$ , denotado por MCD (datos mensuales) o QCD (datos trimestrales), cuando MCD o QCD es grande, se trata de una serie con irregulares muy grandes, ya que por la definición de  $I(k)/N(k)$ , se espera sea una función decreciente de  $k$ .

6.- El total del cambio año con año en los residuales comparado con el total del cambio año con año del componente estacional.

El valor reportado depende de  $I(L)/S(L)$ , donde  $L = 12$  si las observaciones son mensuales y  $L = 4$  si las observaciones son trimestrales.

7.- La relación entre los totales de estacionalidad móvil y estable. El valor reportado es función de los valores del nivel de significancia descriptivo para  $F_s$  y  $F_m$ .

En las estadísticas 8,9,10 y 11,  $S_t^*$  es el valor estandarizado de  $S_t$ ,  $s = 12$  para observaciones mensuales y  $s = 4$  en observaciones trimestrales.

8.- Una medida de la variación año con año del componente estacional. La estadística de prueba depende de

$$\sum_t |S_t^* - S_{t-s}^*|$$

9.- El promedio del movimiento lineal del componente estacional. El valor reportado depende de

$$\sum_{j=1,9} |S_{1,j}^* - S_{n,j}^*| / n$$

donde  $n$  es el número de años que abarca la serie.

10. Igual que 8, pero calculado para años recientes (3).

11. Igual que 9, pero calculado para años recientes (3).

Después de calcular las 11 estadísticas de prueba, se evalúa la estadística final como un promedio ponderado de los valores asignados a las estadísticas individuales, los pesos propuestos por Lothian y Morry(1978) son,

Prueba	1	2	3	4	5	6	7	8	9	10	11	total
Peso	10	11	10	8	11	10	18	7	7	4	4	100

## 2.1 EJEMPLO

Presentaremos y comentaremos la aplicación del método X-11-ARIMA a una serie de datos tomada de Dagum(1983), que abarca los años 1969-1978 y se llama "Freight and Shipping Payments". La salida del programa X-11-ARIMA esta delimitada por líneas del caracter "=".

En la Gráfica 2.1 se encuentran la serie observada y la tendencia estimada, y en la Gráfica 2.2 aparece el componente estacional estimado.

1.- Serie Original con totales por año y por trimestre.

```
=====
```

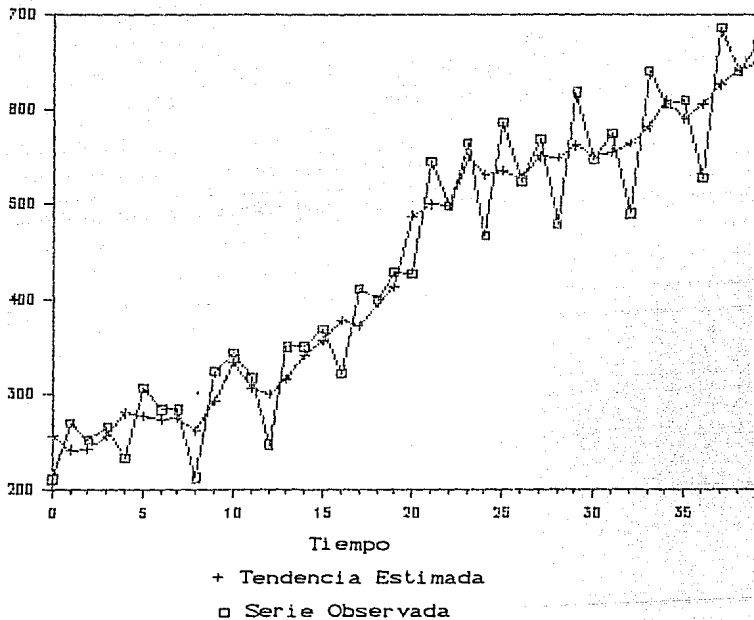
A 1. ORIGINAL SERIES					
YEAR	1ST QUAR	2ND QUAR	3RD QUAR	4TH QUAR	TOTAL
1969	210.0	269.0	252.0	265.0	996.0
1970	232.0	306.0	284.0	284.0	1106.0
1971	212.0	324.0	343.0	317.0	1196.0
1972	247.0	350.0	350.0	368.0	1315.0
1973	322.0	411.0	399.0	428.0	1560.0
1974	427.0	545.0	499.0	565.0	2036.0
1975	467.0	586.0	525.0	569.0	2147.0
1976	479.0	618.0	547.0	575.0	2219.0
1977	490.0	641.0	606.0	609.0	2346.0
1978	528.0	687.0	641.0	669.0	2525.0
AVGE	361.4	473.7	444.6	464.9	

```
TABLE TOTAL- 17446.0 MEAN- 436.1 STD. DEVIATION- 143.4
```

```
=====
```

En la Gráfica 2.1 se puede observar que la serie presenta más variabilidad al principio de la serie, y esto se refleja en que el error de pronóstico para los 3 primeros años es mayor que para los últimos.

Gráfica 2.1  
Método X-11-ARIMA



2.- Modelos ARIMA ajustados para hacer pronósticos hacia atrás (backcast) y hacia adelante (forecast).

FREIGHT AND SHIPPING PAYMENTS

PAGE 2, SERIES

A5. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) EXTRAPOLATION  
ARIMA EXTRAPOLATION MODEL (FORECAST)

THIS PROGRAM WAS DEVELOPED FOLLOWING THE PROCEDURES OUTLINED IN  
'TIME SERIES ANALYSIS' BY G. E. P. BOX AND G. M. JENKINS.  
AVERAGE PERCENTAGE STANDARD  
ERROR IN FORECASTS

MODEL	TRAN.	ADDITIVE CONSTANT	LAST 3 YEARS	LAST YEAR	LAST-1 YEAR	LAST-2 YEAR	CHI-SQ. PROB.
(0,1,1)(0,1,1)	NONE	0.000E+00	3.02	2.03	2.90	4.12	48.48%
(0,2,2)(0,1,1)	NONE	0.000E+00	3.28	2.32	5.00	2.52	48.67%

MODEL	R-SQUARED VALUE	ESTIMATED PARAMETERS		
(0,1,1)(0,1,1)	0.9569	0.2802E+00	0.3780E+00	
(0,2,2)(0,1,1)	0.9461	0.1075E+01	-0.1772E+00	0.3028E+00

THE MODEL CHOSEN IS (0,1,1)(0,1,1) WITH TRANSFORMATION - NONE

FREIGHT AND SHIPPING PAYMENTS

PAGE 3, SERIES

A6. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) EXTRAPOLATION  
ARIMA EXTRAPOLATION MODEL (BACKCAST)

THIS PROGRAM WAS DEVELOPED FOLLOWING THE PROCEDURES OUTLINED IN  
'TIME SERIES ANALYSIS' BY G. E. P. BOX AND G. M. JENKINS.

AVERAGE PERCENTAGE STANDARD  
ERROR IN BACKCASTS

MODEL	TRAN.	ADDITIVE CONSTANT	LAST 3 YEARS	LAST YEAR	LAST-1 YEAR	LAST-2 YEAR	CHI-SQ. PROB.
(0,1,1)(0,1,1)	NONE	0.000E+00	11.92	8.46	11.66	15.63	46.75%

MODEL	R-SQUARED VALUE	ESTIMATED PARAMETERS		
(0,1,1)(0,1,1)	0.9562	0.2897E+00	0.4340E+00	

THE MODEL CHOSEN IS (0,1,1)(0,1,1) WITH TRANSFORMATION - NONE

3.- Serie de datos extrapolada un año atrás y otro año hacia adelante.

=====

B 1. PRIOR ADJUSTED ORIGINAL SERIES

YEAR	1ST QUAR	2ND QUAR	3RD QUAR	4TH QUAR	TOTAL
1968	184.5	256.5	242.4	249.3	932.7
1969	210.0	269.0	252.0	265.0	996.0
1970	232.0	306.0	284.0	284.0	1106.0
1971	212.0	324.0	343.0	317.0	1196.0
1972	247.0	350.0	350.0	368.0	1315.0
1973	322.0	411.0	399.0	428.0	1560.0
1974	427.0	545.0	499.0	565.0	2036.0
1975	467.0	586.0	525.0	569.0	2147.0
1976	479.0	618.0	547.0	575.0	2219.0
1977	490.0	641.0	606.0	609.0	2346.0
1978	528.0	687.0	641.0	669.0	2525.0
AVGE	345.3	454.0	426.2	445.3	

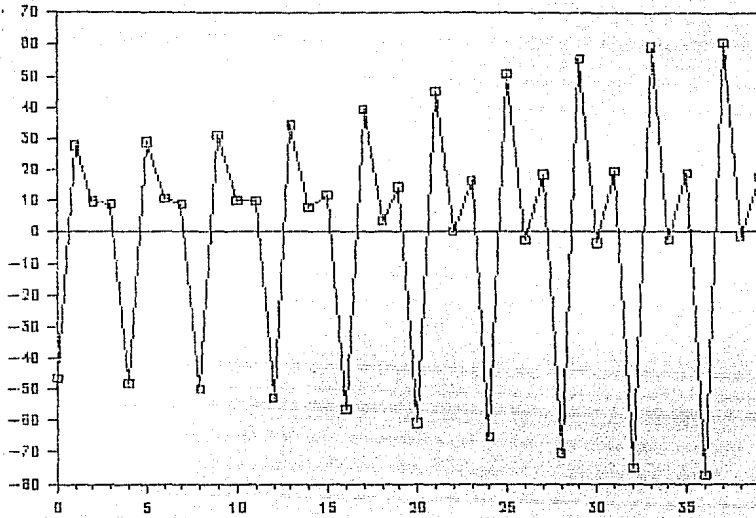
TABLE TOTAL- 18378.7 MEAN- 417.7 STD. DEVIATION- 148.9  
 ORIGINAL SERIES EXTRAPOLATED ONE YEAR AHEAD

YEAR	1ST QUAR	2ND QUAR	3RD QUAR	4TH QUAR	TOTAL
1979	583.0	735.9	690.1	713.2	2722.2

=====

El componente estacional (Gráfica 2.2) estimado tiene un comportamiento regular, incrementando su variabilidad al transcurrir el tiempo.

Gráfica 2.2  
Ajuste Estacional  
Método X-11-ARIMA



Tiempo

□ Componente Estacional Estimado

4.- Factores estacionales finales y pronóstico de factores estacionales.

=====

FREIGHT AND SHIPPING PAYMENTS PAGE 5, SERIES

D10. FINAL SEASONAL FACTORS

YEAR	3X5 MOVING AVERAGE SELECTED.					AVGE
	1ST QUAR	2ND QUAR	3RD QUAR	4TH QUAR		
1969	-46.2	28.0	9.9	8.9	0.2	
1970	-48.1	29.1	10.7	8.9	0.1	
1971	-50.0	31.0	10.0	10.0	0.2	
1972	-52.9	34.4	7.8	11.8	0.3	
1973	-56.4	39.5	3.7	14.5	0.3	
1974	-60.6	45.2	0.1	16.5	0.3	
1975	-65.0	50.9	-2.8	18.6	0.4	
1976	-70.2	55.8	-3.5	19.4	0.4	
1977	-74.8	59.4	-2.8	19.1	0.2	
1978	-77.5	60.9	-1.2	17.8	0.0	
AVGE	-60.2	43.4	3.2	14.5		

TABLE TOTAL- 10.0 MEAN- 0.2 STD. DEVIATION- 38.8

D10A. SEASONAL FACTORS, ONE YEAR AHEAD

YEAR	1ST QUAR	2ND QUAR	3RD QUAR	4TH QUAR	AVGE
1979	-78.3	61.4	-0.3	17.0	-0.1

En la tabla siguiente se muestran las observaciones en las que no se asignó un peso del 100% a sus irregulares.

Obs.	0	8	10	20	22	23	35	36
Peso (%)	34.7	58.5	00.0	36.0	20.7	65.7	0.00	47.7



5. - Serie ajustada estacionalmente.

=====

FREIGHT AND SHIPPING PAYMENTS PAGE 6, SERIES

D11. FINAL SEASONALLY ADJUSTED SERIES

YEAR	1ST QUAR	2ND QUAR	3RD QUAR	4TH QUAR	TOTAL
1969	256.2	241.0	242.1	256.1	995.3
1970	280.1	276.9	273.3	275.1	1105.4
1971	262.0	293.0	333.0	307.0	1195.0
1972	299.9	315.6	342.2	356.2	1313.9
1973	378.4	371.5	365.3	413.5	1558.7
1974	487.6	499.8	498.9	548.5	2034.9
1975	532.0	535.1	527.8	550.4	2145.3
1976	549.2	562.2	550.5	555.6	2217.5
1977	564.8	581.6	608.8	589.9	2345.1
1978	605.5	626.1	642.2	651.2	2524.9
AVGE	421.6	430.3	441.4	450.4	

TABLE TOTAL- 17436.0 MEAN- 435.9 STD. DEVIATION- 136.3

=====

FREIGHT AND SHIPPING PAYMENTS

=====

6. - Estadísticas de Calidad del Ajuste.

Comentarios.

- De F2.G, los irregulares a rezago 1 tienen un correlación alta, aunque la prueba de F3.4 no sea significativa.
- De F2.I, existe efecto estacional pero no es fijo en del tiempo, y la conclusión de la prueba combinada ( no mostrada) es la existencia de estacionalidad identificable.
- Las estadísticas F3.10 y F3.11 se encuentran en el límite de la zona crítica.
- Se acepta el ajuste estacional de la serie.

=====

F 2. SUMMARY MEASURES

F 2.A: AVERAGE DIFFERENCES WITHOUT REGARD TO SIGN OVER THE INDICATED

C18	SPAN						
	IN F	B1	D11	D13	D12	D10	A2
	QUARTERS	O	CI	I	C	S	P

TD	QC						
0.00	16.8	1	58.54	16.88	10.94	11.92	57.45 0.00
0.00	25.9	2	51.97	25.93	7.64	22.90	47.09 0.00
0.00	35.0	3	73.35	35.04	9.51	32.69	58.39 0.00
0.00	43.5	4	43.58	43.50	8.60	42.78	2.61 0.00

F 2.B: RELATIVE CONTRIBUTIONS TO THE VARIANCE OF THE DIFFERENCES IN T SPAN

TOTAL	RATIO (X100)	IN QUARTERS	D13	D12	D10	A2	C18
			I	C	S	P	TD
			1	3.36	3.99	92.65	0.00
100.00	103.95	2	2.08	18.73	79.18	0.00	0.00
100.00	103.67	3	1.98	23.40	74.62	0.00	0.00
100.00	84.91	4	3.87	95.77	0.36	0.00	0.00
100.00	100.58						

F 2.C: AVERAGE DIFFERENCES WITH REGARD TO SIGN AND STANDARD DEVIATION SPAN

D12	C	S. D.	IN QUARTERS	B1		D13		AVGE
				O	I	S. D.	AVGE	
				AVGE	S. D.	AVGE	S. D.	
11.39		1.64	1	11.77	71.66	-0.27	13.80	10.39
20.53		0.92	2	21.87	57.69	-0.20	11.82	21.16
26.77		2.32	3	34.22	75.71	-0.17	11.77	32.07
31.32		-0.02	4	42.47	33.57	-0.13	11.57	42.62

F 2.D: AVERAGE DURATION OF RUN CI I C QCD

1.77 1.26 3.90 1.77

F 2.E: I/C RATIO FOR QUARTERS SPAN

1 2 3 4  
0.92 0.33 0.29 0.20

QUARTERS FOR CYCLICAL DOMINANCE: 1

F 2.F: RELATIVE CONTRIBUTION OF THE COMPONENTS TO  
THE STATIONARY PORTION THE VARIANCE

I	C	S	P	TD	TOTAL
2.89	32.85	66.76	0.00	0.00	102.51

F 2.G: THE AUTOCORRELATION OF THE IRREGULARS  
FOR SPANS 1 TO 6

1	2	3	4	5	6
-0.45	-0.04	0.00	0.06	0.00	-0.04

F 2.H: THE FINAL I/C RATIO FROM TABLE D12: 0.66  
THE FINAL I/S RATIO FROM TABLE D10: 2.08

F 2.I:  
STATISTIC PROBABILITY LEVEL

F-TEST FOR STABLE SEASONALITY FROM TABLE B 1.  
69.101 0.00%

F-TEST FOR STABLE SEASONALITY FROM TABLE D 8  
89.673 0.00%

KRUSKAL-WALLIS CHI SQUARED TEST FOR STABLE SEASONALITY FROM  
TABLE D8  
31.234 0.00%

F-TEST FOR MOVING SEASONALITY FROM TABLE D 8.  
3.064 1.18%

FREIGHT AND SHIPPING PAYMENTS

PAGE 10, SERIES

F 3. MONITORING AND QUALITY ASSESSMENT STATISTICS  
ALL THE MEASURES BELOW ARE IN THE RANGE FROM 0 TO 3 WITH AN  
ACCEPTANCE REGION FROM 0 TO 1.

1. THE RELATIVE CONTRIBUTION OF THE IRREGULAR OVER ONE  
QUARTER SPAN (FROM TABLE F 2.B). M1 = 0.338
2. THE RELATIVE CONTRIBUTION OF THE IRREGULAR COMPONENT  
TO THE STATIONARY PORTION OF M2 = 0.289  
THE VARIANCE (FROM TABLE F 2.F).
3. THE AMOUNT OF QUARTER TO QUARTER CHANGE IN THE IRREGULAR  
COMPONENT AS COMPARED TO THE M3 = 0.487  
AMOUNT OF QUARTER TO QUARTER CHANGE IN THE TREND-CYCLE  
(FROM TABLE F2.H).
4. THE AMOUNT OF AUTOCORRELATION IN THE IRREGULAR AS DESCRIBED  
BY THE AVERAGE DURATION M4 = 0.695  
OF RUN (TABLE F 2.D).

5. THE NUMBER OF QUARTERS IT TAKES THE CHANGE IN THE TREND-CYCLE TO SURPASS THE AMOUNT  $M5 = 0.416$  OF CHANGE IN THE IRREGULAR (FROM TABLE F 2.E).
6. THE AMOUNT OF YEAR TO YEAR CHANGE IN THE IRREGULAR AS COMPARED TO THE AMOUNT OF YEAR  $M6 = 0.767$  TO YEAR CHANGE IN THE SEASONAL (FROM TABLE F 2.H).
7. THE AMOUNT OF MOVING SEASONALITY PRESENT RELATIVE TO THE AMOUNT OF STABLE  $M7 = 0.300$  SEASONALITY (FROM TABLE F 2.I).
8. THE SIZE OF THE FLUCTUATIONS IN THE SEASONAL COMPONENT THROUGHOUT THE WHOLE SERIES.  $M8 = 0.683$
9. THE AVERAGE LINEAR MOVEMENT IN THE SEASONAL COMPONENT THROUGHOUT THE WHOLE SERIES.  $M9 = 0.843$
10. SAME AS 8, CALCULATED FOR RECENT YEARS ONLY.  $M10 = 1.050$
11. SAME AS 9, CALCULATED FOR RECENT YEARS ONLY.  $M11 = 1.050$

\*\*\* ACCEPTED \*\*\* AT THE LEVEL 0.52

\*\*\* CHECK THE 2 ABOVE MEASURES WHICH FAILED.

FREIGHT AND SHIPPING PAYMENTS

## CAPITULO III.

### EL FILTRO DE KALMAN.

En este capítulo se presentan los fundamentos del filtro de Kalman y su aplicación en series de tiempo para ajustar a los llamados modelos estructurales, que representan una alternativa a la metodología de Box y Jenkins.

Un modelo estructural está formado por componentes estocásticos o deterministas, que tengan una interpretación directa, y representen a las características sobresalientes de la serie de tiempo, como pueden ser, los componentes tendencia o estacional, ésta manera de modelar es opuesta a la propuesta por Box y Jenkins, donde se intenta modelar una transformación que convierte a la serie en estacionaria.

No es posible revisar todos los aspectos de los modelos estructurales de series de tiempo, para ésto se recomienda Harvey(1989) que resume mucho del trabajo reciente; Anderson y Moore(1979) hacen un estudio muy amplio del filtro de Kalman.

#### 3.1 INTRODUCCION.

Mostraremos que con la ayuda de los sistemas lineales introducidos en la sección 2.4, es posible representar y generalizar a varios modelos estadísticos comunes, para hacer patente la flexibilidad y generalidad de los sistema lineales.

Ejemplo 1. Modelo de Regresión Lineal.

Sea el modelo de regresión lineal con los supuestos comunes,

$$Y = C\beta + \varepsilon \quad (3.1)$$

donde  $C' = [C_1, \dots, C_n]$ ,  $C \in M(n, p)$ ,

$$\beta \in \mathbb{R}^p, \quad \varepsilon' = [\varepsilon_1, \dots, \varepsilon_p], \quad Y' = [y_1, y_2, \dots, y_n].$$

$$E(\varepsilon) = 0, \quad E(\varepsilon\varepsilon') = \sigma^2 I.$$

el modelo de regresión puede ser escrito como un sistema lineal, donde el vector de estados es fijo y representa al vector de

parámetros desconocidos,

$$\begin{aligned} \beta_k &= I \beta_{k-1}, & k &= 1, \dots, n. \\ y_k &= C_k' \beta_k + \varepsilon_k, & \beta_0 &= \beta. \end{aligned} \quad (3.2)$$

éste modelo puede ser generalizado y manejar coeficientes de regresión variables (La Motte y Mc Worthner, 1978), o diseños de experimentos no balanceados y errores correlacionados (Jones, 1985).

Ejemplo 2. Proceso Autoregresivo.

Sea  $\{z_t\}$  el proceso autoregresivo de orden  $p$  definido por

$$z_t = \sum_{j=1}^p z_{t-j} \alpha_j + \varepsilon_t, \quad (3.3)$$

puede ser representado con el sistema lineal

$$\begin{aligned} x_t &= \Phi x_{t-1} + e_t, \\ z_t &= C x_t + 0, \end{aligned} \quad (3.4)$$

o explícitamente,

$$\begin{bmatrix} x_{t,0} \\ x_{t,1} \\ x_{t,2} \\ \vdots \\ x_{t,p-1} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_p \\ & 1 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 \\ & \vdots & & & 0 \\ & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_{t-1,0} \\ x_{t-1,1} \\ x_{t-1,2} \\ \vdots \\ x_{t-1,p-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$z_t = [1, 0, \dots, 0] [x_{t,0}, x_{t,1}, \dots, x_{t,p-1}]',$$

donde  $x_{t,i}$  es la variable  $i$  del vector de estados en el momento  $t$  y representa el valor del proceso autoregresivo a tiempo  $t-1$ , es decir  $x_{t,i} = z_{t-1}$ , al sustituir el vector  $[1, 0, \dots, 0]$  por  $[1, 1, \dots, 1]$ ,  $z_t$  representaría la suma de  $p$  valores consecutivos de un proceso autoregresivo de orden  $p$ .

Ejemplo 3. Proceso de Promedios Móviles.

El proceso de promedios móviles definido por

$$z_t = \sum_{j=1}^q \beta_j \varepsilon_{t-j} + \varepsilon_t \quad (3.4)$$

puede ser escrito como

$$\begin{bmatrix} x_{t,0} \\ x_{t,1} \\ x_{t,2} \\ \vdots \\ x_{t,q} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1,0} \\ x_{t-1,1} \\ x_{t-1,2} \\ \vdots \\ x_{t-1,q-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$z_t = [1, \beta_1, \beta_2, \dots, \beta_q] [x_{t,0}, x_{t,1}, \dots, x_{t,q}]' + 0,$$

donde  $x_{t,j}$  representa a  $\varepsilon_{t-j}$ .

La representación de un modelo estadístico a través de un sistema lineal en su forma espacio-estado no es única, y después de construir la representación del modelo es importante verificar las propiedades de observabilidad y controlabilidad, pues éstas determinan las propiedades estadísticas del modelo.

Cuando el modelo estadístico se ha expresado en su forma espacio-estado, el filtro de Kalman nos sirve para evaluar la verosimilitud como función de los parámetros no incluidos en el vector de estados, y estimar recursiva y óptimamente al vector de estados; en los dos primeros ejemplos el vector de estados coincide con los parámetros de interés.

Ahora daremos una derivación elemental del filtro de Kalman, junto con sus propiedades básicas.

Comencemos por recordar un conocido resultado,

Resultado.- Si  $X$  y  $Y$  son dos v.a. distribuidas conjuntamente normal con parámetros (3.5)

$$\mathbb{E} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \quad \mathbb{V} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix},$$

entonces  $\mathbb{E}(X | Y = y)$  es una v.a. normal con vector de medias

$$\bar{x} + \Sigma_{xx}^{-1} \Sigma_{xy} (y - \bar{y}) \quad (3.6)$$

y matriz de covarianzas

$$\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} \quad (3.7)$$

Aplicaremos el resultado anterior a la sucesión de v.a.

$\langle X_j \rangle_{j=1}^N$  generadas por (3.8)

$$\begin{aligned} X_{k+1} &= \Phi X_k + Dw_k \quad k = 0, \dots, N, \\ Y_k &= CX_k + v_k, \end{aligned} \quad (3.8)$$

para estimar recursivamente a  $X_j$  con base en la información proporcionada por  $Y^j = [Y_0, \dots, Y_j]$ .

En el sistema (3.8)  $\langle w_k \rangle_{k=1}^N$  y  $\langle v_k \rangle_{k=1}^N$  son sucesiones de v.a.i.i.d. normales con media 0 y matrices de covarianza Q y R, e independientes entre sí. Además  $X_0$  y  $Y_0$  son dos v.a. distribuidas normal con parámetros

$$E \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} = \begin{bmatrix} \bar{x}_0 \\ C\bar{x}_0 \end{bmatrix} \quad V \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} = \begin{bmatrix} P & PC' \\ CP & CPC' + R \end{bmatrix} \quad (3.9)$$

De los supuestos anteriores y usando el resultado (3.8)

$$E\langle X_0 ; Y_0 \rangle = \bar{x}_0 + P C' (CP C' + R)^{-1} (Y_0 - C\bar{x}_0) \quad (3.10)$$

y

$$V\langle X_0 ; Y_0 \rangle = P - PC' (CPC' + R)^{-1} CP \quad (3.11)$$

Si denotamos a  $E\langle X_0 ; Y_0 \rangle$  con  $\hat{X}_{0/0}$  y a  $V\langle X_0 ; Y_0 \rangle$  por  $P_{0/0}$ , por ser  $X_0$  y  $Y_0$  v.a. normales

$$\hat{X}_{1/0} = E\{X_1 ; Y_0\} = \Phi \hat{X}_{0/0}, \quad (3.12)$$

$$P_{1/0} = V\{X_1 ; Y_0\} = \Phi P_{0/0} \Phi' + DQD', \quad (3.13)$$

y también

$$E\{ (X_1 - \hat{X}_{1/0}) (Y_1 - E\langle Y_1 ; Y_0 \rangle)' \} = P_{1/0} C', \quad (3.14)$$

luego entonces el vector  $[X_1', Y_1']'$  condicionado en  $Y_0$  es una v.a normal con vector de medias y matriz de covarianzas dadas por

$$\begin{bmatrix} \hat{X}_{1/0} \\ C\hat{X}_{1/0} \end{bmatrix} \quad \begin{bmatrix} P_{1/0} & P_{1/0} C' \\ CP_{1/0} & CP_{1/0} C' + R \end{bmatrix}$$

entonces la v.a.  $X_1$  condicionada en  $Y_0$  y  $Y_1$  es una v.a. normal con vector de medias

$$(3.16)$$



$$\hat{X}_{1/1} = \hat{X}_{1/0} + P_{1/0} C' C CP_{1/0} C' + R)^{-1} (Y_1 - C\hat{X}_{1/0})$$

y matriz de covarianzas

$$P_{1/1} \equiv P_{1/0} - P_{1/0} C' C CP_{1/0} C' + R)^{-1} CP_{1/0} \quad (3.17)$$

por lo tanto, la v.a.  $X_2$  condicionada con  $Y_0$  y  $Y_1$  es una v.a. normal con vector de medias

$$\hat{X}_{2/1} \equiv \Phi \hat{X}_{1/1} \quad (3.18.a)$$

y matriz de covarianzas

$$P_{2/1} \equiv \Phi P_{1/1} \Phi' + DQD' \quad (3.18.b)$$

repetiendo los argumentos, no es muy difícil probar que

$$(3.19.a) \quad \hat{X}_{k/k} = \hat{X}_{k/k-1} + P_{k/k-1} C' C CP_{k/k-1} C' + R)^{-1} (Y_k - C\hat{X}_{k/k-1})$$

$$(3.19.b) \quad \hat{X}_{k+1/k} = \Phi \hat{X}_{k/k}$$

$$(3.19.c) \quad P_{k/k} = P_{k/k-1} + P_{k/k-1} C' C CP_{k/k-1} C' + R)^{-1} CP_{k/k-1}$$

$$(3.19.d) \quad P_{k+1/k} = \Phi P_{k/k} \Phi' + DQD'$$

$$(3.19.e) \quad K_k = P_{k/k-1} C' C CP_{k/k-1} C' + R)^{-1}$$

que son las ecuaciones que definen al filtro de Kalman.

Algunos comentarios que pueden hacerse son.

Porque las v.a.  $X$  y  $Y$  son v.a. distribuidas conjuntamente normales, su distribución queda caracterizada por el vector de medias y la matriz de covarianzas, entonces el filtro de Kalman es un procedimiento para actualizar la función de densidad condicional de  $X$  dado  $Y$ , y esta idea ha servido para hacer generalizaciones.

La estimación de  $X_k$  cuando  $Y$  es un escalar, es invariante ante las transformaciones  $P_0^* = aP_0$ ,  $Q^* = aQ$ , y  $R^* = aR$ .

La serie de innovaciones o errores de pronóstico  $v_k = Y_k - \hat{C}X_{k/k-1}$  son v.a. tales que

$$E\langle v_k : Y^{k-1} \rangle = 0, \quad (3.20.a)$$

$$E\langle v_k v'_k : Y^{k-1} \rangle = CP_{k/k-1} C' + R, \quad (3.20.b)$$

$$E\langle v_k v'_l : Y^{k-1} \rangle = 0 \quad k > l, \quad (3.20.c)$$

(3.20.c) se obtiene porque  $\hat{C}X_{k/k-1}$  es la mejor estimación lineal de  $Y_k$  en el subespacio generado por  $Y^{k-1} = [Y_0, \dots, Y_{k-1}]$ , entonces  $v_k$  es ortogonal a toda combinación lineal de  $[Y_0, \dots, Y_{k-1}]$ , en particular a  $v_{k-p}$  que es función lineal de  $[Y_0, \dots, Y_{k-p}]$ .

Se ha supuesto que las matrices  $\Phi, D, Q, C$  y  $R$  son constantes y conocidas, pero el argumento cambiaría poco, si permitimos que  $\Phi_k, D_k, Q_k, R_k$  y  $C_k$  sean funciones de  $Y^{k-1} = [Y_0, \dots, Y_{k-1}]$ .

### 3.2 EL FILTRO DE KALMAN.

En esta sección se generalizan los resultados de la sección anterior, y se da una presentación más formal del filtro de Kalman (Chen, 1985).

Sea el sistema lineal

$$x_{k+1} = \Phi_{k+1} x_k + B_{k+1} u_k + D_{k+1} \xi_{k+1} \quad (3.21)$$

$$y_k = C_k x_k + H_k v_{k-1} + F_k \xi_k$$

en el que se incluye a los términos  $B_{k+1} u_k$  y  $H_k v_{k-1}$ , que sirven para representar la intervención de un control y posibles variables exógenas.

Los supuestos sobre el sistema (3.21) son,

A1.  $\Phi_{k+1}, D_{k+1}, B_{k+1}$  y  $u_k$  son funciones  $Y^k$ -medibles, para toda  $k$   $\Phi_{k+1}$  y  $B_{k+1}$  son acotadas por una constante (las cotas pueden depender de  $k$ ) y también

$$E \| u_k \|^2 < \infty, \quad E \| D_{k+1} \|^2 < \infty.$$

B1.  $C_k, F_k, H_k$  y  $v_{k-1}$  son funciones  $Y^{k-1}$  medibles para toda  $k$ ,  $C_k$  es acotada por una constante ( la cota puede depender de  $k$  ) y

$$E \|H_k v_{k-1}\|^2 < \infty \quad E \|F_k\|^2 < \infty$$

C1.  $E \|x_0\|^2 < \infty$  y  $x_0$  se distribuye normal con vector de medias  $x_{0/-1}$  y matriz de covarianzas  $P$ .

D1.  $\{\xi_k\}$  son v.a. mutuamente independientes, e idénticamente distribuidas según  $N(0, I)$ , además son independientes de  $x_0$ .

Teorema 3.0.

Suponga que el sistema (3.21) cumple las condiciones A1-D1, entonces los estimadores insegados de mínima covarianza para  $x_j$  dado  $Y^j$  pueden ser calculados recursivamente usando las ecuaciones que definen al filtro de Kalman.

$$\hat{x}_{k+1/k+1} = \hat{x}_{k+1/k} + K_{k+1} (y_{k+1} - C_{k+1} \hat{x}_{k+1/k} - H_{k+1} v_k) \quad (3.22.a)$$

$$\hat{x}_{k+1/k} = \Phi_{k+1} \hat{x}_{k/k} + B_{k+1} u_k \quad (3.22.b)$$

$$K_{k+1} = (P_{k+1/k} C_{k+1}' + D_{k+1} F_{k+1}') (C_{k+1} P_{k+1/k} C_{k+1}' + F_{k+1} F_{k+1}' + C_{k+1} D_{k+1} F_{k+1}' + F_{k+1} D_{k+1}' C_{k+1}')^{-1} \quad (3.22.c)$$

$$P_{k+1/k+1} = P_{k+1/k} - K_{k+1} (C_{k+1} P_{k+1/k} + F_{k+1} D_{k+1}') \quad (3.22.d)$$

$$P_{k+1/k} = \Phi_{k+1} P_{k/k} \Phi_{k+1}' + D_{k+1} D_{k+1}' \quad (3.22.e)$$

$$\hat{x}_{0/-1} = E x_0 \quad P_{0/-1} = P \quad D_0 = 0$$

De las expresiones (3.22), se observa que la matriz de ganancia  $K_k$  y las matrices de covarianza  $P_{k/k}$  y  $P_{k+1/k}$  dependen en general de  $Y^k$ , pero en el caso especial de que  $\Phi_k, D_k, C_k$  y  $F_k$  son deterministas, entonces  $K_k, P_k$  y  $P_{k/k-1}$  son deterministas y conocidas desde el momento  $t = 0$ .

Tambien, si suponemos  $k \geq j \geq 0$  y definimos

$$\hat{x}_k^j = E \langle x_k | Y^k, x_j \rangle \quad \text{y} \quad P_k^j = E \langle (x_k - \hat{x}_k^j)(x_k - \hat{x}_k^j)' | Y^k, x_j \rangle,$$

el teorema (3.0) da formulas recursivas para  $\hat{x}_k^j$ , si se sustituye  $\hat{x}_k^j$  por  $\hat{x}_k$ ,  $P_k^j$  por  $P_k$ , y se define  $\hat{x}_j^j = x_j$ ,  $P_j^j = 0$ .

Además, si  $\Phi_{k+1}, D_{k+1}, C_k, F_k, B_{k+1}, u_k$  y  $H_k, v_{k-1}$  son deterministas,

$E\xi_i = 0$ ,  $E\xi_i \xi_j' = \delta(i-j)I$  y  $E\xi_i x_0 = 0$  para todas  $i$  y  $j$ , entonces las ecuaciones (3.22), son las formulas recursivas para calcular los estimadores lineales e insesgados de mínima varianza (LUMVUE) para  $x_j$  condicionados en  $Y^j$ .

Porque el sistema (3.21) es muy general, en ésta tesis es suficiente con el sistema (3.23)

$$\begin{aligned} x_{k+1} &= \Phi_k x_k + G_k u_k \\ y_k &= C_k x_k + v_k \end{aligned} \quad k = 0, \dots, N. \quad (3.23)$$

donde  $\{u_k', v_k'\}$  son v.a.i.i.d. normales con media 0 y

$$E\langle [u_k', v_k']' [u_k', v_k'] \rangle = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix} \quad (3.24)$$

**Teorema 3.1.**

Sea el sistema 3.23,

a) Filtro de Kalman. Los estimadores insesgados de mínima covarianza para  $x_k$  dado  $Y^k$ , pueden ser calculados recursivamente por :

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k (y_k - C_k \hat{x}_{k/k-1}) \quad (3.25.a)$$

$$\hat{x}_{k+1/k} = \Phi_k \hat{x}_{k/k} \quad (3.25.b)$$

$$K_k = P_{k/k-1} C_k' C_k P_{k/k-1} C_k' + R_k)^{-1} \quad (3.25.c)$$

$$P_{k+1/k} = \Phi_k P_{k/k} \Phi_k' + G_k Q_k G_k' \quad (3.25.d)$$

$$P_{k/k} = P_{k/k-1} - K_k C_k P_{k/k-1} \quad (3.25.e)$$

$$\hat{x}_{0/0-1} = E x_0 \quad P_{0/0-1} = E \langle (x_0 - x_0)(x_0 - x_0)' \rangle$$

b) Suavizamiento en un punto fijo. Los estimadores insesgados de mínima covarianza para  $x_j$  dado  $Y^k$ , con  $k = j+1, j+2, j+3, \dots$ ; pueden ser calculados recursivamente por:

$$\hat{x}_{j/k} = \hat{x}_{j/k-1} + K_k^\alpha (y_k - C_k \hat{x}_{j/k-1}) \quad (3.26.a)$$

$$K_k^\alpha = P_{k/k-1}^\alpha C_k' (C_k P_{k/k-1}^\alpha C_k' + R_k)^{-1} \quad (3.26.b)$$

$$P_{k+1/k}^\alpha = P_{k/k-1}^\alpha (\Phi_k - K_k C_k') \quad (3.26.c)$$

$$P_{j/k} = P_{j/k-1} - P_{k/k-1}^\alpha C_k' (K_k^\alpha)^{-1} \quad (3.26.d)$$

donde  $P_{j/j-1}^\alpha = P_{j/j-1}$

c) Suavizamiento en un intervalo. Los estimadores insesgados de mínima covarianza para  $x_j$  dado  $Y^k$ , con  $j = k-1, k-2, k-3, \dots, 0$ ; pueden ser calculados recursivamente por:

$$\hat{x}_{j-1/k} = \hat{x}_{j-1/j-1} + AC(j-1) (\hat{x}_{j/k} - \hat{x}_{j/j-1}) \quad (3.27.a)$$

$$AC(j-1) = P_{j-1/j-1} \Phi_{j-1}' P_{j/j-1}^{-1} \quad (3.27.b)$$

$$P_{j-1/k} = P_{j-1/j-1} + AC(j-1) (P_{j/k} - P_{j/j-1}) A'(j-1). \quad (3.27.c)$$

d) Pronóstico. Los estimadores insesgados de mínima covarianza de  $x_j$  dado  $Y^k$ , con  $j = k+1, k+2, k+3, \dots$ ; pueden ser calculados recursivamente por

$$\hat{x}_{j+1/k} = \Phi_j \hat{x}_{j/k} \quad (3.28.a)$$

$$P_{j+1/k} = \Phi_j P_{j/k} \Phi_j' + G_j Q_j G_j' \quad (3.28.b)$$

Existen formulas equivalentes a las del teorema 3.1 en las que se actualizan a las matrices  $P_{t/t}^{-1}$ ,  $P_{t/t}^{1/2}$ , que en algunos casos resultan ser mejores numéricamente, y pueden consultarse en Anderson y Moore(1979) o en Bierman (1977).

$\hat{x}_{j/k}$  con  $k > j$  es conocido con el estimador suavizado de  $x_j$ .

## SERIES DE TIEMPO.

En esta sección se muestra la representación con un sistema lineal en la forma espacio-estado de los componentes de una serie de tiempo.

### Tendencia.

El componente tendencia ( $N_t$ ) puede ser modelado por los siguientes procesos :

a) Tendencia Localmente Lineal,

$$(I - B) \mu_t = \beta_{t-1} + \eta_t, \quad (3.29)$$

$$\beta_t - \beta_{t-1} = \xi_t,$$

donde  $\langle \eta_t \rangle$  y  $\langle \xi_t \rangle$  son v.a.i.i.d. normales no correlacionadas entre sí, con media cero y varianzas  $\sigma_\eta^2$  y  $\sigma_\xi^2$ . La representación con un sistema espacio-estado de (3.29.a) y (3.29.b) es

$$\begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix} \quad (3.30)$$

$$N_t = [1, 0] \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix},$$

b) Caminata Aleatoria Mas Ruido,

$$\mu_t = \mu_{t-1} + \eta_t, \quad (3.31)$$

$$N_t = \mu_t + \varepsilon_t,$$

donde  $\langle \eta_t \rangle$  y  $\langle \varepsilon_t \rangle$  son v.a.i.i.d. normales no correlacionados entre sí, con medias cero y varianzas  $\sigma_\eta^2$ ,  $\sigma_\varepsilon^2$ .

### Estacional.

El componente estacional es modelado por el proceso

$$(I + B + \dots + B^{L-1})s_t = w_t \quad (3.32)$$

donde  $\langle w_t \rangle$  son v.a.i.i.d normales con media cero y varianzas  $\sigma_w^2$ , cuando  $L = 4$  su representación en la forma espacio-estado es

$$\begin{bmatrix} s_{t,0} \\ s_{t,1} \\ s_{t,2} \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} s_{t-1,0} \\ s_{t-1,1} \\ s_{t-1,2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} w_t \quad (3.33)$$

una manera alternativa de modelar el comportamiento estacional es con el uso funciones trigonométricas, si  $\lambda_j = 2\pi j/L$ ,  $j = 1 \dots [L/2]$ , el efecto estacional al tiempo  $t$  es

$$S_t = \sum_{i=1}^{[L/2]} \gamma_i \cos \lambda_i t + \gamma_i^* \sin \lambda_i t \quad (3.34)$$

donde los coeficientes  $\gamma_i$  y  $\gamma_i^*$  son fijos o satisfacen una ecuación en diferencias estocástica.

Ruido Autoregresivo.

La representación de un proceso AR, ya fue dada al inicio de este capítulo.

Variables Exógenas

Un manera de representar al componente determinista

$$D_t = \sum_{i=1}^p x_{ti} \beta_i \quad (3.35)$$

es con el sistema

$$\begin{bmatrix} \beta_{t,1} \\ \beta_{t,2} \\ \vdots \\ \beta_{t,p} \end{bmatrix} = \begin{bmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & \ddots & \\ 0 & 0 & & 1 \end{bmatrix} \begin{bmatrix} \beta_{t-1,1} \\ \beta_{t-1,2} \\ \vdots \\ \beta_{t-1,p} \end{bmatrix} \quad (3.36)$$

$$D_t = [x_{t1}, \dots, x_{tp}] [\beta_{t,1}, \beta_{t,2}, \dots, \beta_{t,p}]'$$

donde  $\beta_{t,i}$  es la estimación del  $j$ -ésimo coeficiente en el momento  $t$ , y con una adecuada selección de las variables exógenas puede representarse una tendencia polinomial o trigonométrica, o una intervención.

Hasta ahora cada componente se ha representado por un sistema de la forma

$$x_t(j) = F(j) x_{t-1}(j) + G(j) u_t(j),$$

$$y_t(j) = C(j) x_t(j),$$

donde el índice  $i$  indica la observación  $i$ -ésima y  $j$  representa al  $j$ -ésimo componente, por lo tanto, bajo el supuesto de independencia entre los  $p$  componentes de la serie, un modelo para toda la serie es

$$\begin{bmatrix} x_t(1) \\ x_t(2) \\ \vdots \\ x_t(p) \end{bmatrix} = \begin{bmatrix} F(1) & & & \\ 0 & F(2) & & \\ & & \ddots & \\ 0 & & & 0 \quad F(p) \end{bmatrix} \begin{bmatrix} x_{t-1}(1) \\ x_{t-1}(2) \\ \vdots \\ x_{t-1}(p) \end{bmatrix} + \begin{bmatrix} G(1) & & & \\ & G(2) & & \\ & & \ddots & \\ & & & G(p) \end{bmatrix} \begin{bmatrix} u_t(1) \\ u_t(2) \\ \vdots \\ u_t(p) \end{bmatrix}$$

$$y_t = [C(1), \dots, C(p)] [x_t(1), \dots, x_t(p)] + \varepsilon_t$$

$$\begin{aligned} \text{o,} \quad x_t &= \Phi x_{t-1} + G u_t \\ y_t &= C x_t + \varepsilon_t \end{aligned} \quad (3.38)$$

donde  $u_t$  y  $\varepsilon_t$  son v.a.i.i.d. normales tales que

$$E\langle u_t \rangle = 0 \quad E\langle \varepsilon_t \rangle = 0 \quad E\langle [u_t, \varepsilon_t] [u_j, \varepsilon_j]' \rangle = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \delta_{(1-j)}$$

Con el modelo (3.37) podremos estimar optimamente a los componentes no observables tendencia y estacional, quienes aparecen en el vector de estados, de manera condicional en los valores de las matrices  $\Phi$ ,  $Q$ ,  $R$  y  $C$ .

Un modelo típico para el ajuste estacional de series de tiempo es el Modelo Básico Estructural (MBS) definido por,

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad (3.39.a)$$

donde  $\mu_t$  es la tendencia definida por (3.29),  $\gamma_t$  es el componente estacional definido por (3.32) y  $\varepsilon_t$  es un ruido blanco.

Despejando a  $\mu_t$  de  $\Delta^2 \mu_t = \xi_{t-1} + \Delta \eta_t$ , el modelo implicado por el MBS para  $y_t$  es

$$y_t = \frac{\eta_t}{\Delta} + \frac{\xi_{t-1}}{\Delta^2} + \frac{w_t}{S(B)} + \varepsilon_t \quad (3.39.b)$$

donde  $S(B) = I + B + \dots + B^{L-1}$  y  $\Delta = I - B$ .

La forma estacionaria de un modelo estructural se obtiene al aplicar el operador que convierta a  $y_t$  en estacionaria, por lo tanto, la forma estacionaria del MBS es

$$S(B) \Delta^2 y_t = \Delta S(B) \eta_t + S(L) \xi_{t-1} + S(L) \Delta^2 \varepsilon_t \quad (3.39.c)$$

y la forma reducida del MBS es



$$S(B)\Delta^2 y_t = (I + \theta_1 B + \dots + \theta_{L+1} B^{L+1}) a_t \quad (3.39.d)$$

donde los diversos errores han sido sustituidos por sólo uno ( $a_t$ ).

De (3.39.c) se puede obtener la función de autocovarianzas para  $S(B)\Delta^2 y_t$

$$\gamma(0) = 2\sigma_\eta^2 + L\sigma_\xi^2 + 6\sigma_w^2 + 4\sigma_e^2 \quad (3.39.e)$$

$$\gamma(1) = (L-1)\sigma_\xi^2 - 4\sigma_w^2 - 2\sigma_e^2$$

$$\gamma(2) = (L-2)\sigma_\xi^2 + \sigma_w^2$$

$$\gamma(k) = (L-k)\sigma_\xi^2 \quad k = 3, \dots, L-2$$

$$\gamma(L-1) = \sigma_\xi^2 + \sigma_e^2$$

$$\gamma(L) = -\sigma_\eta^2 - 2\sigma_e^2$$

$$\gamma(L+1) = \sigma_e^2$$

donde se puede observar, que el MBS corresponde a un  $MA(L+1)$  para  $S(B)\Delta^2 y_t$  con restricciones no lineales sobre los parámetros..

Hasta ahora no se han impuesto restricciones para el número y tipo de los componentes de una serie de tiempo, y para obtener buenas propiedades de los estimadores es necesario imponer ciertas restricciones, que se dan a continuación.

#### IDENTIFICABILIDAD.

Comenzaremos por distinguir entre una estructura y un modelo. Con un modelo se especifica una distribución para la variable de interés, y una estructura especifica los parámetros de esa distribución. Las definiciones de identificabilidad para un modelo y una estructura son (Harvey, 1989):

1) si dos estructuras tienen la misma función de densidad conjunta, decimos que son observacionalmente equivalentes,

2) una estructura es identificable, si no existe otra estructura que sea observacionalmente equivalente,

3) un modelo es identificable si todas sus posibles estructuras son identificables, si ninguna estructura es identificable se dice que el modelo está sub-identificado.

Puede decirse, que dos estructuras Gaussianas son definidas como observacionalmente equivalentes si en su forma estacionaria tienen una misma función de autocovarianzas, además, porque una única función de autocovarianzas esta asociada con algún proceso ARMA invertible (Box y Jenkins, 1976, p.195), cuando dos estructuras son observacionalmente equivalentes tienen la misma forma reducida.

Cuando el modelo estructural es invariante, puede ser escrito en su forma UCARIMA (Componentes no observados ARIMA), y es posible dar condiciones para su identificabilidad.

Sea la representación UCARIMA de  $y_t$

$$y_t = \sum_{m=0}^M \frac{\theta_m(B)}{\phi_m(B)\Delta_m(B)} \xi_{mt} \quad (3.40)$$

con los supuestos,

- 1a) todos los polinomios  $\phi_m$  son estacionarios y de orden  $p_m$ ,
- 1b) todos los polinomios  $\theta_m$  son invertibles aunque no necesariamente estrictamente invertibles, y de orden  $q_m$ ,
- 2) cada par de polinomios  $(\phi_m, \theta_m)$  no tienen factores en común,
- 3) los polinomios no estacionarios  $(\Delta_m)$  no tienen factores en común, y son de orden  $d_m$ ,
- 4) los ruidos  $\xi_{mt}$ ,  $m = 0, \dots, M$ , son mutuamente no correlacionados,
- 5) los ruidos son normales.

Bajo los supuestos (1-5), una condición suficiente para tener identificabilidad del modelo (3.40) es que,  $(p_m + d_m) \geq q_m$  para al menos  $M$  ( de los  $M + 1$  ) componentes. Si no existen restricciones en los polinomios  $\theta_m$  y  $\phi_m$  la condición sobre los grados de los polinomios es necesaria.

Tomemos como ejemplo el MBS, que puede ser escrito como

$$y_t = \frac{(1 + \theta B)}{\Delta^2} v_t + \frac{1}{S(B)} w_t + \varepsilon_t \quad (3.41)$$

donde  $(1 + \theta B)v_t = \eta_t - \eta_{t-1} + \xi_{t-1}$ , por el resultado  $(\theta, \sigma_v^2,$

$(\sigma_v^2, \sigma_e^2)$  es identificable, y porque la relación entre  $(\theta, \sigma_v^2)$  y  $(\sigma_\eta^2, \sigma_\xi^2)$  es uno a uno,  $(\sigma_\eta^2, \sigma_\xi^2, \sigma_w^2, \sigma_e^2)$  es identificable.

#### PROPIEDADES DE LOS ESTIMADORES.

Una de las razones para usar a los estimadores de máxima verosimilitud es su normalidad asintótica. Los estimadores de máxima verosimilitud para los parámetros de un sistema espacio-estado tenderan a tener distribución normal, si se imponen restricciones adicionales y propias de los sistemas lineales.

Las condiciones bajo las cuales  $\hat{\psi}$ , estimador de M.V. para el vector de parámetros desconocidos  $\psi$  se distribuya normal con

a) media  $\psi$  y

b) matriz de covarianzas  $M(\hat{\psi})$ ,

$$M(\hat{\psi}) = T^{-1} I A^{-1} C(\psi)$$

(3.42)

$$I A C(\psi) = \lim T^{-1} I C(\psi)$$

donde  $I(\psi)$  es la matriz de información.

son

- 1)  $\psi$  es un punto interior del espacio paramétrico.
- 2)  $\psi$  es identificable.
- 3) las derivadas de  $\log L$ , hasta de orden tres existen y son continuas en una vecindad de  $\psi$ .
- 4) el modelo espacio-estado es detectable y estabilizable.
- 5)  $I I - \mathfrak{E} L$  puede ser factorizado como  $\Delta(L)\phi(L)$ , donde  $\Delta(L)$  no tiene parámetros desconocidos y  $\phi(L)$  tiene todas sus raíces fuera del círculo unitario.

Debe recordarse que un sistema estabilizable, no es necesariamente estacionario.

Si las condiciones 2 a 5 son satisfechas y un sólo parámetro esta en la frontera, la distribución asintótica de su estimador máximo-verosímil esta relacionado con la distribución normal, y puede ser una distribución mitad-normal con varianza dada por la matriz de información y masa de  $1/2$  en la frontera. La

distribución de los otros parámetros no es afectada, y la distribución conjunta con el parámetro que está en la frontera sí es afectada.

Las condiciones para que se cumpla (4) en algunos sistemas son las siguientes.

- Caminata Aleatoria Mas Ruido.

El sistema lineal es detectable y estabilizable si  $\sigma_{\eta}^2$  es positiva, aunque  $\sigma_{\epsilon}^2$  sea cero.

- Tendencia Localmente Lineal.

El sistema es detectable y estabilizable si y solo si  $\sigma_{\xi}^2 > 0$ .

- Modelo Basico Estructural.

El modelo MBS es detectable y estabilizable si y solo si  $\sigma_{\xi}^2 > 0$  y  $\sigma_{\psi}^2 > 0$ , y cuando los modelos mencionados son detectables y estabilizables, las condiciones de regularidad son satisfechas si  $\psi$  es un punto interior.

#### PRUEBAS DE HIPOTESIS.

Un aspecto importante en el ajuste estacional es averiguar si los componentes tendencia y estacional son deterministas, en el MBS quiere decir  $\sigma_{\eta}^2 = \sigma_{\psi}^2 = \sigma_{\xi}^2 = 0$ , encontrandonos en la frontera del espacio paramétrico.

El método de cociente de Verosimilitudes supone que el parámetro que especifica a la distribución nula no se encuentra en la frontera del espacio paramétrico. Para probar la hipótesis  $\sigma_{\eta}^2 = \sigma_{\psi}^2 = \sigma_{\xi}^2 = 0$ , definamos los dos modelos y después daremos la estadística de prueba según Harvey y Franzini (1983).

En el primer modelo los componentes son deterministas, es decir

$$y_t = \alpha + \beta t + \sum_j \delta_j z_{tj} + \epsilon_t, \quad (3.43)$$

donde  $y_1, \dots, y_T$  son las observaciones,  $\alpha$  y  $\beta$  son los parámetros de la tendencia lineal,  $z_{tj}$  son las variables indicadoras de las

estaciones.  $\delta_j$  es el efecto de la estación  $j$ , y  $\varepsilon_t$  son v.a.i.i.d normales; en este modelo se tienen  $s+1$  parámetros y el método adecuado de estimación es mínimos cuadrados ordinarios (OLS).

En el segundo modelo, se supone que los componentes son estocásticos,

$$y_t = \mu_t + \gamma_t + \varepsilon_t \quad (3.44)$$

donde  $\mu_t$  y  $\gamma_t$  son los componentes tendencia y estacional del MBS, la tendencia es definida por

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \beta_t = \beta_{t-1} + \xi_t, \quad t=1, \dots, T, \quad (3.45)$$

y el componente estacional por

$$\sum_j \gamma_{t-j} = w_t, \quad t=1, \dots, T, \quad (3.46)$$

donde  $\{\xi_t\}$ ,  $\{\eta_t\}$  y  $\{w_t\}$  son v.a.i.i.d. normales no correlacionadas entre sí, y con varianzas  $\sigma_\xi^2$ ,  $\sigma_\eta^2$  y  $\sigma_w^2$ .

Para probar la hipótesis  $H_0 : \sigma_\eta^2 = \sigma_\xi^2 = \sigma_w^2 = 0$ , Harvey y Franzini expresan el MBS como un modelo de regresión. En primer lugar se expresa al MBS en su forma espacio-estado, definiendo a

$$a_t = ( \mu_t, \beta_t, \gamma_t, \gamma_{t-1}, \gamma_{t-2} )', \quad (3.47)$$

$$b_t = ( \eta_t, \xi_t, w_t, 0, 0 )'$$

se tiene  $a_t = C a_{t-1} + b_t$ ,

$$y_t = \bar{x}'_t a_t + \varepsilon_t,$$

para adecuadas elecciones de  $C$  y  $\bar{x}'_t$  (Ver 3.29 y 3.33).

Si  $a_0$  es fijo y con sucesivas sustituciones de  $a_t$ , es posible expresar a  $y_t$  por

$$y_t = x'_t a_0 + v_t \quad t = 0, \dots, T. \quad (3.48)$$

donde  $x_t = C^t \bar{x}'_t$   $t = 0, \dots, T$ .

$$v_t = \bar{x}'_t \left( \sum_{j=1}^t C^{t-j} b_j \right) + \varepsilon_t,$$

entonces el modelo de regresión para  $Y' = [ y_1, \dots, y_T ]$  es

$$Y = X a_0 + v, \quad (3.49)$$

con  $E\langle v \rangle = 0$ ,  $E\langle vv' \rangle = \sigma^2 \Omega = \sigma^2 ( I + \Omega^* )$ , donde,

$$[\Omega]_{st} = \delta(s-t) + [\Omega^*]_{st}, \quad [\Omega^*]_{st} = \frac{\min(s,t)}{\sum_{j=1}^s x'_{t-j} Q x_{s-j}}$$

$$Q = \sigma^{-2} E\langle b'_i b_i \rangle = \text{Diag}(\bar{\sigma}_\eta^2, \bar{\sigma}_\xi^2, \bar{\sigma}_v^2, 0, 0)$$

y  $\bar{\sigma}_\eta^2 = \sigma_\eta^2 / \sigma^2$ ,  $\bar{\sigma}_\xi^2 = \sigma_\xi^2 / \sigma^2$ ,  $\bar{\sigma}_v^2 = \sigma_v^2 / \sigma^2$  son las varianzas relativas

Conociendo a  $\bar{\sigma}_\eta^2, \bar{\sigma}_\xi^2$  y  $\bar{\sigma}_v^2$ , la suma ponderada de los residuales al cuadrado y el estimador de mínimos cuadrados generalizados (GLS)  $\hat{a}_o$ , pueden ser estimados sin invertir a la matriz  $\Omega$ , es decir

$$\begin{aligned} S\langle \bar{\sigma}_\eta^2, \bar{\sigma}_\xi^2, \bar{\sigma}_v^2 \rangle &= \langle y - X \hat{a}_o \rangle \Omega^{-1} \langle y - X \hat{a}_o \rangle & (3.50) \\ &= \frac{T}{\sum_{t=s+2}^T v_t^2 / f_t} \end{aligned}$$

donde  $v_t$  es el error de predicción a tiempo  $t$  y  $f_t = \sigma^{-2} \text{var}(v_t)$ .

Si la matriz  $\Omega$  es conocida, la región crítica e invariante para las transformaciones  $y^* = ry + XR$ ,  $r$  escalar y  $R$  vector, más potente para probar (Harvey y Franzini, 1983)

$$H_0: E\langle vv' \rangle = \sigma^2 I \quad \text{Vs.} \quad H_1: E\langle vv' \rangle = \sigma^2 \Omega$$

es

$$\frac{\langle y - \hat{X}a_o \rangle \Omega^{-1} \langle y - \hat{X}a_o \rangle}{\langle y - X \tilde{a} \rangle \langle y - X \tilde{a} \rangle} < C \quad (3.51)$$

donde  $\tilde{a}_o$  es el estimador de OLS para  $a_o$ .

Los valores críticos de la prueba pueden ser calculados con el método de Imhof (1961), y dependen de  $s$ ,  $T$  y los valores asignados a las varianzas relativas.

Para usar la prueba propuesta, es necesario asignar valores a  $\bar{\sigma}_\xi^2, \bar{\sigma}_\eta^2$  y  $\bar{\sigma}_v^2$ ; planteandose dos alternativas,

1) hacer todas las varianzas relativas iguales, es decir

$$Q = q \text{diag}(1, 1, 1, 0, 0). \quad (\text{Prueba } b^+),$$

2) hacer  $\bar{\sigma}_\eta^2 = \bar{\sigma}_v^2$  y  $\bar{\sigma}_\xi^2 = 0$  (Prueba  $b$ ),

$$Q = q \text{diag}(1, 0, 1, 0, 0),$$

después de escoger estas hipótesis alternativas, falta determinar un buen valor para  $q$ .

Si  $\theta$  es la proporción del valor esperado de la suma de los residuales al cuadrado de un ajuste OLS bajo la hipótesis alternativa explicada por la variabilidad de los coeficientes de la regresión,  $\theta$  y  $q$  están relacionados (La Motte y Mc Worther, 1978) por

$$\theta = \frac{q \operatorname{tr}(M \Omega^*)}{\operatorname{tr}(CM) + q \operatorname{tr}(M \Omega^*)} \quad (3.52)$$

$$\Omega = I + q\Omega^*, \quad M = I - X(X'X)^{-1}X'$$

Harvey y Franzini fijaron a  $q$  como función de  $\theta$ , y después de examinar numericamente la potencia de la prueba para elecciones de  $\theta = 0, 0.5$  y  $0.9$  con valores verdaderos de  $\theta = 0.5$  a  $0.9$ , se encontró que la selección de  $\theta = 1/2$  es un buen valor para las pruebas  $b$  y  $b^+$ .

Al comparar las potencias de las pruebas  $b(\theta=1/2)$ ,  $b^+(\theta=1/2)$ , Durbin-Watson, y la Durbin-Watson de cuarto orden, cuando  $T = 20$  y un nivel de significancia del 5%, contra las alternativas

$$1) Q = q \operatorname{diag}(1, 1, 1, 0, 0)$$

$$2) Q = q \operatorname{diag}(1, 0, 1, 0, 0)$$

$$3) Q = q \operatorname{diag}(1, 0, 0, 0, 0)$$

$$4) Q = q \operatorname{diag}(0, 0, 1, 0, 0)$$

concluyen que la estadística  $b(\theta=1/2)$ , es la que mejor se porta en los cuatro casos, y es la estadística elegida.

Los valores críticos para  $b(\theta=1/2)$  pueden encontrarse en Harvey y Franzini (1983), y aquí se dan algunos.

T	q	C	T	q	c
12	0.384	0.437	28	0.201	0.504
16	0.314	0.447	32	0.180	0.523
20	0.265	0.465	36	0.162	0.540
24	0.229	0.485	40	0.148	0.555

Otro artículo relacionado es el de Nyblom (1986), donde se da otra manera de escoger  $\theta$  para un modelo con sólo tendencia.

## EVALUACION DE LA VEROSIMILITUD.

Uno de los conceptos básicos en la Estadística es la verosimilitud, y bajo el supuesto de que  $y_t$  sólo depende de su pasado, es posible escribir a la función de densidad conjunta como

$$L(Y; \psi) = \prod_{t=1}^T p(y_t | Y^{t-1}), \quad (3.53)$$

donde  $Y^t = [y_1, \dots, y_t]$ .

En un sistema lineal donde el vector inicial y las variables de ruido tienen una distribución normal,  $y_t$  condicionado en  $Y^{t-1}$  tiene una distribución normal, cuyo vector de medias y matriz de covarianzas pueden ser actualizados con las formulas del filtro de Kalman, y de su derivación, puede observarse que  $y_t$  condicionada en  $Y^{t-1}$  es una v.a normal con media  $C_t \hat{x}_{t|t-1}$  y matriz de covarianzas  $F_t = CP_{t|t-1}C' + R$ . luego entonces, la verosimilitud puede ser escrita como

$$\log(L) = - (NT/2) \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log |F_t| - (1/2) \sum_{t=1}^T v_t' F_t^{-1} v_t$$

donde  $v_t = y_t - C_t \hat{x}_{t|t-1}$ . La expresión (3.54) es conocida como la descomposición en errores de predicción de la verosimilitud.

Debe observarse que el filtro de Kalman nos da una manera de evaluar a la verosimilitud, pero no siempre es la mejor numericamente, por ejemplo, si tenemos un proceso AR(p) debemos multiplicar y sumar matrices de  $p \times p$  cuando menos T veces, método que no tiene comparación con la evaluación de la expresión

$$\sum_{t=1}^T (y_t - \alpha_1 y_{t-1} - \dots - \alpha_p y_{t-p})^2 / w_t \quad (3.55)$$

con  $w_t = \text{Var}(y_t - \alpha_1 y_{t-1} - \dots - \alpha_p y_{t-p})$ ,  $y_t = 0$  para  $t < 1$ .

En el modelo univariado es posible concentrar a la verosimilitud como función de R, varianza del error de observación, y entonces las varianzas de los errores son expresadas relativas a R

$$\text{Var}(\varepsilon) = R/R = 1, \quad \text{Var}(\xi) = Q/R, \quad f_t = F_t/R, \quad (3.56)$$



con esta reparametrización del modelo es posible expresar a la verosimilitud como

$$\begin{aligned} \text{Log } L = & -(T/2)\log(2\pi) - (T/2)\log(R) - (1/2)\sum_{t=1}^T \log(f_t) \\ & - (1/(2R)) \left( \sum_{t=1}^T v_t^2 / f_t \right), \end{aligned} \quad (3.57)$$

el estimador máximo verosimil de  $R$  es

$$\hat{R} = \left( \sum_{t=1}^T v_t^2 / f_t \right) / T, \quad (3.58)$$

la verosimilitud concentrada es (3.59)

$$\text{Log}^* L = -(T/2)(\log(2\pi) + 1) - (1/2)\sum_{t=1}^T \log(f_t) - (T/2)\log(\hat{R})..$$

Cuando se desconoce el valor inicial del vector de estados y su matriz de covarianzas, existen varias alternativas.

1) Si el vector de estados es estacionario, es decir  $|\lambda_t(\Phi)| < 1$ , su matriz de covarianzas ( $P$ ) es la solución a la ecuación

$$P = \Phi P \Phi' + G Q G', \quad (3.60)$$

un algoritmo para evaluar la verosimilitud de un proceso ARMA con el filtro de Kalman después de resolver la ecuación (3.60), está en Gardner, Harvey y Phillips(1979).

2) Si existe la matriz límite de las matrices de  $P_{t/t}$ , Gersh y Kitagawa(1984) proponen usarla como matriz inicial, y como vector inicial sugieren  $x_{0/-1} = 0$  o el pronóstico de  $x(0)$ , obtenido al pronosticar la serie en orden invertido.

3.a) Cuando el vector de estados no es estacionario, el uso de una distribución inicial difusa es equivalente a la construcción de una distribución inicial propia a partir de los primeras  $n$  (dim. del vector de estados) observaciones, si el modelo es observable (Harvey, 1989, p. 122), y la verosimilitud de  $y_{n+1}, \dots, y_T$ , condicionada en  $y_1, \dots, y_n$  es igual a la expresión (3.54), con las sumas evaluadas a partir de  $n+1$ .

3.b) Si algunos elementos del vector de estados son no estacionarios y otros son estacionarios, sin pérdida de generalidad se puede suponer que los primeros  $d$  elementos del vector de estados no son estacionarios, entonces, la matriz de transición y el vector de observación deben tener la siguiente forma

$$\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 \\ 0 & \Phi_3 \end{bmatrix} \quad C = [C_1, C_2] \quad (3.61)$$

donde  $\Phi_1$  tiene dimensiones  $d \times d$ ,  $\Phi_2$  es  $d \times (n-d)$ ,  $\Phi_3$  es  $(n-d) \times (n-d)$  con  $|\lambda_i(\Phi_3)| < 1$ , y  $C$  particionado conformablemente; entonces es posible construir una distribución propia a partir de las primeras  $d$  observaciones, si  $(\Phi_1, C_1)$  es observable y la matriz de covarianzas inicial es

$$P_{0/-1} = \begin{bmatrix} kI & 0 \\ 0 & P \end{bmatrix}, \quad (3.62)$$

donde  $P$  es la matriz de covarianzas para los elementos estacionarios del vector de estados.

De Jong (1988) da un método para evaluar a la verosimilitud como función de cualesquiera valores asignados a el vector de estados inicial ( $\mu$ ) y su matriz de covarianzas ( $C$ ). Denotando a  $l(Y^T; \mu, C)$  por menos 2 veces el logaritmo de la verosimilitud de  $Y^T = [y_1, \dots, y_T]$ , tres de sus conclusiones son :

$$a) \lim_{C \rightarrow \infty} l(Y^T; \mu, C) - \log |C| = l(y; *, \infty),$$

i.e. si  $C = kI$  y  $k$  tiende a infinito la verosimilitud tiende a cero,

$$b) l(Y^T; *, \infty) - K = l(Y^T; \hat{\mu}, \hat{C}) = l(Y^T; \hat{\mu}, \hat{C} = 0),$$

porque  $l(Y^T; \hat{\mu}, C)$ , la menos 2 veces log-verosimilitud concentrada respecto a  $\mu$  es una función decreciente de  $C$ ,

c)  $l(Y^T; *, \infty)$  es menos dos veces la log-verosimilitud de la transformación lineal de  $Y^T$ , que convierte a los datos en invariantes respecto a  $\mu$ , y corresponde a la log-verosimilitud dada por Ansley y Kohn (1985) cuando  $C = kI$  y  $k = \infty$ .

Ansley y Kohn (1985) muestran que las operaciones de límite y filtrado no son intercambiables, dan un filtro de Kalman modificado ( $k$  es infinito), y estudian el efecto de  $k$  en  $P_{t/-1}$  y  $\hat{x}_{t/-1}$ .

## SISTEMAS INVARIANTES.

Si se usa el filtro de Kalman en un sistema invariante, es de interés saber, cuando se da la convergencia de la matriz de covarianzas, pues entonces las ecuaciones para el vector de ganancia y las varianzas del vector de inovaciones tambien convergen, lo que puede reducir mucho el tiempo de cómputo, las ecuaciones estacionarias para el filtro de Kalman son,

$$P_{t+1/t} = \bar{P}, \quad (3.63.a)$$

$$\bar{K} = \bar{\Phi} \bar{P} C' (C \bar{P} C' + R)^{-1} \quad (3.63.b)$$

$$\bar{\Phi} = \Phi - \bar{K} C, \quad (3.63.c)$$

$$\hat{x}_{t+1/t} = \bar{\Phi} \hat{x}_{t/t-1} + \bar{K} y_t, \quad (3.63.d)$$

$$\bar{P} - \bar{\Phi} \bar{P} \bar{\Phi}' + \bar{\Phi} \bar{P} C' (C \bar{P} C' + R)^{-1} C \bar{P} \bar{\Phi}' - G G' = 0, \quad (3.63.e)$$

las condiciones bajo las cuales se satisface la ecuación algebraica de Riccati (3.63.e) y convergen las matrices  $P_{t+1/t}$  se dan en los resultados siguientes (Chan, Goodwin y Smith, 1984, y Souza, Gevers y Goodwin, 1986)

### Resultado 1.

Si el modelo es estable ( $|\lambda_i(\bar{\Phi})| < 1$ ) y la matriz de covarianzas iniciales ( $P_{0/-1}$ ) es positiva definida,

entonces  $\lim_{t \rightarrow \infty} P_{t+1/t} = \bar{P}$ ,

con  $\bar{P}$  independiente de  $P_{0/-1}$ , y la convergencia se da una manera exponencial.

### Resultado 2.

Suponga que  $P_{0/-1} - \bar{P}$  es positiva definida o  $P_{0/-1} = \bar{P}$

entonces  $\lim_{t \rightarrow \infty} P_{t/t-1} = \bar{P}$ ,

si y solo si  $(\bar{\Phi}, C)$  es detectable.

Cuando modelamos a una serie de tiempo por

$$(I-B)N_t = \xi_t, \quad (I-B^k)S_t = \eta_t, \quad (3.64)$$

$$y_t = N_t + S_t,$$

donde  $\{\xi_t\}$  y  $\{\eta_t\}$  son v.a.i.i.d. normal e independientes entre sí, la matriz de transición  $\Phi$  no es detectable, por lo tanto, el límite de  $P_{t/t-1}$  no es independiente de  $P_{0/-1}$ , lo que se tiene es,

a) para cada  $P_{0/-1} = \lambda I$  existe  $\lim P_{t/t-1} = P(\lambda)$ , evidencia empírica,

b) Burrige y Hall(1987) muestran que la varianza de las inovaciones y el vector de ganancias,

$$F_t = CP_{t/t-1}C' + R, \quad (3.65)$$

$$K_t = \Phi P_{t/t-1}C'F_t^{-1},$$

si convergen independientemente de  $P_{0/-1}$ .

Los conceptos de observabilidad y detectabilidad han sido dados para sistemas invariantes, para el sistema no invariante

$$x_{k+1} = \Phi_k x_k + D_k \xi_{k+1}, \quad (3.66)$$

$$y_k = C_k x_k + F_k \xi_k,$$

Chen(1985) define a

$$\Phi_{i,i} = I \quad \Phi_{k,i} = \Phi_k \Phi_{k-1,i}$$

$$\Phi_k' = [ C_{1,0}' \Phi_{1,0}' C_{1,0}' \dots \Phi_{k,0}' C_{k,0}' ]$$

y demuestra que el LUMVE de  $x_j$  basado en  $Y^k$  ( $j \leq k$ ) denotado por  $\hat{x}_j(k)$  y  $P_j(k)$  su matriz de covarianzas, cumplen a) y b)

a)  $\lim_{\lambda \rightarrow \infty} P_j(k; P_{0/-1} = \lambda I) = I \lambda$  existe

b)  $E\{ \lim_{\lambda \rightarrow \infty} \hat{x}_j(k; P_{0/-1} = \lambda I) \} = x_j$

si y sólo si  $(A^+)$  es una inversa generalizada

$$\Phi_{j,0} \Phi_k^+ \Phi_k = \Phi_{j,0}. \quad (3.67)$$

También Chen(1985), generaliza los conceptos de observabilidad y detectabilidad, define que un sistema es estocasticamente observable, si  $\Phi_k$  es de rango completo a partir de cierta  $k$ , mientras que, un sistema es estocasticamente detectable, si se cumple (3.67) para alguna  $k$ .

## DIAGNOSTICOS.

Los diagnósticos son importantes porque permiten detectar cuando un modelo propuesto no esta bien especificado.

Las estadísticas básicas para calificar el ajuste de un modelo son los residuos, definidos como las inovaciones estandarizadas

$$v_t' = v_t / (f_t')^{1/2} \quad t = n, \dots, T. \quad (3.88)$$

que son v.a.i.i.d.  $N(0, \sigma^2)$  si el modelo es conocido con certeza.

En la practica se recomienda trabajar con los residuos centrados y estandarizados

$$v_t^* = (v_t' - \bar{v}_t') / \text{Var}(v_t')^{1/2} \quad t = n, \dots, T. \quad (3.89)$$

y De Jong(1988) hace notar que las inovaciones estimadas son funciones lineales del vector de estados inicial, es decir

$$v_t(\mu) = v_t(0) - M_t \mu.$$

Las variables  $v_t^*$  seran aproximadamente  $N(0,1)$  (Harvey,1989,p.257), solo si conocemos a todos los parámetros con excepción de  $\sigma^2$ .

A las estadísticas  $v_t^*$  pueden ser aplicados todos los diagnósticos que comunmente se usan en los residuos de una serie de tiempo, Box y Jenkins (1976, p. 285-294), Harvey (1989,p. 256-271) y Harvey(1981, p.146-155), teniendo en cuenta que el número de grados de libertad que se pierden son

$$g.l. = \# \text{parametros estimados} - 1.$$

porque los residuos estimados son independientes de  $\sigma^2$ .

Algunos de los diagnósticos que se pueden usar son :

- a) la suma acumulada de residuales CUSUM, Harvey (1981 p.151-4),
- b) probar que un conjunto de autocorrelaciones muestrales son cero (Ljung y Box,1978),
- c) el periodograma acumulado (Box y Jenkins,1976, p. 294 ).

## OBSERVACIONES FALTANTES.

Una de las virtudes de usar el Filtro de Kalman (KF), para evaluar la verosimilitud de un modelo, es la facilidad par manejar observaciones faltantes.

Sea el sistema

$$\begin{aligned}x_{k+1} &= \Phi x_k + Gu_k, \\y_k &= Cx_k + \varepsilon_k, \quad .k = 1, \dots, T,\end{aligned} \quad (3.70)$$

y supongamos  $r$  es una observación faltante, este hecho puede ser representado muy facilmente con el sistema

$$\begin{aligned}x_{k+1} &= \Phi x_k + G u_k, \quad k+1 = 1, \dots, r-1, r+2, \dots, T, \\x_{r+1} &= \Phi^2 x_{r-1} + \Phi G u_{r-1} + G u_r, \\y_k &= C x_k + e_k \quad k = 1, \dots, r-1, r+1, \dots, T,\end{aligned} \quad (3.71)$$

y en terminos operativos, las ecuaciones del filtro de Kalman son modificadas por las siguientes reglas (Jones 1985)

$$\begin{aligned}K_r &= 0, \\P_{r/r} &= P_{r/r-1}.\end{aligned} \quad (3.72)$$

## OBSERVACIONES EXTREMAS.

Las observaciones extremas son observaciones inconsistentes con el modelo especificado, y mencionaremos 2 alternativas para manejarlos.

Si es posible asociar a la observación discordante un evento conocido, la manera de tratarlo es con la inclusión de una variable indicadora, que toma el valor de 1 en la observación discordante y 0 en otro caso.

Otra alternativa es suponer el siguiente sistema

$$\begin{aligned}x_{k+1} &= \Phi x_k + Gu_k (1 + w_t^i), \\y_t &= Cx_t + \varepsilon_t (1 + w_t^a),\end{aligned} \quad (3.73)$$

donde  $w_t^i$  y  $w_t^a$  son escalares, cuando  $w_t^i$  es diferente de cero,

representa un ruido extraordinario que influirá en  $y_{t+1}, y_{t+2}, \dots$ .  
Con  $w_t^a$  es posible modelar un ruido que sólo se manifiesta en  $y_t$   
(Tiao, 1985).

Por último, con el filtro de Kalman, podemos monitorear la  
estimación de los parámetros que aparecen en el vector de estados,  
ayudandonos a detectar posibles cambios estructurales.

### 3.4 EJEMPLO.

En esta sección se muestra y comenta el ajuste estacional de la serie de tiempo ya analizada en la sección 2.1. La salida del programa K (Apéndice B) está delimitada por el caracter "=".

#### 1.- Descripción de los componentes de la serie de tiempo.

##### Matriz de Estados

=====  
Variables regresoras : 0  
Efectos no observables : 2

##### Descripción de los efectos no observables

tipo : Tendencia  
    rezago max. : 2  
    Coeficientes  
        a( 1 ) = 2.000000 a( 2 ) = -1.000000  
tipo : Estacional  
    rezago max. : 3  
    Coeficientes  
        a( 1 ) = -1.000000 a( 2 ) = -1.000000 a( 3 ) = -1.000000

##### Matriz de Observacion

=====  
Variables regresoras : 0  
Efectos no observables : 2

##### Descripción de efectos no observables

Tipo	Rezago	Lugar
Tendencia	0	0
Estacional	0	2



2. Parámetros estimados y matriz de covarianzas del último vector de estados (dividida por R ).

=====

Var. de ruidos

Tipo

Tendencia 0.100000

Estacional 0.120000

Varianza del error de medicion 142.587603

log-likelihood -166.620

hiper-par  $q_1 = .10$   $q_2 = .12$   $r = 11.9410$

err. pred = 28.76 #par = 3 AIC = 339.24

Matriz de covarianzas

	0	1	2	3	4
0	0.918				
1	0.449	0.390			
2	-0.378	-0.146	0.665		
3	-0.056	-0.066	-0.123	0.481	
4	0.082	0.016	-0.123	-0.166	0.471

Vector de estados

652.782 638.253 14.458 2.736 61.839

=====

Los cocientes señal a ruido estimados no pertenecen a la malla propuesta por Gersh y Kitagawa(1984), y se obtuvieron de maximizar la verosimilitud a través de una búsqueda. Aunque los cocientes señal a ruido estimados son diferentes a los encontrados al maximizar en la malla, los componentes estimados difieren poco, cambian significativamente la función de autocorrelación para los residuales y las matrices de covarianza.

### 3.- Pronósticos para la serie y sus componentes.

Datos : 40            Componentes : 3

de 40 a 45

	Serie	Obs y Est	Tendencia	Estacional
40	0.000	588.278	667.311	-79.033
41	0.000	743.679	681.840	61.839
42	0.000	699.105	696.369	2.736
43	0.000	725.356	710.898	14.458
44	0.000	646.394	725.427	-79.033
45	0.000	801.796	739.956	61.839

Varianzas de los componentes y de la Variable de respuesta

	Tendencia	Estacional	V. Ser.
40	2.572	1.138	4.218
41	5.993	1.164	7.370
42	11.810	1.173	12.957
43	20.655	1.358	20.710
44	33.162	1.831	36.122
45	49.962	1.856	52.296

Los factores estacionales son casi iguales a los pronosticados por X-11-ARIMA, y el pronóstico para la serie es muy parecido para las tres observaciones siguientes.

La función de autocorrelación implícita para  $(1-B)(1-B^{12})x_t$  es

Rezago	1	2	3	4	5
	-.4579	.0672	.1890	-.4201	.210

que no es idéntica a la obtenida con X-11 y posibles razones son:

a) el modelo  $(0,1,1)(0,1,1)$ , ajustado por X-11-ARIMA, minimiza la suma de cuadrados de los residuales e impone como restricción que  $\rho(2) = 0$ ,

b) KF maximiza la verosimilitud con otras restricciones sobre la función de autocorrelación para  $(1-B)(1-B^{12})x_t$  (Ver 3.39),

c) KF maximizó la verosimilitud de  $y_{d+1}, \dots, y_T$ , condicionada en  $y_1, \dots, y_d$ , d es la dimensión del vector de estados (Khon y Ansley, 1985, p.1292, y Harvey, 1989, p.127).

d) la parametrización de la verosimilitud como función de los cocientes señal a ruido, dá como resultado una función mas o menos plana.

4.- Estimación de los componentes tendencia después de aplicar el algoritmo de suavizamiento (Gráfica 3.1).

```

=====
Datos      :          40          Componentes      :          3
de 0 a 39
=====

```

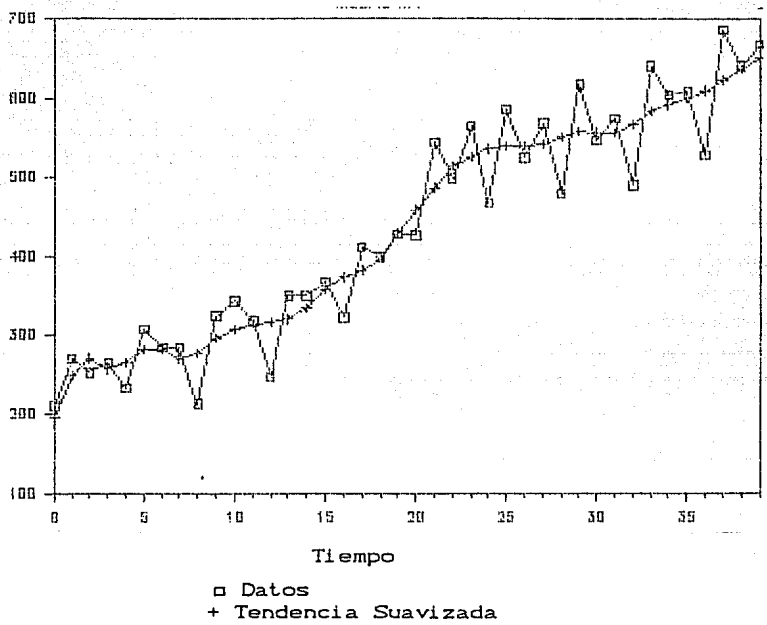
	Serie	Obs y Est	Tendencia	Estacional
0	210.000	210.000	194.476	15.524
1	269.000	269.000	248.381	20.619
2	252.000	252.000	270.629	-18.630
3	265.000	266.917	256.812	10.105
4	232.000	225.491	264.212	-38.721
5	306.000	312.981	280.661	32.320
6	284.000	289.781	282.009	7.773
7	284.000	277.848	270.159	7.689
8	212.000	220.269	276.783	-56.514
9	324.000	326.782	295.299	31.483
10	343.000	328.248	306.730	21.518
11	317.000	323.580	313.544	10.035
12	247.000	253.329	315.663	-62.334
13	350.000	350.893	319.636	31.257
14	350.000	348.647	335.193	13.454
15	368.000	373.043	358.458	14.584
16	322.000	318.353	373.269	-54.916
17	411.000	417.360	382.303	35.057
18	399.000	397.370	397.098	0.272
19	428.000	441.816	429.052	12.763
20	427.000	407.653	459.454	-51.801
21	545.000	534.411	487.269	47.142
22	499.000	501.070	513.633	-12.563
23	565.000	551.497	526.144	25.353
24	467.000	475.129	536.505	-61.376
25	586.000	590.587	540.342	50.244
26	525.000	526.408	539.834	-13.425
27	569.000	585.127	541.501	23.627
28	479.000	485.543	552.107	-66.564
29	618.000	616.431	558.516	57.915
30	547.000	549.043	557.554	-8.510
31	575.000	574.058	556.818	17.240
32	490.000	495.036	567.917	-72.881
33	641.000	644.019	584.433	59.586
34	606.000	595.006	592.367	2.640
35	609.000	613.392	600.237	13.155
36	528.000	532.516	611.158	-78.642
37	687.000	685.953	623.823	62.130
38	641.000	641.157	638.428	2.728
39	669.000	667.240	652.782	14.458

```

=====

```

Gráfica 3.1  
Ajuste Estacional  
Método KF.



5.- Varianzas de los componentes después de aplicar el algoritmo de suavizamiento.

=====

Varianzas de los componentes y de la Variable de respuesta

	Tendencia	Estacional	V. Ser.	Inovac.	Var. c	Valor T
0	-1778865.590	-1778860.508	1.000	210.000	8000001.6	0.006
1	-5238.953	-5238.072	1.000	85.250	7875004.0	0.003
2	-24847.991	-24847.264	1.000	-110.635	2634925.9	-0.006
3	0.210	0.802	0.955	-114.120	2048197.3	-0.007
4	0.217	0.427	0.445	-262.141	752948.2	-0.025
5	0.258	0.446	0.614	14.000	7.343	0.464
6	0.258	0.413	0.475	-0.022	6.535	-0.001
7	0.260	0.414	0.495	-15.006	6.530	-0.492
8	0.266	0.389	0.491	-46.636	6.090	-1.583
9	0.272	0.390	0.510	32.115	5.912	1.106
10	0.271	0.383	0.508	60.452	5.891	2.086
11	0.274	0.383	0.512	-15.641	5.871	-0.541
12	0.274	0.378	0.510	-31.288	5.791	-1.089
13	0.275	0.379	0.512	-3.221	5.765	-0.112
14	0.275	0.377	0.512	1.564	5.755	0.055
15	0.275	0.377	0.512	33.500	5.752	1.170
16	0.275	0.376	0.512	27.252	5.739	0.953
17	0.276	0.376	0.513	-10.350	5.735	-0.362
18	0.276	0.376	0.513	-24.325	5.732	-0.851
19	0.276	0.376	0.513	13.562	5.731	0.474
20	0.276	0.376	0.513	58.574	5.728	2.049
21	0.276	0.376	0.513	37.203	5.728	1.302
22	0.276	0.376	0.513	-45.040	5.727	-1.578
23	0.276	0.376	0.513	16.716	5.727	0.585
24	0.276	0.376	0.513	-68.921	5.726	-2.412
25	0.276	0.377	0.514	-18.346	5.726	-0.642
26	0.276	0.377	0.514	-24.010	5.726	-0.840
27	0.276	0.377	0.514	-4.750	5.726	-0.166
28	0.276	0.378	0.515	-8.830	5.726	-0.309
29	0.276	0.380	0.518	27.540	5.726	0.984
30	0.276	0.381	0.519	-3.111	5.726	-0.109
31	0.276	0.382	0.520	-20.438	5.726	-0.715
32	0.276	0.389	0.528	-4.664	5.726	-0.163
33	0.277	0.395	0.541	25.995	5.726	0.910
34	0.277	0.397	0.544	41.089	5.726	1.438
35	0.285	0.408	0.544	-24.791	5.726	-0.868
36	0.304	0.465	0.686	-14.301	5.726	-0.500
37	0.304	0.478	0.697	14.823	5.726	0.519
38	0.328	0.479	0.713	7.179	5.726	0.251
39	0.916	0.665	0.825	10.077	5.726	0.353

=====

Las primeras varianzas son negativas debido a errores numéricos, originados porque las primeras matrices tienen elementos muy grandes.

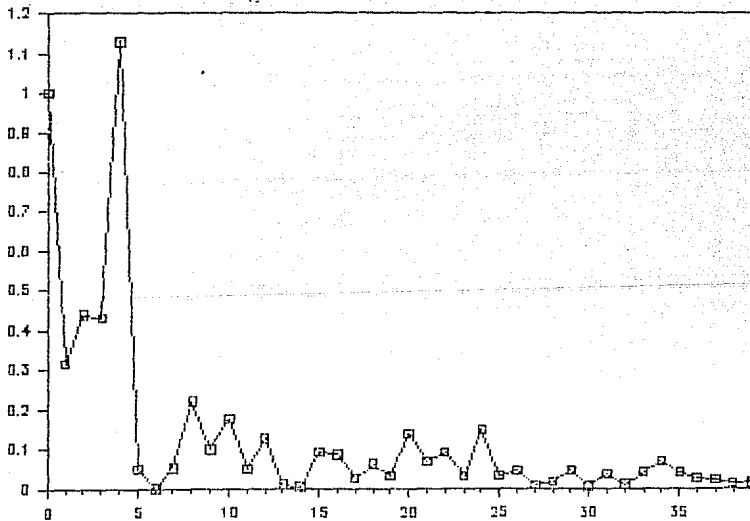
V. Ser. representa a la varianza de  $E(Cx_t; Y^t)$  dividida por R, es decir  $CCP_{t-1}/RDC'$

Inovac representa a los errores de predicción paso a paso, en la Gráfica 3.2 se muestran los errores relativos de predicción, y tienen como promedio a 0.056 (Comitiendo los primeros cinco).

Var. c. es igual a la varianza de los errores de predicción dividida por R, es decir  $CCP_{t-1}/RDC' + 1$ .

Valor T es el cociente entre la inovación y su desviación estandar, y nos permite identificar como posibles observaciones extremas a las observaciones 10,20, y 24.

Gráfica 3.2.  
Ajuste Estacional  
Método KF.



Tiempo

□ Error Relativo de Predicción

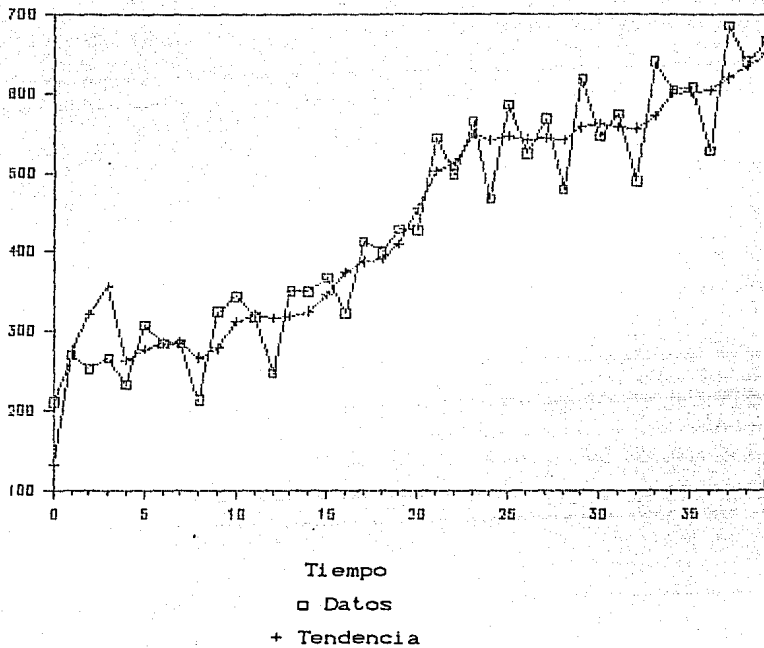
6.- Estimación de los componentes tendencia y estacional paso a paso, sin aplicar el suavizamiento (Gráfica 3.3).

Datos : 40            Componentes : 3

de 0 a 39

	Serie	Obs y Est	Tendencia	Estacional
0	210.000	210.000	131.250	78.750
1	269.000	269.000	274.952	-5.952
2	252.000	252.000	320.096	-68.096
3	265.000	265.000	355.553	-90.553
4	232.000	232.000	262.750	-30.750
5	306.000	303.957	275.927	28.030
6	284.000	284.003	284.583	-0.580
7	284.000	288.298	285.481	0.817
8	212.000	219.658	265.857	-46.199
9	324.000	318.568	277.610	40.958
10	343.000	332.737	312.047	20.691
11	317.000	319.664	319.807	-0.143
12	247.000	252.403	315.476	-63.073
13	350.000	350.559	318.900	31.659
14	350.000	349.728	324.139	25.589
15	368.000	362.176	346.902	15.274
16	322.000	317.251	374.180	-56.929
17	411.000	412.805	387.637	25.168
18	399.000	403.244	391.146	12.098
19	428.000	425.634	409.315	16.318
20	427.000	416.775	454.884	-38.109
21	545.000	538.505	502.719	35.786
22	499.000	506.864	515.061	-8.196
23	565.000	562.081	550.035	12.046
24	467.000	479.035	542.875	-63.840
25	586.000	589.204	546.722	42.481
26	525.000	529.193	543.212	-14.019
27	569.000	569.829	544.419	25.411
28	479.000	480.542	542.314	-61.772
29	618.000	613.190	557.698	55.492
30	547.000	547.543	563.067	-15.524
31	575.000	578.569	558.385	20.185
32	490.000	490.815	557.384	-66.570
33	641.000	636.460	571.779	64.681
34	606.000	598.824	600.401	-1.577
35	609.000	613.330	603.243	10.086
36	528.000	530.497	605.902	-75.405
37	687.000	684.411	620.866	63.545
38	641.000	639.746	635.201	4.545
39	669.000	667.240	652.782	14.458

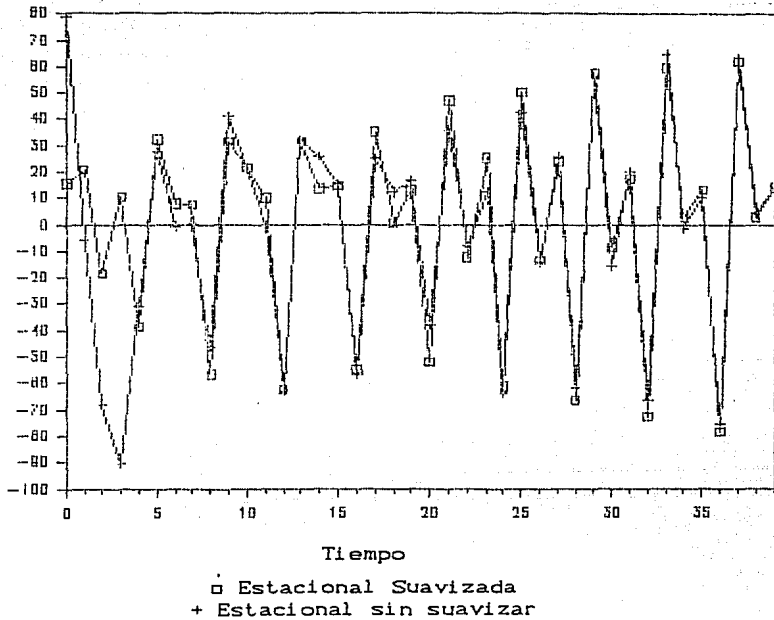
Gráfica 3.3  
Ajuste Estacional  
Método KF



En la Gráfica 3.4 se muestran las estimaciones del componente estacional antes y después de suavizar, y son significativamente diferentes sólo para los primeros cinco datos.



Gráfica 3.4  
Ajuste Estacional  
Método KF



Notar que el procedimiento de suavizamiento, tiene mayor efecto para las primeras observaciones, y después de cuatro observaciones, ya se tienen estimaciones creíbles de la tendencia.

7.- Varianza sin suavizar de la serie, componentes tendencia y estacional, las otras tres columnas se repiten.

=====  
 Varianzas de los componentes y de la

Variable de respuesta

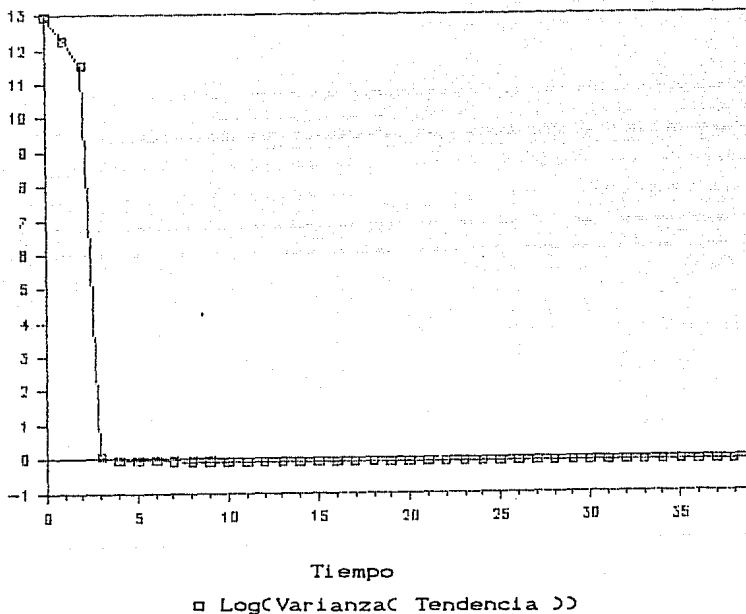
	Tendencia	Estacional	V. Ser.	Inovac.	Var. c	Valor T
0	428572.409	428571.885	1.000	210.000	8000001.663	0.006
1	210844.220	210844.148	1.000	85.250	7875004.046	0.003
2	105883.387	105883.517	1.000	-110.635	2634925.952	-0.006
3	1.071	0.821	1.000	-114.120	2048197.316	-0.007
4	0.977	0.817	0.864	-262.141	752948.246	-0.025
5	0.974	0.784	0.847	14.999	7.343	0.464
6	0.963	0.775	0.847	-0.022	6.535	-0.001
7	0.935	0.700	0.836	-15.006	6.530	-0.492
8	0.934	0.697	0.831	-46.636	6.090	-1.583
9	0.927	0.686	0.830	32.115	5.912	1.106
10	0.923	0.684	0.830	60.452	5.891	2.086
11	0.919	0.672	0.827	-15.641	5.871	-0.541
12	0.919	0.672	0.827	-31.288	5.791	-1.089
13	0.918	0.669	0.826	-3.221	5.765	-0.112
14	0.917	0.669	0.826	1.564	5.755	0.055
15	0.916	0.667	0.826	33.500	5.752	1.170
16	0.916	0.667	0.826	27.252	5.739	0.953
17	0.916	0.666	0.826	-10.350	5.735	-0.362
18	0.916	0.666	0.826	-24.325	5.732	-0.851
19	0.916	0.665	0.825	13.562	5.731	0.474
20	0.916	0.665	0.825	58.574	5.728	2.049
21	0.916	0.665	0.825	37.203	5.728	1.302
22	0.916	0.665	0.825	-45.040	5.727	-1.576
23	0.916	0.665	0.825	16.716	5.727	0.585
24	0.916	0.665	0.825	-68.921	5.726	-2.412
25	0.916	0.665	0.825	-18.346	5.726	-0.642
26	0.916	0.665	0.825	-24.010	5.726	-0.840
27	0.916	0.665	0.825	-4.750	5.726	-0.166
28	0.916	0.665	0.825	-8.830	5.726	-0.309
29	0.916	0.665	0.825	27.540	5.726	0.964
30	0.916	0.665	0.825	-3.111	5.726	-0.109
31	0.916	0.665	0.825	-20.438	5.726	-0.715
32	0.916	0.665	0.825	-4.664	5.726	-0.163
33	0.916	0.665	0.825	25.995	5.726	0.910
34	0.916	0.665	0.825	41.089	5.726	1.438
35	0.916	0.665	0.825	-24.791	5.726	-0.868
36	0.916	0.665	0.825	-14.301	5.726	-0.500
37	0.916	0.665	0.825	14.823	5.726	0.519
38	0.916	0.665	0.825	7.179	5.726	0.251
39	0.916	0.665	0.825	10.077	5.726	0.353

=====

Las dos primeras columnas muestran la convergencia de los elementos de la diagonal de las matrices  $P_{t|t}$ , como se ve es muy rápida, y practicamente para la cuarta observación se tiene un valor cercano al valor límite.

En la Gráfica 3.5 se muestra el logaritmo de la primera columna.

Gráfica 3.5  
Ajuste Estacional  
Método KF.



8.- Correlograma para las inovaciones estandarizadas, junto con las estadísticas q1 (Ljung y Box,1978) y q2(Box y Pierce,1970).

Correlogramas para la serie de inovaciones

- 1.- Calculado con media 0
- 2.- Calculado con media 0.46
- Media de las Inov. 0.46
- Se omitieron las primeras 5 inovaciones

k	r1(k)	r2(k)	q1(k)	q2(k)
0	1.0000	1.0000	0.0000	0.0000
1	0.0938	0.0924	0.3788	0.3680
2	-0.2796	-0.2815	3.8354	3.8717
3	-0.2729	-0.2748	7.2180	7.3012
4	0.0644	0.0630	7.4117	7.4866
5	0.1429	0.1416	8.3915	8.4492
6	0.1212	0.1199	9.1170	9.1591
7	-0.1306	-0.1323	9.9859	10.0505
8	-0.2366	-0.2384	12.9246	13.0348
9	-0.1334	-0.1351	13.8892	14.0239
10	0.1209	0.1196	14.7073	14.8244
11	0.0914	0.0900	15.1909	15.2939
12	-0.1589	-0.1606	16.7049	16.8408
13	0.0666	0.0652	16.9811	17.1057
14	0.0824	0.0810	17.4195	17.5298
15	-0.2080	-0.2098	20.3258	20.4863
16	-0.0432	-0.0448	20.4567	20.6267
17	0.0222	0.0207	20.4927	20.6581
18	0.1472	0.1460	22.1480	22.2852
19	0.1020	0.1007	22.9804	23.0960

Los cuantiles para una  $\chi^2$  son

$\alpha$	g.l.	
	8	17
0.9	13.36	24.77
0.95	15.51	27.59

por lo tanto, si usamos a q1(10) y q1(19), no se rechaza la hipótesis de que los residuales sean ruido blanco (por un margen pequeño), esto puede ser originado por la falta un término al modelo estructural.

## CAPITULO IV.

### SIMULACIONES.

En éste capítulo, se medirá la bondad del método X-11-ARIMA y del filtro de Kalman, en la estimación del componente tendencia de una serie de tiempo. En primer lugar, se explica la metodología de la simulación, para después, comparar a los métodos, y analizar aspectos propios de cada método.

#### 4.1 CONDICIONES DE SIMULACION.

La series son generadas por el modelo

$$Y = S + T + I \quad (4.1. a)$$

$$T_t - \delta T_{t-1} = \varepsilon_t(T), \quad \varepsilon_t(T) \sim N(0, \sigma_T^2), \quad (4.1. b)$$

$$S_t - \delta S_{t-4} = \varepsilon_t(S), \quad \varepsilon_t(S) \sim N(0, \sigma_S^2), \quad (4.1. c)$$

$$I_t = \varepsilon_t(I), \quad \varepsilon_t(I) \sim N(0, \sigma_I^2), \quad (4.1. d)$$

donde  $\varepsilon_t(T)$ ,  $\varepsilon_t(S)$  y  $\varepsilon_t(I)$  son variables aleatorias no correlacionadas a cualquier rezago.

Los parámetros de la simulación y las razones de su elección son :

a) el número de repeticiones por cada conjunto de parámetros fue de 200,

b) la estructura de las series es muy simple, esperando que ésta elección no favorezca a algún método,

c) el método X-11-ARIMA puede analizar con moderada exactitud series que satisfacen  $(I - B)(I - B^{12})X_t = A(B)c_t$ , para algún polinomio  $A(B)$  (Cleveland y Tiao, 1976),

d) las longitud de las series fueron 40, 80 y 116 datos, el último valor se eligió porque el programa X-11-ARIMA no ajusta un modelo ARIMA a series que abarquen más de 29 años (116 observaciones trimestrales),

e) se escogió trabajar con datos trimestrales, porque de haberse realizado la simulación con observaciones mensuales, el

uso del filtro de Kalman habría consumido mucho tiempo,

f) el valor de  $\delta$  fue muy cercano a 1 ( $\delta = 0.99$ ), para que las series se parecieran a las supuestas por Gersh y Kitagawa(1984), pero sin dejar de ser estacionarias, y así poder calcular la varianza de los componentes,

g) se decidió mantener fijo a  $\sigma_T^2 = 1$ , y modificar  $\sigma_S^2$  para que  $\lambda = \text{Var}(S) / \text{Var}(Y)$  tome los valores 0.1, 0.2 y 0.4,  $\lambda$  representa al porcentaje de varianza explicada por el componente estacional,

h) en todas las series  $T_0$  fue 0, porque de la tendencia sólo nos importa su nivel, ya que se puede trabajar con la serie  $x_t - x_0$  para evitar problemas numéricos,

i) los valores  $S_{-4}$ ,  $S_{-3}$ ,  $S_{-2}$  y  $S_{-1}$  fueron obtenidos de generar y centrar cuatro v.a.  $N(0, \sigma_S^2)$ ,

j) las variables aleatorias Normales simuladas, fueron obtenidas con el método de Box y Muller(1958),

k) las variables aleatorias uniformes usadas por el método de Box y Muller, fueron generadas con el algoritmo AS183 de la revista Applied Statistics(Wichman y Hill, 1982).

En la tabla siguiente, se muestran los parámetros resultantes de las condiciones impuestas,

Tabla 4.1.1

Varianzas.

$\lambda$	$\sigma^2(T)$	$\sigma^2(S)$	$\sigma^2(I)$	$\text{Var}(S)$
0.1	1.0	0.110912	0.01	5.5734
0.2	1.0	0.428372	0.01	21.5262
0.4	1.0	0.668467	0.01	33.4908

$$\text{Var}(S_t) = \sigma_S^2 / (1 - \delta^2) \quad \text{Var}(T_t) = \sigma_T^2 / (1 - \delta^2) \quad (4.2)$$

y los modelos ARIMA para las series simuladas son

$$(I - B)(I + B + B^2 + B^3)Y_t = C I - \psi_1 B^1 - \psi_2 B^2 - \psi_3 B^3 - \psi_4 B^4 \varepsilon_t \quad (4.3)$$

$$\varepsilon_t \sim N(0, \sigma^2)$$

Tabla 4.1.2

Var(S)/Var(T)	Parámetros				
$\lambda$	$\psi_1$	$\psi_2$	$\psi_3$	$\psi_4$	$\sigma^2$
0.1	.8299	.7758	.6681	-.0066	1.2284
0.2	.6904	.6088	.4677	-.0046	1.4872
0.4	.6283	.5400	.3957	-.0039	1.5946

#### 4.2 ESTIMACION DEL COMPONENTE TENDENCIA.

En esta sección, compararemos la eficiencia de los métodos X-11-ARIMA y KF en la estimación del componente tendencia, para los diferentes modelos y tamaños de muestra especificados en la sección 4.1.

El programa X-11-ARIMA se uso con dos opciones :

- a) el programa elige un modelo, opción denotada en todas las tablas por XA,
- b) el programa estima un modelo ARIMACO,0,4)CO,1,0)4, denominada opción XB.

la estimación de la tendencia se tomó de la tabla D11 del programa X-11-ARIMA, y tiene como título : "Estimación final de la tendencia despues de aplicar el filtro de Henderson".

El filtro de Kalman (KF) se uso con las condiciones siguientes:

- a) una distribución inicial impropia,
- b) se tomó como estimación final del componente tendencia a la estimación de la tendencia suavizada,
- c) el modelo estructural usado fue

$$y_t = \frac{\xi_t}{(I - B)} + \frac{\eta_t}{(I + B + B^2 + B^3)} + \varepsilon_t,$$

donde  $\{\xi_t\}$ ,  $\{\eta_t\}$  y  $\{\varepsilon_t\}$  son v.a.i.i.d con media cero y varianzas

$\sigma_{\xi}^2$ ,  $\sigma_{\eta}^2$  y  $\sigma_{\varepsilon}^2$ . El modelo corresponde a un sistema controlable y detectable si  $\sigma_{\xi}^2 > 0$  y  $\sigma_{\eta}^2 > 0$ , además, los parámetros son identificables porque se cumplen las condiciones de identificabilidad (Ver sección 3.2).

En algunas tablas obtenidas poco se pueden comentar, se presentan por completez, y se tuvo que recurrir a la definición de índices más o menos arbitrarios, porque el método X-11-ARIMA no supone modelo para los componentes.

Antes de pasar a comparar los métodos, analicemos algunos resultados particulares de cada método.

#### 4.2.1 METODO X-11-ARIMA.

En ésta sub-sección se dan los resultados de la simulación que son relevantes para el método X-11-ARIMA. Las tablas tienen un nombre y dos índices : los tamaños de muestra (columnas) y los valores del cociente  $\lambda = \text{Var}(S_t) / \text{Var}(Y_t)$  (englones).

Tabla 4.2.1.

	Sel		
( $\lambda$ )	40	80	(N)
0.1	187	179	179
0.2	191	184	183
0.4	191	185	184

Sel muestra el número de veces (de un total de 200) que el programa X-11 no pudo identificar un modelo para la serie, e hizo un ajuste estacional sin extender a la serie con pronósticos, los resultados de ésta tabla motivaron la necesidad de ejecutar el programa X-11-ARIMA con la opción XB.

Tabla 4.2.2.a

	Fail			XA	(N)
( $\lambda$ )	40	80	116		
0.1	0.795	0.93	1.095		
0.2	0.735	1.085	1		
0.4	0.885	1.205	1.1		

Tabla 4.2.2.b

	Fail			XB	(N)
( $\lambda$ )	40	80	116		
0.1	0.9600	0.960	1.095		
0.2	0.9000	1.090	1.030		
0.4	1.0600	1.240	1.125		



Fail indica el promedio de estadísticas de control para las que se obtuvo un valor de rechazo, de las tablas se puede ver que al aumentar  $T$  o  $\lambda$  aumenta Fail, no existen diferencias significativas entre XA y XB, y si tomamos en cuenta que Fail e [0,11] el valor que toma Fail es bajo.

Tabla 4.2.3.a

	Accept	XA	CN
( $\lambda$ )	40	80	116
0.1	196	198	200
0.2	197	198	200
0.4	197	198	200

Tabla 4.2.3.b

	Accept	XB	CN
( $\lambda$ )	40	80	116
0.1	196	198	200
0.2	197	198	200
0.4	198	198	200

Accept indica el número de veces que fue aceptado el ajuste estacional por la estadística global, en las tablas se observa poca diferencia entre XA y XB, y el porcentaje de aceptación es casi total.

Después de estudiar el comportamiento de la estadística global, analicemos a las 11 estadísticas que la componen.

Qual1.- Contribución relativa del componente irregular en periodos de tres meses.

Tabla 4.2.4.a

	Qual1	XA	CN
( $\lambda$ )	40	80	116
0.1	0.5986	0.614	0.5268
0.2	0.3196	0.3081	0.2516
0.4	0.2734	0.2592	0.2111

Tabla 4.2.4.b.

	Qual1	XB	CN
( $\lambda$ )	40	80	116
0.1	0.5510	0.6037	0.5212
0.2	0.2770	0.2988	0.2461
0.4	0.2339	0.2508	0.2062

Qual2.- Contribución relativa del componente irregular a la parte estacionaria de la varianza.

Tabla 4.2.5.a

	Qual2	XA	CN
( $\lambda$ )	40	80	116
0.1	0.3151	0.2204	0.1613
0.2	0.2292	0.1767	0.1367
0.4	0.2139	0.1694	0.1333

Tabla 4.2.5.b

	Qual2	XB	CN
( $\lambda$ )	40	80	116
0.1	0.2864	0.2150	0.1590
0.2	0.1979	0.1692	0.1328
0.4	0.1810	0.1616	0.1292

Qual1 y Qual2 decrecen al aumentar la varianza explicada por el componente estacional y toman en promedio valores bajos (de no rechazo).

Qual3.- El total de cambio cuarto a cuarto en el componente irregular, comparado con el cambio cuarto a cuarto del componente tendencia.

Tabla 4.2.5.a

	Qual3	XA	(N)
(λ)	40	80	116
0.1	0.4725	0.4841	0.4868
0.2	0.6336	0.6686	0.6583
0.4	0.7244	0.7737	0.7483

Tabla 4.2.5.b

	Qual3	XB	(N)
(λ)	40	80	116
0.1	0.4669	0.4769	0.4826
0.2	0.6282	0.6604	0.6544
0.4	0.7211	0.7663	0.7459

Qual4.- El total de autocorrelación en el componente irregular, medida por la duración promedio de las corridas.

Tabla 4.2.6.a

	Qual4	XA	(N)
(λ)	40	80	116
0.1	0.5729	0.9014	1.0671
0.2	0.6124	0.9888	1.1494
0.4	0.6156	1.0153	1.1804

Tabla 4.2.6.b

	Qual4	XB	(N)
(λ)	40	80	116
0.1	0.5786	0.8925	1.0684
0.2	0.6166	0.9909	1.1588
0.4	0.6387	1.0143	1.1873

Qual3 y Qual4 presentan tendencia a crecer cuando aumenta λ o N. Qual4 es la única estadística para la que se obtienen valores promedio de rechazo.

Qual5.- El número de cuartos necesarios para que la variabilidad de la tendencia sea mayor que la variabilidad de los irregulares.

Tabla 4.2.7.a

	Qual5	XA	(N)
(λ)	40	80	116
0.1	0.4625	0.449	0.4373
0.2	0.8066	0.5976	0.5948
0.4	0.6796	0.6694	0.6664

Tabla 4.2.7.b

	Qual5	XB	(N)
(λ)	40	80	116
0.1	0.4349	0.4403	0.4313
0.2	0.5666	0.5881	0.5860
0.4	0.6408	0.6583	0.6578

Qual6.- El total del cambio año con año en el componente irregular comparado con el total de cambio año con año en el componente estacional.

Tabla 4.2.8.a

Qual6 XA		(N)	
(λ)		40	80
0.1	0.4324	0.3315	0.3081
0.2	0.5752	0.6009	0.6301
0.4	0.6296	0.677	0.7191

Tabla 4.2.8.b

Qual6 XB		(N)	
(λ)		40	80
0.1	0.4638	0.338	0.3077
0.2	0.6396	0.611	0.6364
0.4	0.7017	0.689	0.7252

Qual5 y Qual 6 incrementan su valor cuando aumenta la variabilidad explicada por el componente tendencia, y presentan valores de no rechazo.

Las estadísticas restantes tuvieron en promedio valores de no rechazo y tienen un comportamiento de poco interés.

Qual7.- La estacionalidad estable comparada con la estacionalidad móvil.

Tabla 4.2.9.a.

Qual7 XA		(N)	
(λ)		40	80
0.1	0.2641	0.233	0.2002
0.2	0.2039	0.1943	0.1744
0.4	0.1937	0.1888	0.1711

Tabla 4.2.9.b.

Qual7 XB		(N)	
(λ)		40	80
0.1	0.2654	0.232	0.2004
0.2	0.2063	0.194	0.1746
0.4	0.1967	0.188	0.1713

Qual8.- El tamaño de las fluctuaciones del componente estacional en la serie completa.

Tabla 4.2.10.a.

Qual8 XA		(N)	
(λ)		40	80
0.1	0.5264	0.6016	0.5721
0.2	0.4486	0.5241	0.501
0.4	0.4359	0.5135	0.4922

Tabla 4.2.10.b.

Qual8 XB		(N)	
(λ)		40	80
0.1	0.5970	0.6205	0.5823
0.2	0.5218	0.5436	0.5133
0.4	0.5106	0.5339	0.5052

Qual9.- El promedio de los movimientos lineales del componente estacional en la serie completa.

Tabla 4.2.11.a.

Qual9 XA			
(λ)	40	80	CN)
0.1	0.427	0.3166	0.2368
0.2	0.381	0.3033	0.2348
0.4	0.3732	0.3021	0.2348

Tabla 4.2.11.b

Qual9 XB			
(λ)	40	80	CN)
0.1	0.4690	0.3216	0.2388
0.2	0.4273	0.3091	0.2372
0.4	0.4225	0.3089	0.2372

Qual10.- Igual que 8, pero calculado para los últimos tres años.

Tabla 4.2.12.a

Qual10 XA			
(λ)	40	80	CN)
0.1	0.665	0.6573	0.5852
0.2	0.5816	0.575	0.4865
0.4	0.5673	0.5625	0.4763

Tabla 4.2.12.b

Qual10 XB			
(λ)	40	80	CN)
0.1	0.7008	0.6870	0.5865
0.2	0.6193	0.6085	0.5192
0.4	0.6096	0.5995	0.5117

Qual11.- Igual que 9, pero calculada para los últimos tres años.

Tabla 4.2.13.a

Qual11 XA			
(λ)	40	80	CN)
0.1	0.6333	0.6273	0.5395
0.2	0.5622	0.5562	0.4691
0.4	0.5496	0.545	0.4601

Tabla 4.2.13.b

Qual11 XB			
(λ)	40	80	CN)
0.1	0.6659	0.6508	0.5517
0.2	0.5936	0.5849	0.4949
0.4	0.5860	0.5772	0.4883

La conclusión general es: con excepción de la estadística que mide la aleatoriedad de los residuales, se obtuvieron valores bastante bajos y de no rechazo.

#### 4.2.2 FILTRO DE KALMAN.

En ésta sub-sección, se estudia el comportamiento de la verosimilitud y las estimaciones de la tendencia en tres escenarios de simulación, cuando se ajustó estacionalmente a la serie con el filtro de Kalman.

La descripción de las tablas es la siguiente :

- a) ELEGIDOS muestra el número de veces que resultó elegido un juego de parámetros, es decir, maximizó a la verosimilitud dentro de la malla de búsqueda,
- b) NCASOS es el número de veces que se evaluó la verosimilitud en un punto. NCASOS incluye evaluaciones hechas en puntos vecinos a los máximos, ésto se hizo, para estudiar con más cuidado a la verosimilitud cerca de los puntos elegidos,
- c) el promedio de log-verosimilitud evaluada NCASOS veces se da en la tabla VEROSIMILITUD,
- e) en la tabla  $R^2(T, \hat{T})$  se muestra el promedio de los coeficientes de determinación entre las tendencias verdadera y estimada.

Las tablas de esta sección muestran la malla donde Gersh y Kitagawa sugieren maximizar a la verosimilitud, el índice de las columnas es el cociente  $\tau_s = \sigma_s^2 / \sigma_I^2$ , relación señal a ruido del componente estacional, y en los renglones aparece  $\tau_t = \sigma_t^2 / \sigma_I^2$ .

Escenario 1.

Tamaño muestra 40,  $\lambda = 0.1$ .

Tabla 4.2.14.a.

Número de evaluaciones de la verosimilitud.

$(\tau_t)$	NCASOS								
	1	2	4	8	16	32	64	128	256
1	5	6	7	2	1				
2	7	9	11	4	2				
4	11	13	15	5	3		1	1	1
8	10	15	17	7	2		1	1	1
16	14	24	28	14	4		1	1	1
32	53	91	109	57	19	1	21	21	21
64	64	102	120	57	19	1	54	54	54
128	58	90	105	48	16	1	54	54	54
256	15	19	19	4			33	33	33

Tabla 4.2.14.b.

Distribución empírica de  $(\tau_t, \tau_s)$

$(\tau_t)$	ELEGIDOS									
	1	2	4	8	16	32	64	128	256	$(\tau_s)$
1	1	0	0	0	0					
2	2	2	1	1	0					
4	1	1	1	1	0				0	
8	2	2	0	1	0		0	1	0	
16	3	1	4	0	0		0	0	0	
32	4	2	6	3	0	0	0	0	0	
64	25	18	28	15	1	0	0	21	0	
128	9	6	4	0	0	0	0	33	0	
256	0	0	0	0			0	0	0	

Tabla 4.2.14.c

Promedio de la log-verosimilitud.

VEROSIMILITUD

$(\tau_t)$	VEROSIMILITUD									
	1	2	4	8	16	32	64	128	256	$(\tau_s)$
1	-61.30	-60.00	-60.12	-55.28	-57.76					
2	-59.56	-57.96	-58.27	-54.89	-57.94					
4	-60.50	-59.24	-59.53	-56.06	-59.12			-77.02	-57.68	-57.68
8	-59.51	-58.14	-58.65	-56.71	-59.62			-75.38	-57.68	-57.68
16	-59.54	-59.20	-59.15	-58.65	-58.06			-72.92	-57.68	-57.68
32	-59.55	-59.58	-59.59	-59.32	-58.81	-57.79		-72.06	-58.43	-58.43
64	-59.34	-59.41	-59.19	-58.91	-57.97	-56.86	-66.46		-58.71	-58.71
128	-59.66	-60.17	-61.09	-62.15	-63.33	-62.15	-76.21		-58.66	-58.66
256	-57.31	-57.11	-57.39	-56.25			-73.62		-58.69	-58.69

Tabla 4.2.14.d.

Coef. de determinacion entre la tendencia observada y estimada.

$R^2(T, \hat{T})$

$(\tau_t)$	$R^2(T, \hat{T})$									
	1	2	4	8	16	32	64	128	256	$(\tau_s)$
1	0.971	0.968	0.958	0.922	0.869					
2	0.942	0.945	0.926	0.894	0.810					
4	0.941	0.943	0.923	0.884	0.815		0.975	0.981	0.981	
8	0.922	0.922	0.906	0.883	0.791		0.975	0.981	0.981	
16	0.932	0.938	0.935	0.934	0.913		0.976	0.980	0.980	
32	0.954	0.954	0.952	0.945	0.931	0.892	0.930	0.944	0.944	
64	0.947	0.948	0.949	0.943	0.935	0.898	0.931	0.941	0.941	
128	0.956	0.950	0.944	0.929	0.919	0.882	0.912	0.944	0.944	
256	0.938	0.917	0.910	0.821			0.908	0.943	0.943	

En éste escenario la log-verosimilitud es una función suave de los parámetros, y se tienen estimaciones razonables del componente tendencia en un rango amplio de los parámetros. Las selecciones del componente tendencia se agrupan en un valor, mientras que, existe dispersión en la selección del componente estacional.

Los parámetros estimados aparecen en toda la malla, pero se concentran en el valor de  $\sigma_T^2/\sigma_I^2 = 64$ .

Escenario 2.

Tamaño muestra 40,  $\lambda = 0.4$ .

Tabla 4.2.15.a

Distribución empírica de  $(\tau_t, \tau_s)$ .

$(\tau_t)$	NCASOS								
	$(\tau_s)$								
	1	2	4	8	16	32	64	128	256
1	2	9	9	11	4	4			
2	5	16	17	17	6	5	1	1	1
4	10	22	25	22	10	7	1	1	1
8	10	20	33	27	17	4	1	1	1
16	13	29	59	58	44	14	2		
32	15	39	90	102	88	38	12	2	1
64	33	59	101	94	78	37	22	12	11
128	27	43	67	57	49	26	20	12	11
256	20	27	28	8	1		10	10	10

Tabla 4.2.15.b

Distribución empírica de  $(\tau_t, \tau_s)$

$(\tau_t)$	ELEGIDOS								
	$(\tau_s)$								
	1	2	4	8	16	32	64	128	256
1	1	0	3	0	1	0			
2	1	0	4	0	3	0	0	0	0
4	1	2	4	1	1	0	0	1	0
8	2	3	1	2	2	0	0	0	0
16	2	0	5	10	1	0	0		
32	4	2	10	18	9	2	0	0	0
64	2	5	9	23	17	8	1	1	0
128	10	10	7	1	0	0	0	10	0
256	0	0	0	0	0		0	0	0

Tabla 4.2.15.c

Promedio de la log-verosimilitud.

VEROSIMILITUD									
$(\tau_t)$	$(\tau_a)$								
	1	2	4	8	16	32	64	128	256
1	-65.21	-64.89	-64.89	-63.93	-61.87	-62.51			
2	-63.87	-64.48	-64.56	-64.17	-62.64	-62.56	-64.50	-66.65	-66.65
4	-63.61	-64.17	-64.01	-64.18	-63.62	-64.64	-63.67	-66.65	-66.65
8	-64.22	-64.59	-64.89	-65.22	-65.70	-66.19	-62.27	-66.65	-66.65
16	-64.47	-64.89	-65.07	-65.46	-65.79	-66.40	-62.93		
32	-64.04	-65.06	-65.67	-66.08	-66.03	-65.85	-63.84	-69.02	-65.27
64	-65.22	-65.07	-65.39	-65.88	-65.67	-65.26	-65.69	-65.35	-64.68
128	-64.64	-64.69	-66.01	-67.96	-69.71	-70.46	-73.56	-65.12	-64.44
256	-65.03	-64.29	-64.39	-62.13	-61.80		-77.45	-64.35	-64.35

Tabla 4.2.15.d

Coef. de determinación entre la tendencia observada y estimada.

$R^2(T, \hat{T})$									
$(\tau_t)$	$(\tau_a)$								
	1	2	4	8	16	32	64	128	256
1	0.778	0.843	0.841	0.758	0.574	0.554			
2	0.827	0.844	0.849	0.789	0.678	0.604	0.933	0.946	0.946
4	0.853	0.858	0.863	0.814	0.759	0.685	0.933	0.945	0.945
8	0.881	0.874	0.885	0.876	0.878	0.834	0.933	0.943	0.943
16	0.892	0.902	0.916	0.919	0.911	0.883	0.743		
32	0.918	0.922	0.928	0.931	0.926	0.919	0.892	0.876	0.957
64	0.891	0.912	0.923	0.934	0.931	0.925	0.898	0.888	0.897
128	0.917	0.928	0.928	0.936	0.928	0.930	0.899	0.896	0.903
256	0.908	0.912	0.908	0.899	0.824		0.883	0.897	0.897

Al aumentar el porcentaje de variabilidad explicada por el componente estacional, y mantener el tamaño de muestra, la verosimilitud sigue siendo una función suave, los parámetros elegidos ya no son tan dispersos, y disminuye significativamente la estadística  $R^2$ .



Escenario 3.

Tamaño muestra 116  $\lambda = 0.4$

Tabla 4.2.16.a

Número de evaluaciones de la verosimilitud.

$(\tau_t)$	NCASOS							$(\tau_s)$
	1	2	4	8	16	32	64	
1	5	10	13	8	3			
2	11	20	23	12	3			
4	16	29	37	21	8			
8	19	42	59	41	18	1		
16	14	49	94	83	48	3		
32	10	47	112	130	95	30	2	
64	21	45	98	104	82	29	2	
128	20	28	53	58	52	27	2	
256	19	21	21	2				

Tabla 4.2.16.b

Distribución empírica de  $(\tau_t, \tau_s)$

$(\tau_t)$	ELEGIDOS							$(\tau_s)$
	1	2	4	8	16	32	64	
1	0	2	1	2	0			
2	0	3	4	1	0			
4	2	4	4	0	0			
8	2	3	4	5	0	0		
16	1	7	15	12	1	0		
32	0	1	16	28	2	0	0	
64	0	1	8	25	25	2	0	
128	5	14	2	0	0	0	0	
256	0	0	0	0				

Tabla 4.2.16.c

Promedio de la log-verosimilitud.

VEROSIMILITUD

$(\tau_s)$	VEROSIMILITUD							$(\tau_t)$
	1	2	4	8	16	32	64	
1	-203.92	-202.80	-203.25	-203.19	-204.91			
2	-202.52	-202.53	-203.14	-203.60	-204.67			
4	-203.00	-202.98	-203.18	-203.17	-204.05			
8	-200.71	-202.60	-203.25	-204.44	-205.27	-192.91		
16	-200.96	-202.58	-202.67	-203.44	-205.18	-202.98		
32	-199.99	-202.73	-202.37	-201.94	-202.95	-204.07	-202.83	
64	-208.12	-204.33	-202.91	-201.56	-201.14	-201.13	-198.73	
128	-201.08	-201.37	-203.29	-206.58	-213.24	-219.43	-219.16	
256	-201.28	-201.35	-202.38	-208.65				

Tabla 4.2.16.d.

Coef. de determinación entre tendencia observada y estimada.

$(\tau_t)$	$R^2(T, \hat{T})$						$(\tau_s)$
	1	2	4	8	16	32	
1	0.912	0.927	0.933	0.941	0.942		
2	0.933	0.929	0.931	0.925	0.942		
4	0.942	0.937	0.929	0.914	0.889		
8	0.966	0.954	0.947	0.936	0.925	0.954	
16	0.974	0.961	0.955	0.950	0.943	0.915	
32	0.979	0.960	0.958	0.955	0.953	0.948	0.979
64	0.941	0.953	0.955	0.954	0.954	0.949	0.979
128	0.954	0.957	0.956	0.953	0.951	0.948	0.978
256	0.952	0.956	0.956	0.977			

Al aumentar el tamaño de la muestra en relación al escenario dos, la verosimilitud sigue siendo una función suave de los parámetros, los parámetros elegidos no muestran tanta variabilidad como en el caso 2, y mejora el promedio de  $R^2$ .

Es claro que maximizar en una malla discreta no es del todo correcto, aunque la verosimilitud sea muy suave, se usó éste método porque fue la referencia inicial de trabajo, y también, porque Gersh y Kitagawa(1984) fueron los primeros autores en aplicar el filtro de Kalman al ajuste estacional.

Para entender la relación entre la representación espacio-estado y ARIMA de una serie de tiempo, consideremos el modelo correspondiente a una caminata aleatoria mas ruido

$$y_t = \frac{\xi_t}{I - B} + \varepsilon_t,$$

donde  $\langle \xi_t \rangle$  y  $\langle \varepsilon_t \rangle$  son v.a.i.i.d.  $(0, \sigma_\xi^2)$  y  $(0, \sigma_\eta^2)$ . En éste modelo  $(I-B)y_t$  es un MAC(1) con parametro  $\theta$ , y la relación entre  $\theta$  y  $q = \sigma_\xi^2 / \sigma_\varepsilon^2$  es  $q = \theta + 1/\theta - 2$ , donde se observa que cambios grandes de  $q$  se traducen en cambios pequeños para  $\theta$ .

La conclusión de esta sección es : la función de verosimilitud expresada a través de los cocientes señal a ruido es una función suave.

### 4.2.3 COMPARACION.

En esta sub-sección se comparan los métodos KF y X-11-ARIMA en la estimación de la tendencia.

El primer índice de comparación que definimos es (4.4)

$$E.C.M(Y) = \frac{\sum_{\text{simulaciones}} \sum_{\text{datos}} (y - \hat{y})^2}{\sum_{\text{simulaciones}} \sum_{\text{datos}} y^2}$$

e intenta medir la magnitud de las diferencias entre la serie observada y suavizada.

Tabla 4.2.17.a

C(\lambda)	Indice ECMCY), XA (N)		
	40	80	116
0.1	0.9923	0.9943	0.9956
0.2	0.9919	0.9936	0.9949
0.4	0.9917	0.9933	0.9946

Tabla 4.2.17.b

C(\lambda)	Indice ECMCY), XB (N)		
	40	80	116
0.1	0.9931	0.9944	0.9957
0.2	0.9930	0.9939	0.9951
0.4	0.9930	0.9936	0.9948

Tabla 4.2.17.c

C(\lambda)	Indice ECMCY), KF (N)		
	40	80	116
0.10	0.9997	0.9999	0.9999
0.40	0.9998	0.9999	0.9999
0.60	0.9999	0.9999	0.9999

De manera análoga a (4.4) se define ECMCS) y ECMCTD).

Tabla 4.2.18.a

C(\lambda)	Indice ECMCTD), XA (N)		
	40	80	116
0.1	0.9793	0.9809	0.9835
0.2	0.9518	0.9488	0.953
0.4	0.9313	0.9247	0.93

Tabla 4.2.18.b

	Indice ECMCTD), XB (N)		
	40	80	116
0.1	0.9795	0.9809	0.9836
0.2	0.9525	0.9489	0.9531
0.4	0.9326	0.9250	0.9302

Tabla 4.2.18.c

C(\lambda)	Indice ECMCTD), KF (N)		
	40	80	116
0.10	0.9804	0.9875	0.9901
0.40	0.9531	0.9659	0.9711
0.60	0.9339	0.9505	0.9575

Tabla 4.2.19.a

C(λ)	Indice ECMCS), XA (N)		
	40	80	116
0.1	0.9333	0.9061	0.9049
0.2	0.945	0.9171	0.9138
0.4	0.9465	0.9184	0.9149

Tabla 4.2.19.b

C(λ)	Indice ECMCS), XB (N)		
	40	80	116
0.1	0.9348	0.9064	0.9051
0.2	0.9473	0.9176	0.9141
0.4	0.9490	0.9191	0.9153

Tabla 4.2.19.c

C(λ)	Indice ECMCS), KF (N)		
	40	80	116
0.10	0.9133	0.8985	0.8937
0.40	0.9448	0.9258	0.9168
0.60	0.9496	0.9303	0.9207

De la estimación del componente tendencia (tabla 4.2.18) se observa que no existen diferencias significativas entre XA, XB y KF.

De la estimación del componente estacional (tabla 4.2.19) podemos decir que :

- entre XA y XB practicamente no existe diferencia,
- cuando  $\lambda = 0.1$  parecen ser XA y XB mejores.

En el ajuste estacional de series de tiempo, el componente más importante es la tendencia, por lo tanto, se tabuló el promedio y las varianzas de los coeficientes de determinación entre las tendencias observada y estimada ( $R^2(T, \hat{T})$ ).

Tabla 4.2.20.a

C(λ)	$R^2(T, \hat{T})$ , XA (N)		
	40	80	116
0.1	0.9323	0.958	0.9665
0.2	0.8974	0.9243	0.9314
0.4	0.8741	0.9015	0.9067

Tabla 4.2.20.b

C(λ)	$R^2(T, \hat{T})$ , XB (N)		
	40	80	116
0.1	0.9323	0.9581	0.9666
0.2	0.8987	0.9247	0.9316
0.4	0.8762	0.9020	0.9071

Tabla 4.2.20.c

$R^2(T, \hat{T}), KF$			
$C(\lambda)$	40	80	116
0.10	0.9175	0.9552	0.9687
0.40	0.8496	0.9108	0.9337
0.60	0.8122	0.8844	0.9116

Varianzas de los índices  $R^2$

Tabla 4.2.21.a

$Var(R^2(T, \hat{T}), XA$			
$C(\lambda)$	40	80	116
0.1	0.0034	0.0013	0.0006
0.2	0.0073	0.0045	0.0035
0.4	0.0103	0.0075	0.0067

Tabla 4.2.21.b

$Var(R^2(T, \hat{T}), XB$			
$C(\lambda)$	40	80	116
0.1	0.0036	0.0013	0.0006
0.2	0.0074	0.0045	0.0035
0.4	0.0106	0.0074	0.0067

Tabla 4.2.21.c

$Var(R^2(T, \hat{T}), KF$			
$C(\lambda)$	40	80	116
0.10	0.0048	0.0016	0.0007
0.40	0.0159	0.0064	0.0033
0.60	0.0237	0.0107	0.0060

De las tablas se observa que éste es un índice más sensible, y con este índice se tienen las conclusiones siguientes :

a) cuando el tamaño de muestra es 40, XA y XB son mejores que KF, pero al aumentar el tamaño de la muestra, la diferencia entre los métodos se hace mínima,

b) es notorio el mejoramiento de la estimación de la tendencia en KF al aumentar el tamaño de la muestra,

c) entre XA y XB practicamente no existe diferencia.

Para tener otra medida de comparación entre los métodos, se definió el índice TOTAL

$$TOTAL = \sum_{\text{datos}} \frac{(y - \hat{y})^2}{\sigma_I^2} + \frac{(\hat{S}_t - \hat{S}_t)^2}{\sigma_T^2} + \frac{(\hat{T}_t - \hat{T}_t)^2}{\sigma_S^2} \quad (4.5)$$

que trata de medir el cumplimiento de las hipótesis con las cuales fueron estimados los componentes, y en las tablas siguientes, se

muestra los promedios del índice TOTAL.

Tabla 4.2.22.a

	Indice TOTAL, XA		
(λ)	40	80	(N)
0.1	16.0943	16.0358	15.493
0.2	25.9678	24.5785	24.0042
0.4	33.3278	31.1679	30.5753

Tabla 4.2.22.b

	Indice TOTAL, XB		
(λ)	40	80	(N)
0.1	14.5512	15.6606	15.2590
0.2	22.4489	23.6438	23.3278
0.4	28.3929	29.7440	29.6124

Tabla 4.2.22.c

	Indice TOTAL, KF		
(λ)	40	80	(N)
0.10	7.9499	6.9708	6.2923
0.40	4.9859	4.2752	3.8718
0.60	4.6422	3.8850	3.5022

De éstas tablas se concluye que K es superior a XA y XB en éste aspecto, pero indudablemente, las tablas de mayor interés son las tablas 4.2.20.

### 4.3 INCLUSION DE UNA VARIABLE AJUSTADA ESTACIONALMENTE EN UN MODELO DE REGRESION.

La pregunta de cuando es conveniente ajustar estacionalmente una serie de tiempo para incluirla en un modelo de regresión, surge de estudiar modelos con errores de observación en las variables.

Zellner (1971) estudia el modelo

$$y_t = \eta_t + u_t, \quad x_t = \xi_t + w_t, \quad (4.6)$$

$$\eta_t = \alpha + \beta \xi_t, \quad t = 1, \dots, n,$$

donde  $\langle u_t \rangle$  y  $\langle w_t \rangle$  son v.a.i.i.d.  $N(0, \sigma_u^2)$  y  $N(0, \sigma_w^2)$  respectivamente, e independientes entre sí, los parámetros desconocidos son  $\alpha, \beta, \langle \xi_t \rangle$  y  $\langle \eta_t \rangle$ , y se observa a  $\langle x_t \rangle$  y  $\langle y_t \rangle$ . Cuando el cociente  $\phi = \sigma_u^2 / \sigma_w^2$  es conocido, se pueden obtener estimadores máximo-verosímiles para  $(\alpha, \beta, \langle \xi_t \rangle, \sigma_w^2)$ .

Si se sabe que  $\langle u_t \rangle$  o  $\langle w_t \rangle$  son efectos estacionales modelables a través de regresión. Thomas y Wallis(1971), Jorgenson (1971) y Lovell(1966) han estudiado el problema, en particular, Lovell muestra que la estimación de  $\beta$  por OLS es la misma para los modelos (4.7) y (4.8)

$$Y = X \beta + D \alpha + \varepsilon \quad (4.7)$$

$$Y_a = X_a \beta + \varepsilon \quad (4.8)$$

donde  $X_a = AX \quad Y_a = AY \quad (4.9.a)$

$$A = I - D(D'D)^{-1}D' \quad (4.9.b)$$

X es la matriz de variables explicativas, D es la matriz de variables indicadoras para modelar un efecto estacional y  $\varepsilon$  es  $N(0, \sigma^2 I)$ .

En el modelo (4.8), es equivalente estimar a  $\beta$  con las series ajustadas  $(X_a, Y_a)$  y las series originales  $(X, Y)$ , si definimos a los residuales del modelo  $E\langle Y \rangle = D\alpha$ , como la serie ajustada estacionalmente para Y.

Pero modelar a un componente estacional con regresión no es la manera más adecuada, por ésta razón, Wallis(1974) y Sims(1974) estudian el modelo (4.6) cuando los ruidos no son v.a.i.d.d., es decir,

$$y_t = x_t * \beta(L) + e_t, \quad (4.10)$$

$$x'_t = x_t + w_t, \quad y'_t = y_t + u_t, \quad (4.11)$$

$$u_t = C(L) w_t + \varepsilon_t \quad (4.12)$$

$$\text{donde } \beta(L) = \sum_{-\infty}^{\infty} \beta_k L^k \quad C(L) = \sum_{-\infty}^{\infty} C_k L^k \quad (4.13)$$

$$\beta^*(\lambda) = \sum_{-\infty}^{\infty} \beta_k e^{-i\lambda k} \quad C^*(\lambda) = \sum_{-\infty}^{\infty} C_k e^{-i\lambda k} \quad (4.14)$$

observamos a  $\{x_t\}$  y  $\{y_t\}$  contaminadas por los ruidos  $\{w_t\}$  y  $\{v_t\}$  respectivamente.

Si usamos  $\{y'_t\}$  y  $\{x'_t\}$  para estimar a  $\beta$ , el estimador asintótico de la transformada de Fourier de  $\beta(L)$  es

$$b^*(\lambda) = \beta^*(\lambda)q(\lambda) + C^*(\lambda)(1 - q(\lambda)), \quad (4.15)$$

$$\text{donde } q(\lambda) = \frac{f_x(\lambda)}{f_x(\lambda) + f_v(\lambda)}, \quad (4.16)$$

$f_x$  y  $f_v$  son las densidades espectrales de  $x_t$  y  $w_t$ , por lo tanto, en el dominio de las frecuencias  $b^*(\lambda)$  es un promedio de  $\beta^*(\lambda)$  y  $C^*(\lambda)$ .

Cuando estimamos a  $\beta(\lambda)$  con las series  $A_x(L)x'$ ,  $A_y(L)y'$ , y  $C(L) = 0$ , el estimador asintótico de  $\beta^*(\lambda)$  es (Sims,1974),

$$b^*(\lambda) = \frac{A_y(\lambda)}{A_x(\lambda)} q(\lambda) \beta^*(\lambda), \quad (4.17)$$

escogiendo a  $A_y(\lambda) = 1$  y  $A_x(\lambda) = q(\lambda)$ , el estimador asintótico de  $\beta^*(\lambda)$  es insesgado, pero  $q(L)$  es el filtro óptimo para la estimación de  $\{x_t\}$  a partir de  $\{x'_t\}$ , entonces,  $q(\lambda)$  depende del modelo seleccionado para los componentes de la serie  $x'$ , y en terminos operativos, se ajusta estacionalmente a  $x'$  y con la



serie ajustada se estima  $\beta$ .

Aplicaremos el procedimiento descrito en el párrafo anterior a un caso particular. Sea el modelo de regresión donde observamos con error a las variables explicativas

$$y_t = T_t \beta + \sigma \varepsilon_t \quad t = 1, \dots, n. \quad (4.18.a)$$

$$x_t = S_t + T_t + I_t \quad t = 1, \dots, n. \quad (4.18.b)$$

donde

$$y_t \in \mathbb{R}^1, \quad x_t = [x_{t1}, x_{t2}, \dots, x_{tp}] \in \mathbb{R}^p, \quad \beta \in \mathbb{R}^p, \quad \varepsilon \sim \text{NID}(0, 1),$$

$y_t$  y  $x_t$  son variables observables, mientras que,  $T_t, S_t, I_t$  y  $\varepsilon_t$  son variables  $p$ -dimensionales no observables e independientes entre sí.

Si definimos las matrices  $M_T, \hat{M}_T, Y$  y  $e$  por

$$M_T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix}, \quad \hat{M}_T = \begin{bmatrix} \hat{T}_1 \\ \hat{T}_2 \\ \vdots \\ \hat{T}_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad e = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (4.19)$$

el estimador de  $\beta$  por mínimos cuadrados, usando  $\hat{M}_T$  en sustitución de  $M_T$  es

$$\beta^* = ( \hat{M}_T' \hat{M}_T )^{-1} \hat{M}_T' Y, \quad (4.20)$$

una expresión para  $\beta^*$  es

$$\beta^* = \beta + A^* \beta + B^* e, \quad (4.21)$$

donde  $A^* = ( \hat{M}_T' \hat{M}_T )^{-1} \hat{M}_T' ( M_T - \hat{M}_T )$ , y  $B^* = ( \hat{M}_T' \hat{M}_T )^{-1} \hat{M}_T' e$ . (4.22)

Cuando se estima  $\beta$  con  $M_T$ , la expresión para  $\hat{\beta}$  es

$$\hat{\beta} = ( M_T' M_T )^{-1} M_T' Y \quad (4.23)$$

y sustituyendo  $Y$  por  $( \hat{M}_T + M_T - \hat{M}_T ) \beta + e$ , se tiene,

$$\hat{\beta} = \beta + A \beta + B e \quad (4.24)$$

donde  $A = 0$ , y  $B = ( M_T' M_T )^{-1} M_T' e$ , (4.25)

Un criterio para calificar a  $\beta^*$  es su sesgo, que tiene la expresión siguiente

$$E(\beta^* - \beta) = E(A^*)\beta + E(B^*)\sigma \quad (4.26)$$

donde se observa :

- 1)  $\beta^*$  es insesgado si  $\hat{M}_T = M_T$  y  $e_t$  no está correlacionado con  $S_t$  y  $T_t$ ,
- 2)  $\beta^*$  es insesgado si  $\hat{M}_T$  y  $\hat{M}_T - M_T$  son no correlacionados y  $e_t$  no está correlacionado con  $S_t$ ,  $T_t$  y  $I_t$ ,
- 3) el sesgo de  $\beta^*$  es una función lineal de  $\beta$ .

Una medida de la variabilidad de  $\beta^*$  es la cantidad

$$(\beta^* - \beta)'(\beta^* - \beta) \quad (4.27)$$

para la que se tiene la expresión

$$(\beta^* - \beta)'(\beta^* - \beta) = \beta' A^* A^* \beta + 2\beta' A^* B^* \sigma + \sigma^2 B^* B^* \quad (4.28)$$

con valor esperado

$$\beta' E(A^* A^*) \beta + 2\beta' E(A^* B^*) \sigma + \sigma^2 E(B^* B^*) \quad (4.29)$$

Finalmente, la expresión para la suma de los residuales al cuadrado es

$$(Y - \hat{Y})(Y - \hat{Y}) = \sigma^2 \langle T3^* \rangle + \beta' \langle M3^* \rangle \beta + 2\beta' \langle V3^* \rangle \sigma \quad (4.30)$$

donde

$$T3^* = e'e - e' \hat{M}_T B, \quad (4.31)$$

$$M3^* = (M_T - \hat{M}_T)'(M_T - \hat{M}_T) - (M_T - \hat{M}_T)' \hat{M}_T A^*, \quad (4.32)$$

$$V3^* = (M_T - \hat{M}_T)' e - (M_T - \hat{M}_T)' \hat{M}_T B^*. \quad (4.33)$$

Expresiones similares a (4.26)-(4.33), se obtienen para  $\hat{\beta}$  si se sustituye  $B$  por  $B^*$  y se tiene en cuenta que  $M_T = \hat{M}_T$ , en este caso, las formas cuadráticas (4.28) y (4.30) sólo son función de  $\sigma^2$ .

Las expresiones obtenidas no son fáciles de estudiar analíticamente cuando el método usado para estimar a la tendencia contiene elementos no lineales, por lo que, se estudiará a

través de simulaciones el comportamiento de los coeficientes que definen a las expresiones (4.26), (4.28) y (4.30).

Para aprovechar el trabajo de la sección 4.2, trabajaremos un modelo particular de (4.18),

$$y_k = \beta_0 + \beta_1 t_k + \varepsilon_k, \quad (4.34.a)$$

$$x_k = s_k + t_k + i_k, \quad k = 1, \dots, n. \quad (4.34.b)$$

donde  $x_k$  es la serie definida en (4.1),  $s_k, t_k$  e  $i_k$  son los componentes de  $x_k$ , y  $\varepsilon_k$  es un ruido blanco no correlacionado con los componentes de  $x_k$ .

El modelo (4.34) tiene dos variables explicativas y sólo  $t_k$  es observada con error, este hecho ocasiona que se den simplificaciones para los coeficientes de interés, y las expresiones (4.36)-(4.42), particulares del modelo (4.34) lo muestran.

$$\text{Sean} \quad \hat{T}' = [ \hat{t}_1, \hat{t}_2, \dots, \hat{t}_n ] \quad (4.35.a)$$

$$T' = [ t_1, t_2, \dots, t_n ] \quad (4.35.b)$$

$$1' = [ 1_1, 1_2, \dots, 1_n ] \quad (4.35.c)$$

entonces

$$M_T = [ 1 \ T ] \quad \hat{M}_T = [ 1 \ \hat{T} ] \quad (4.36)$$

$$M_T - \hat{M}_T = [ 0, \ T - \hat{T} ] \quad (4.37)$$

$$\hat{M}_T' (M_T - \hat{M}_T) = \begin{bmatrix} 0 & 0 \\ 0 & \hat{T}'(T - \hat{T}) \end{bmatrix} \quad (4.38)$$

$$(M_T - \hat{M}_T)' (M_T - \hat{M}_T) = \begin{bmatrix} 0 & 0 \\ 0 & (T - \hat{T})'(T - \hat{T}) \end{bmatrix} \quad (4.39)$$

$$(M_T - \hat{M}_T)' \hat{M}_T = \begin{bmatrix} 0 & 0 \\ (T - \hat{T})' 1 & (T - \hat{T})' \hat{T} \end{bmatrix} \quad (4.40)$$

$$C M_T - \hat{M}_T' e = \begin{bmatrix} 0 \\ (T-T)' e \end{bmatrix} \quad (4.41)$$

y las simplificaciones que se dan son,

$$A^* = \begin{bmatrix} 0 & a1^* \\ 0 & a2^* \end{bmatrix} \quad A^* \cdot A^* = \begin{bmatrix} 0 & 0 \\ 0 & aa^* \end{bmatrix} \quad (4.42)$$

$$A^* \cdot B^* = \begin{bmatrix} 0 \\ ab^* \end{bmatrix} \quad M3^* = \begin{bmatrix} 0 & 0 \\ 0 & m3^* \end{bmatrix}$$

$$V3^* = \begin{bmatrix} 0 \\ v3^* \end{bmatrix}$$

Para el modelo (4.34), las formas cuadráticas (4.28) y (4.30) son función de los parámetros  $\sigma^2$  y  $\beta_1$ , y no dependen de  $\beta_0$ , coeficiente de la variable que observamos sin error.

En las tablas de esta sección, las columnas señalan los tamaños de muestra (N) y los renglones indican el cociente  $\lambda = \text{Var}(S_T) / \text{Var}(Y_T)$ .

En primer lugar, se dan los promedios de los coeficientes del vector  $a^* = [a1^*, a2^*]$  (expresión 4.26), junto con sus varianzas.

Elementos del vector  $a^*$ .

Coficiente  $a1^*$ .

Tabla 4.3.1.a

	$a1^* \quad XA$		(N)
( $\lambda$ )	40	80	118
0.1	-0.0214	0.0184	0.0419
0.2	-0.0388	0.0459	0.0872
0.4	-0.0504	0.0578	0.1112

Tabla 4.3.1.b

	$a1^* \quad XB$		(N)
( $\lambda$ )	40	80	118
0.1	-0.0250	0.0195	0.0422
0.2	-0.0468	0.0461	0.0889
0.4	-0.0609	0.0577	0.1143

Tabla 4.3.1.c

	a1*	KF	CN)
Cλ)	40	80	116
0.10	0.0082	0.0743	0.0111
0.40	-0.0029	0.1384	0.0182
0.60	-0.0123	0.1695	0.0214

Varianzas del coeficiente a1\*

Tabla 4.3.2.a

	v_a1*	XA	CN)
Cλ)	40	80	116
0.1	0.086	0.1418	0.158
0.2	0.3027	0.5126	0.597
0.4	0.4549	0.7767	0.912

Tabla 4.3.2.b

	v_a1*	XB	CN)
Cλ)	40	80	116
0.1	0.0867	0.1406	0.1575
0.2	0.3006	0.5129	0.5961
0.4	0.4451	0.7790	0.9100

Tabla 4.3.2.c

	v_a1*	KF	CN)
Cλ)	40	80	116
0.10	0.1562	0.1637	0.2113
0.40	0.4299	0.5589	0.7935
0.60	0.6287	0.8394	1.2187

Coeficiente a2\*

Tabla 4.3.3.a

	a2*	XA	CN)
Cλ)	40	80	116
0.1	-0.0075	-0.0075	-0.0022
0.2	-0.0437	-0.0369	-0.0292
0.4	-0.0664	-0.0567	-0.0498

Tabla 4.3.3.b

	a2*	XB	CN)
Cλ)	40	80	116
0.1	-0.0059	-0.0075	-0.0021
0.2	-0.0398	-0.0369	-0.0290
0.4	-0.0606	-0.0560	-0.0494

Tabla 4.3.3.c

	a2*	KF	CN)
Cλ)	40	80	116
0.10	-0.0356	-0.0157	-0.0082
0.40	-0.0934	-0.0485	-0.0309
0.60	-0.1303	-0.0703	-0.0474

Varianza del coeficiente  $a_2^*$ .

Tabla 4.3.4.a

	$v_{a_2^*}$	XA	CND
( $\lambda$ )	40	80	116
0.1	0.0033	0.0032	0.0027
0.2	0.0118	0.0114	0.0093
0.4	0.0173	0.0166	0.0135

Tabla 4.3.4.b

	$v_{a_2^*}$	XB	CND
( $\lambda$ )	40	80	116
0.1	0.0033	0.0031	0.0027
0.2	0.0119	0.0115	0.0092
0.4	0.0175	0.0168	0.0134

Tabla 4.3.4.c

	$v_{a_2^*}$	KF	CND
( $\lambda$ )	40	80	116
0.10	0.0051	0.0030	0.0026
0.40	0.0150	0.0110	0.0096
0.60	0.0222	0.0168	0.0141

Se puede comentar que las varianzas de los coeficientes  $a_2^*$  disminuyen al aumentar el tamaño de muestra, y las varianzas de  $a_1^*$  aumentaron con el tamaño de la muestra.

Los coeficientes del vector  $b^* = [b_1^*, b_2^*]'$  que aparecen en la expresión del sesgo de  $\beta^*$ .

Coeficiente  $b_1^*$

Tabla 4.3.5.a

	$b_1^*$	XA	CND
( $\lambda$ )	40	80	116
0.1	-0.0039	0.0113	0.0119
0.2	-0.0006	0.0135	0.0143
0.4	0.0006	0.015	0.0153

Tabla 4.3.5.b

	$b_1^*$	XB	CND
( $\lambda$ )	40	80	116
0.1	-0.0056	0.0113	0.0119
0.2	-0.0018	0.0134	0.0143
0.4	0.0002	0.0151	0.0153

Tabla 4.3.5.b

	$b_1^*$	KF	CND
( $\lambda$ )	40	80	116
0.10	0.0285	0.0252	-0.0060
0.40	0.0282	0.0259	-0.0050
0.60	0.0280	0.0250	-0.0045

Coeficiente  $b_2^*$

Tabla 4.3.6.a

C $\lambda$	$b_2^*$ XA		CND
	40	80	
0.1	0.0073	-0.0054	0.0007
0.2	0.0097	-0.0043	0.0014
0.4	0.0106	-0.0039	0.0018

Tabla 4.3.6.b

C $\lambda$	$b_2^*$ XB		CND
	40	80	
0.1	0.0073	-0.0055	0.0006
0.2	0.0104	-0.0043	0.0013
0.4	0.0114	-0.0040	0.0017

Tabla 4.3.6.c

C $\lambda$	$b_2^*$ KF		CND
	40	80	
0.10	-0.0007	-0.0037	-0.0014
0.40	0.0027	-0.0047	-0.0012
0.60	0.0038	-0.0051	-0.0012

Varianzas de los elementos del vector  $b^*$ .

Varianzas del coeficiente  $b_1^*$

Tabla 4.3.7.a

C $\lambda$	$v_{b_1}^*$ XA		CND
	40	80	
0.1	0.088	0.04	0.0246
0.2	0.1162	0.0497	0.028
0.4	0.139	0.0568	0.0308

Tabla 4.3.7.b

C $\lambda$	$v_{b_1}^*$ XB		CND
	40	80	
0.1	0.0875	0.0399	0.0245
0.2	0.1184	0.0496	0.0280
0.4	0.1424	0.0569	0.0308

Tabla 4.3.7.c

C $\lambda$	$v_{b_1}^*$ KF		CND
	40	80	
0.10	0.0775	0.0381	0.0258
0.40	0.0739	0.0379	0.0257
0.60	0.0718	0.0379	0.0258

Varianzas del coeficiente  $b_2^*$ .

Tabla 4.3.8.a

	$v_{b_2}^*$	XA	(N)
( $\lambda$ )	40	80	116
0.1	0.0007	0.0025	0.0013
0.2	0.0077	0.0026	0.0014
0.4	0.0085	0.0027	0.0014

Tabla 4.3.8.b

	$v_{b_2}^*$	XB	(N)
( $\lambda$ )	40	80	116
0.1	0.0070	0.0025	0.0013
0.2	0.0078	0.0026	0.0014
0.4	0.0087	0.0027	0.0014

Tabla 4.3.8.c

	$v_{b_2}^*$	KF	(N)
( $\lambda$ )	40	80	116
0.10	0.0067	0.0018	0.0007
0.40	0.0058	0.0018	0.0007
0.60	0.0053	0.0017	0.0007

Las varianzas de los coeficientes  $b_1^*$  y  $b_2^*$  disminuyen al aumentar el tamaño de la muestra.

Al calcular el cociente  $|E(T)| / (\text{Var}(T))^{1/2}$  para los estimadores de  $a_1^*$ ,  $a_2^*$ ,  $b_1^*$  y  $b_2^*$  resultó ser menor que 1. Se esperaba que  $b_1^*$  y  $b_2^*$  tuvieran valor esperado de cero porque  $\epsilon_k$  no está correlacionado con  $s_k$ ,  $t_k = 1/k$ .

En seguida, aparecen los promedios de los coeficientes de la norma al cuadrado del vector de sesgos para  $\beta_0$  y  $\beta_1$  (expresión 4.28).

El escalar  $aa^*$ , coeficiente de  $\beta^2$ .

Tabla 4.3.9.a

	$aa^*$	XA	(N)
( $\lambda$ )	40	80	116
0.1	0.0899	0.1454	0.1625
0.2	0.3179	0.5275	0.6148
0.4	0.4791	0.7999	0.9404

Tabla 4.3.9.b

	$aa^*$	XB	(N)
( $\lambda$ )	40	80	116
0.1	0.0906	0.1442	0.1620
0.2	0.3163	0.5278	0.6141
0.4	0.4699	0.8023	0.9389



Tabla 4.3.9.c

	aa* KF		CND
Cλ)	40	80	116
0.10	0.1626	0.1726	0.2141
0.40	0.4536	0.5915	0.8044
0.60	0.6680	0.8899	1.2355

El escalar bb\*, coeficiente de  $\sigma^2$ .

Tabla 4.3.10.a

	bb* XA		CND
Cλ)	40	80	116
0.1	0.0951	0.0426	0.026
0.2	0.1239	0.0525	0.0296
0.4	0.1476	0.0597	0.0324

Tabla 4.3.10.a

	bb*, XB		CND
Cλ)	40	80	116
0.1	0.0946	0.0426	0.0259
0.2	0.1264	0.0524	0.0295
0.4	0.1513	0.0599	0.0325

Tabla 4.3.10.a

	bb* KF		CND
Cλ)	40	80	116
0.10	0.0850	0.0406	0.0265
0.40	0.0805	0.0403	0.0265
0.60	0.0779	0.0403	0.0265

El valor ab\*, coeficiente de  $2\beta\sigma$ .

Tabla 4.3.11.a

	ab* XA		CND
Cλ)	40	80	116
0.1	0.0073	0.0053	-0.003
0.2	0.0092	0.0141	-0.0031
0.4	0.0088	0.0197	-0.0012

Tabla 4.3.11.b

	ab* XB		CND
Cλ)	40	80	116
0.1	0.0061	0.0054	-0.0029
0.2	0.0089	0.0143	-0.0030
0.4	0.0107	0.0200	-0.0011

Tabla 4.3.11.c

	ab* KF		CND
Cλ)	40	80	116
0.10	-0.0225	0.0033	0.0051
0.40	-0.0318	0.0062	0.0087
0.60	-0.0364	0.0076	0.0090

Los coeficientes de la expresión que mide la suma de los residuos al cuadrado (expresión 4.30).

El coeficiente de  $\sigma^2$ .

Tabla 4.3.12.a

C $\lambda$	t3* XA		CN
	40	80	
0.1	40.0761	81.1945	116.745
0.2	45.8788	88.6394	124.915
0.4	50.2036	94.223	131.02

Tabla 4.3.12.b

C $\lambda$	t3* XB		CN
	40	80	
0.1	40.0701	81.1938	116.7470
0.2	45.8683	88.6413	124.9150
0.4	50.1864	94.2206	131.0180

Tabla 4.3.12.c

C $\lambda$	t3* KF		CN
	40	80	
0.10	37.6928	78.5030	114.8194
0.40	37.7377	78.5104	114.7951
0.60	37.7618	78.5141	114.7798

El coeficiente de  $\beta_1^2$ .

Tabla 4.3.13.a

C $\lambda$	m3* XA		CN
	40	80	
0.1	9.0726	20.4753	32.3071
0.2	14.0337	37.0246	66.2108
0.4	17.4875	48.7836	90.7193

Tabla 4.3.13.b

C $\lambda$	m3* XB		CN
	40	80	
0.1	9.0066	20.4107	32.2342
0.2	13.6539	36.8161	65.9932
0.4	16.9279	48.4497	90.3338

Tabla 4.3.13.c

C $\lambda$	m3* KF		CN
	40	80	
0.10	12.6686	26.4715	39.7355
0.40	23.6542	53.4299	84.4728
0.60	30.2222	70.1107	113.4614

El coeficiente de  $2\beta_1\sigma$ .

Tabla 4.3.14.a

C(\lambda)	v3* XA		CN)
	40	80	116
0.1	0.0246	-0.2558	-0.3248
0.2	-0.0963	-0.537	-0.505
0.4	-0.199	-0.624	-0.6036

Tabla 4.3.14.c

C(\lambda)	v3* XB		CN)
	40	80	116
0.1	0.0225	-0.2315	-0.2821
0.2	-0.0623	-0.5231	-0.4308
0.4	-0.1626	-0.5809	-0.5530

Tabla 4.3.14.c

C(\lambda)	v3*, KF		CN)
	40	80	116
0.10	-0.1103	0.4294	-0.5970
0.40	-0.3678	0.6938	-0.6912
0.60	-0.4628	0.7599	-0.7400

Para medir la eficiencia del estimador  $\beta^*$  en relacion con el estimador de minimos cuadrados ( $\hat{\beta}$ ), se evaluaron las expresiones

$$\frac{E(\beta - \beta^*)(\beta - \hat{\beta})}{E(\beta - \hat{\beta})(\beta - \hat{\beta})} = \frac{(bb^*)\sigma^2 + 2\beta\sigma(ab^*) + Caa^*)\beta^2}{(bb)\sigma^2} \quad (4.43)$$

$$\frac{E(y - y(\beta^*))^2}{E(y - y(\hat{\beta}))^2} = \frac{(t3^*)\sigma^2 + 2\beta\sigma(v3^*) + (m3^*)\beta^2}{(t3)\sigma^2} \quad (4.44)$$

donde E significa, un promedio sobre todas las simulaciones realizadas, y  $y(\hat{\beta})$  es el valor estimado de y por  $\hat{\beta}$ .

Las expresiones (4.43) y (4.44) son polinomios de segundo orden como función del cociente  $\theta = \beta/\sigma$ , entonces, (4.43) y (4.44) son funciones no acotadas de  $\theta$ , y cuando mas grande sea  $\theta$ , más ineficiente es  $\beta^*$ , faltandonos por saber, la magnitud de la ineficiencia, reescribiendo (4.43) y (4.44) obtenemos (4.45) y (4.46)

$$(h1) + (h2)\theta + (h3)\theta^2 \quad (4.45)$$

$$(k1) + (k2)\theta + (k3)\theta^2 \quad (4.46)$$

y los coeficientes h1, h2, h3, k1, k2 y k3 se dan en las siguientes tablas.

Coefficiente h1 (expresión 4.45).

Tabla 4.3.15.a  
h1, XA

C(λ)	40	80	116	CND
0.1	1.0007	0.9998	0.9998	
0.2	1.0003	1.0002	0.9997	
0.4	0.9996	1.0004	0.9996	

Tabla 4.3.15.b  
h1, XB

C(λ)	40	80	116	CND
0.1	1.0008	0.9998	0.9998	
0.2	1.0001	1.0002	0.9997	
0.4	0.9993	1.0004	0.9996	

Tabla 4.3.15.c

h1, KF

C(λ)	40	80	116	CND
0.10	1.0017	1.0001	0.9999	
0.40	1.0029	1.0002	0.9997	
0.80	1.0036	1.0002	0.9995	

Coefficiente h2 (expresión 4.45).

Tabla 4.3.16.a

h2, XA

C(λ)	40	80	116	CND
0.1	0.0012	-0.0063	-0.0056	
0.2	-0.0042	-0.0121	-0.0081	
0.4	-0.0079	-0.0132	-0.0092	

Tabla 4.3.16.b

h2, XB

C(λ)	40	80	116	CND
0.1	0.0011	-0.0057	-0.0048	
0.2	-0.0027	-0.0118	-0.0069	
0.4	-0.0065	-0.0123	-0.0084	

Tabla 4.3.17.a

h2, KF

C(λ)	40	80	116	CND
0.10	-0.0059	0.0109	-0.0104	
0.40	-0.0195	0.0177	-0.0120	
0.80	-0.0246	0.0194	-0.0129	

Coefficiente h3 (expresión 4.46).

Tabla 4.3.18.a

h3, XA

C(λ)	40	80	116	CND
0.1	0.2265	0.2521	0.2787	
0.2	0.306	0.4178	0.5299	
0.4	0.3482	0.5179	0.6921	

Tabla 4.3.18.b

h3, XB

C(λ)	40	80	116	CND
0.1	0.2249	0.2513	0.2780	
0.2	0.2977	0.4154	0.5281	
0.4	0.3371	0.5144	0.6892	

Tabla 4.3.18.c

h3, KF		CND	
Cλ	40	80	116
0.10	0.3367	0.3372	0.3480
0.40	0.6287	0.6807	0.7356
0.60	0.8032	0.8932	0.9881

Coeficiente k1 (Expresión 4.46).

Tabla 4.3.19.a

k1, XA		CND	
Cλ	40	80	116
0.1	1.0055	1.0304	1.0039
0.2	0.9906	0.9748	0.9916
0.4	0.9953	0.9442	0.9801

Tabla 4.3.19.b

k1, XB		CND	
Cλ	40	80	116
0.1	1.0003	1.0293	1.0003
0.2	1.0099	0.9727	0.9910
0.4	1.0199	0.9471	0.9813

Tabla 4.3.19.c

k1, KF		CND	
Cλ	40	80	116
0.10	0.9551	0.9988	0.9664
0.40	0.9045	0.9924	0.9661
0.60	0.8752	0.9903	0.9691

Coeficiente k2 (Expresión 4.46).

Tabla 4.3.20.a

k2, XA		CND	
Cλ	40	80	116
0.1	0.1541	0.2573	-0.2301
0.2	0.1466	0.5257	-0.2075
0.4	0.1192	0.6241	-0.0744

Tabla 4.3.20.b

k2, XB		CND	
Cλ	40	80	116
0.1	0.1293	0.2611	-0.2272
0.2	0.1424	0.5325	-0.1992
0.4	0.1441	0.6323	-0.0659

Tabla 4.3.20.c

k2, KF		CND	
Cλ	40	80	116
0.10	-0.5065	0.1632	0.3721
0.40	-0.7145	0.3060	0.6321
0.60	-0.8177	0.3744	0.8600

Coefficiente k3 (expresión 4.48).

Tabla 4.3.21.a

	k3, XA		CND
(λ)	40	80	116
0.1	0.9499	3.5151	6.2632
0.2	2.5405	9.7974	20.618
0.4	3.2303	12.6488	28.4107

Tabla 4.3.21.b

	k3, XB		CND
(λ)	40	80	116
0.1	0.9578	3.4860	6.2464
0.2	2.5277	9.8040	20.5942
0.4	3.1686	12.6866	28.3655

Tabla 4.3.21.c

	k3 KF		CND
(λ)	40	80	116
0.10	1.8275	4.2452	7.8187
0.40	5.0982	14.5507	29.3694
0.60	7.5086	21.8914	45.1105

El incremento de los coeficientes (k3,h3) al aumentar el tamaño de la muestra, indica que, por cada observación que añadimos se pierde eficiencia respecto a hacer OLS con la tendencia verdadera.

La pérdida de eficiencia es mucho mayor en la estimación de  $y(\beta)$ .

Después de hacer los histogramas de las variables  $a_1^*$ ,  $a_2^*$ ,  $b_1^*$  y  $b_2^*$ , se observa que muchas de ellas parecen ser simétricas (Gráficas 4.1, 4.2, 4.3 y 4.4), y por cada variable se calculó la estadística Anderson-Darling (D'Agostino y Stephens, 1986), para probar la hipótesis

H : Normalidad con media y varianza desconocida.

Lo cuantiles de la estadística de prueba son :

α	.5	.75	.85	.90	.95	.975
	.341	.470	.561	.631	.752	.873

En las tablas siguientes son señalados con \*, las estadísticas con valores de rechazo al 5%.

Tabla 4.3.22.a

	Anderson a1 XA		(CN)
(λ)	40	80	116
0.1	0.6759	0.1446	0.3126
0.2	0.7078	0.2479	0.3673
0.4	0.6567	0.278	0.3862

Tabla 4.3.22.b

	Anderson a1 XB		(CN)
(λ)	40	80	116
0.1	0.5390	0.1634	0.3081
0.2	0.5678	0.2272	0.3697
0.4	0.6415	0.2827	0.3780

Tabla 4.3.22.c

	Anderson a1 KF		(CN)
(λ)	40	80	116
0.10	0.8557*	0.5135	0.1928
0.40	0.4918	0.4734	0.1695
0.60	0.4442	0.4644	0.1273

Tabla 4.3.23.a

	Anderson a2, XA		(CN)
(λ)	40	80	116
0.1	0.284	0.5874	0.7021
0.2	0.2689	0.8625*	0.6121
0.4	0.2455	0.8352*	0.4897

Tabla 4.3.23.b

	Anderson a2, XB		(CN)
(λ)	40	80	116
0.1	0.5444	0.5802	0.6871
0.2	0.4550	0.8558*	0.5776
0.4	0.3840	0.8372*	0.4888

Tabla 4.3.23.c

	Anderson a2, KF		(CN)
(λ)	40	80	116
0.10	0.6339	0.3232	0.6559
0.40	0.3676	0.2000	0.8274*
0.60	0.5013	0.2221	0.7175

Tabla 4.3.24.a

	Anderson b1 XA		(CN)
(λ)	40	80	116
0.1	0.7302	1.5004*	0.9898*
0.2	1.2088*	2.0746*	0.9227*
0.4	1.5135*	2.3372*	0.9741*

Tabla 4.3.24.b

	Anderson b1 XB		(CN)
(λ)	40	80	116
0.1	0.6281	1.4631*	0.9642*
0.2	1.2585*	2.0322*	0.9186*
0.4	1.6961*	2.3226*	0.9867*

Tabla 4.3.24.c

	Anderson b1 KF		(CN)
(λ)	40	80	116
0.10	1.1117*	1.4396*	0.5229
0.40	0.9336*	1.3982*	0.4228
0.60	0.8194*	1.4983*	0.3751

Tabla 4.3.25.a

	Anderson b2 XA		CND
Cλ)	40	80	116
0.1	0.5749	1.4794*	1.0709*
0.2	0.6055	0.9745*	1.256*
0.4	0.5364	0.9403*	1.3593*

Tabla 4.3.5.b

	Anderson b2 XB		CND
Cλ)	40	80	116
0.1	0.5380	1.4595*	1.0808*
0.2	0.5659	0.9819*	1.2768*
0.4	0.5272	0.9670*	1.3518*

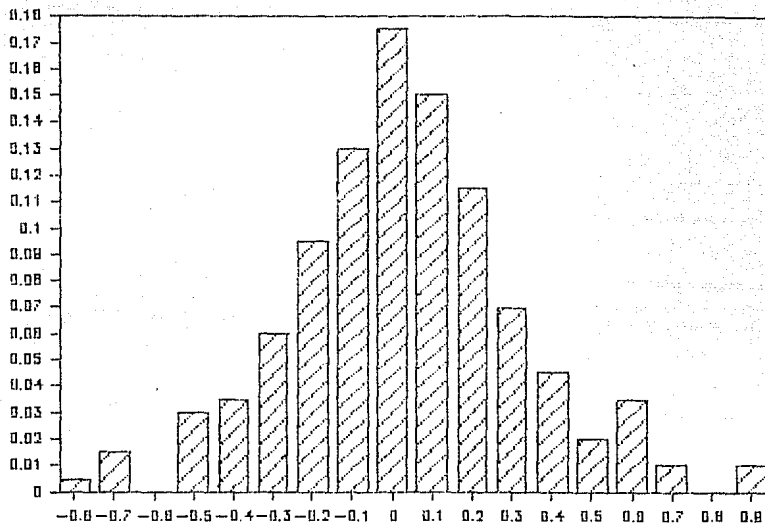
Tabla 4.3.25.c

	Anderson b2 KF		CND
Cλ)	40	80	116
0.10	1.0939*	0.9693*	1.1325*
0.40	0.9250*	1.1393*	1.2015*
0.60	0.7346*	1.2004*	1.2017*

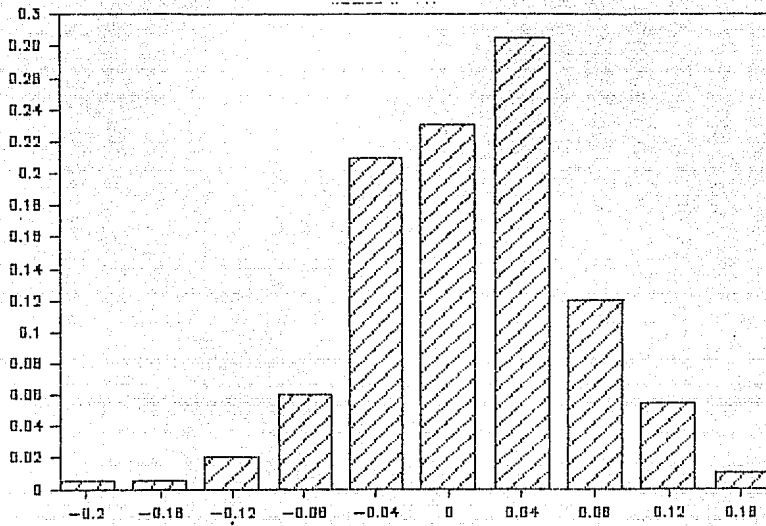
Al aumentar el tamaño de muestra, aumenta la evidencia para suponer la normalidad de  $a_1$ , en  $a_2$  el comportamiento es un poco extraño, pero si parecen ser normales  $a_1$  y  $a_2$ , mientras que  $b_1$  y  $b_2$  no parecen ser normales.



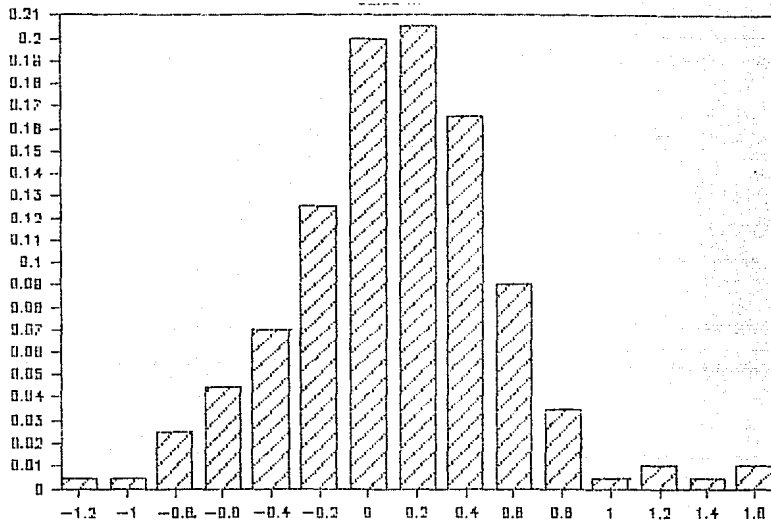
Gráfica 3.1  
Histograma de las Estimaciones del  
Parámetro  $a_1$   
Método X-11



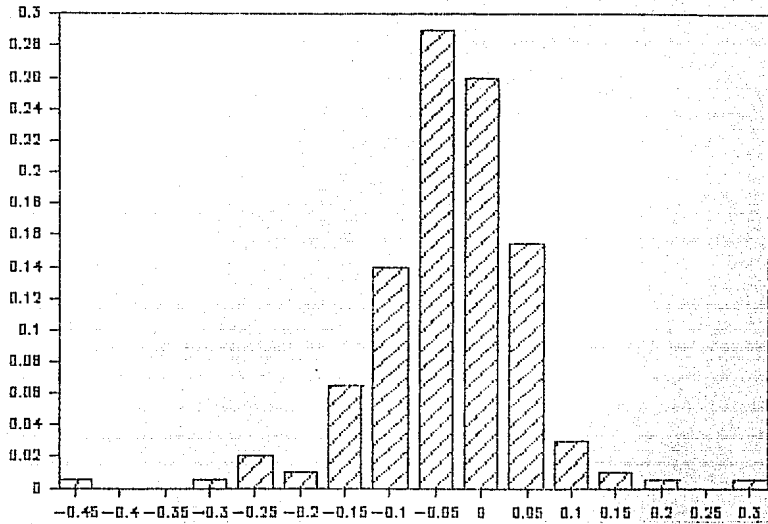
Gráfica 3.2  
Histograma de las Estimaciones del  
Parámetro  $a_2$   
Método X-11



Gráfica 3.3  
Histograma de las Estimaciones del  
Parámetro  $a_1$   
Método KF



Gráfica 3.4  
Histograma de las Estimaciones del  
Parámetro  $a_2$   
Método KF



## CONCLUSIONES.

Después de haber terminado éste trabajo, se han obtenido las conclusiones siguientes.

El filtro de Kalman es una herramienta teóricamente muy poderosa, y es fácil de implementar porque sus algoritmos son sencillos, además, bajo un mismo esquema permite estudiar a una gran variedad de modelos y simplifica su entendimiento.

El método X-11-ARIMA no tiene comparación con el filtro de Kalman en la parte teórica, es poco flexible, y no es fácil de implementar.

Debe reconocerse, que el método X-11-ARIMA es muy fácil de usar, es casi automático, y usa elementos de Estadística Robusta.

Buena parte del trabajo estadístico se orienta al estudio y extensión del filtro de Kalman, pero falta ganar experiencia en su aplicación, para que sea la herramienta de uso común en el modelado de series de tiempo y ajuste estacional, para ésto, se presenta como un obstáculo, que el método X-11-ARIMA y la metodología Box-Jenkins, han funcionado bien en un gran número de casos reales.

En el ajuste estacional de series de tiempo, es preferible el uso de los modelos estructurales, porque con éstos, se selecciona un modelo para la serie y se obtiene de manera simultanea la estimación del componente tendencia, además, la parte lineal del método X-11 puede ser reconstruida con un modelo estructural.

Aunque sólo se trabajó con series univariadas, teóricamente no existe dificultad para que el filtro de Kalman modele series multivariadas, siendo las principales limitaciones

computacionales, por lo tanto, el mejoramiento de los equipos y programas de cómputo, influirá decisivamente en la aplicación común de éstas técnicas.

Como resultados de éste trabajo, se implementaron los dos métodos en estudio, y se comprobó que, para algunos modelos de series de tiempo, la estimación del componente tendencia es muy similar con el método X-11-ARIMA y el uso del filtro de Kalman.

Las simulaciones realizadas han servido para aprender y no obtener resultados concluyentes, porque sólo se exploró un conjunto muy pequeño de modelos.

No ha sido posible mostrar todo el impacto del filtro de Kalman en la Estadística, por lo que, se mencionan tres desarrollos:

- a) intentos para desarrollar un filtro de Kalman robusto (Peña y Guttman, 1988),
- b) filtros no lineales para sistemas no lineales cuyo vector de estados evoluciona continuamente y las observaciones son realizadas en tiempos discretos (Daum, 1988),
- c) Poole, West y Harrison (1988) presentan varios modelos no lineales y/o no normales.

Finalmente, ha aparecido el caracter interdisciplinario de la ciencia, pues el filtro de Kalman fue desarrollado en la Ingeniería.

APENDICE A.

Programa X-11-ARIMA

El programa proporcionado por la Dra. E.B. Dagum, se instaló en la computadora VAX-11/750 del Departamento de Probabilidad y Estadística, del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la UNAM, y a continuación se describen los aspectos relevantes de la instalación.

1. El programa de aproximadamente 7000 líneas, fue dividido en los archivos : X11\_1, X11\_2, X11\_3, X11\_4 y X11\_5.

2. El programa leía los datos en tarjetas, ahora se pide de la terminal el nombre de un archivo con extensión ".mt0", que contenga los datos y opciones del programa, tal como aparecen en el manual (Dagum 1983).

3. En el programa se hace referencia a las rutinas :

```
DVZCHK J )
```

```
OVERFLC J )
```

encargadas de indicar el momento en que ha ocurrido una división por cero o una condición de "overflow", éstas rutinas no existen en el lenguaje FORTRAN de VAX, pero afortunadamente están escritas en un manual de D.E.C.(1985), fueron implantadas, y para su uso se requiere insertar,

a) en el encabezado,

```
include 'ssdef'  
external fixup_result  
integer*4 old_handler  
call lib$flt_under(1)  
old_handler = lib$establish(fixup_result)  
call ini_fz
```

b) y al finalizar la rutina,

```
call lib$establish(Old_handler)
```

por cada rutina que haga uso de las rutinas DVZCHK y OVERFL.

Las rutinas DVZCHK y OVERFL pueden ser usadas en otros programas FORTRAN y estan escritas en el archivo FZ1.FOR.

El valor del parámetro J al finalizar las rutinas DVZCHK y OVERFL se interpreta como sigue,

RUTINA	INTERPRETACION
OVERFL	J = 1 overflow J = 2 no overflow y no underflow J = 3 underflow
DVZCHK	J = 1 division por cero J = 2 no division por cero

4. Para almacenar resultados, en algunas rutinas del programa se definieron las estructuras

```
COMMON/memori/RMEM(5000)
COMMON/bande/I_LEE,I_GUA,I_TAB,I_GRA
COMMON/datos/I_LOC,NDAT
```

y su significado en el programa es

VARIABLE	SIGNIFICADO
I_LEE>0	leer datos del arreglo RMEM
I_GUA>0	escribir resultados al arreglo RMEM
I_TAB<0	no imprimir las tablas de salida del programa
I_GRA<0	no imprimir las graficas de salida
N_DAT	número de datos
I_LOC	localidad de RMEM donde almacenar o leer N_DAT datos



APENDICE B.

El programa K.

En éste apéndice se describe al programa K, escrito para implementar el ajuste estacional de series de tiempo, como lo propusieron Gersh y Kitagawa(1984).

El programa que se escribió, supone que la serie  $Z_t$  puede ser escrita como,

$$Z = D + T + S + W + I, \quad (B.1)$$

donde,

T es la tendencia y satisface  $(I-B)^d T_k = \xi_t(k)$ ,

S es el componente estacional y es generado por

$$(I + B + B^2 + \dots + B^{L-1}) S_k = \xi_s(k) \quad \circ$$

$$(I - B^L) S = \xi_s(k),$$

W es un ruido autoregresivo generado por

$$(I - \alpha_1 B - \dots - \alpha_p B^p) W_k = \xi_w(k),$$

I es una v.a. normal con media 0 y varianza R,

D es el componente determinista, y

$\xi(k) = [\xi_T(k), \xi_S(k), \xi_W(k)]$ , donde

$$E(\xi(k)\xi(l)') = \delta(k-l) * \text{diag}(\sigma_t^2, \sigma_s^2, \sigma_w^2) = Q.$$

El modelo (B.1) puede ser representado con un sistema lineal

$$X_{k+1} = \Phi X_k + G \xi_k \quad (B.2)$$

$$Y_k = C_k X_k + \eta_k \quad k = 0, \dots, N.$$

donde  $\xi_k$  y  $\eta_l$  son variables normales, con media cero y

segundos momentos dados por

$$E(\xi_k \xi_l') = \delta(k-l)Q \quad E(\eta_k \eta_l') = \delta(k-l)R \quad E(\xi_k \eta_l') = 0$$

El listado del programa no se presenta porque es muy grande.

Con el filtro de Kalman, podemos calcular la verosimilitud como función de los parámetros  $(\Phi, G, C_k, R, Q)$ , y en principio, maximizar la verosimilitud como función de todos los parámetros, pero esto no se hace por las razones siguientes :

a) la matriz  $G$  y los vectores  $C_k$  se fijan al elegir los modelos para los componentes,

b) la matriz  $\Phi$  es conocida, cuando no aparece el ruido autoregresivo en el modelo,

c) en el caso univariado es posible concentrar a la verosimilitud como función de  $R$ ,

d) por el supuesto de independencia de los componentes, la matriz  $Q$  es diagonal y se estiman sus elementos por máxima verosimilitud.

Inicialización del vector  $x(0;0)$ .

Si conocemos el valor esperado de alguna componente del vector, debe ser usado para inicializar la correspondiente componente. Cuando se desconoce el valor de  $x(0;0)$ , se inicializa con ceros, o se puede pronosticar con la serie invertida..

Para inicializar a la matriz  $P_{0/-1}$  existen dos opciones,

a) hacer  $P_{0/-1} = \lambda I$ , con  $\lambda$  muy grande ( $\lambda = 1e06$ ),

b) hacer  $P_0 = \lambda I$  y encontrar la solución estacionaria de la ecuación de Riccati, ésta opción duplica el tiempo de ejecución.

Optimización sobre  $R$ .

Ya se mostró en la sección 3.2 como es posible concentrar a la verosimilitud como función de  $R$ .

Optimización sobre  $Q$ .

Al concentrar la verosimilitud en  $R$ , ésta depende de los cocientes  $\sigma_c^2/R, \sigma_b^2/R, \sigma_v^2/R$ , que representan los cocientes de señal a ruido, y Gersh y Kitagawa(1984) proponen maximizar a la verosimilitud, buscand a los cocientes en el conjunto

$$\langle 2^j, j = 0, \dots, 8 \rangle,$$

con el supuesto, de que la verosimilitud es una función suave y mas o menos plana de los cocientes señal a ruido.

A través de simulaciones se observó, que fijando el valor de  $\sigma_t^2$ , la verosimilitud como función de  $\sigma_s^2$  es concava hacia arriba, y como resulta costoso evaluar la verosimilitud en cada uno de los puntos de la malla, existe la opción de hacer una búsqueda sobre  $\tau_s$ , para cada valor de  $\tau_t$  fijo, hasta que la verosimilitud ya no incremente su valor.

Estimación de los parámetros del ruido autoregresivo.

El algoritmo consiste de :

- 1) evaluar la verosimilitud suponiendo que el modelo no tiene componente autoregresiva,
- 2) a partir de las inovaciones estimar inicialmente a  $\alpha = [ \alpha_1, \dots, \alpha_p ]$  por mínimos cuadrados,
- 3) escoger como valor inicial de  $\alpha$  ( $\alpha(0)$ ), aquel valor que maximice la verosimilitud entre 1) y 2),
- 4) repetir 5) y 6) hasta que el incremento en la verosimilitud sea pequeño ( $1E-3$ ),
- 5) si  $\alpha(i)$  es la estimación de  $\alpha$  en la  $i$ -ésima iteración, obtener la dirección de búsqueda ( $h(i)$ ),
- 6) el siguiente valor de  $\alpha$  ( $\alpha(i+1)$ ) a partir de  $\alpha(i)$ , es el primer vector de la sucesión 
$$\alpha(i) + h(i) / (2^j), \quad j = 1, \dots, 12,$$
 que dé un incremento en la verosimilitud.

Falta por definir, la manera de encontrar la dirección de búsqueda, es decir, una dirección que incluya algún punto donde se incremente la verosimilitud.

Sólo para encontrar la dirección de búsqueda, se aproximó a la log-verosimilitud por una función que mide :

a) las magnitudes de las inovaciones

$$v(t) = y(t) - T(t) - S(t) - w(t-1)\alpha_1 - \dots - w(t-p)\alpha_p,$$

b) la consistencia de  $\alpha$  con las estimaciones del ruido autoregresivo, es decir,  $\alpha$  debe hacer pequeñas a las cantidades

$$u(t, \alpha) = w(t) - w(t-1)\alpha_1 - \dots - w(t-p)\alpha_p,$$

para medir la discrepancia de éstos objetivos, se trabajó con la función F

$$F(\alpha) = \sum v^2(t)/\sigma_1^2 + \sum u^2(t, \alpha)/\sigma_2^2, \quad (B.3)$$

y se definió a la "Log-verosimilitud" como

$$L = -F - (n/2) * \log(\sigma_1^2) - (n/2)\log(\sigma_2^2). \quad (B.4)$$

#### Método 1.

Con el vector  $\alpha(i)$ , obtener las inovaciones  $v(t)$  y las estimaciones del componente autoregresivo  $w(t)$ ; para encontrar  $\alpha(i+1)$ , el método 1 encuentra el vector  $\alpha(i+1)$  que minimiza la función F, y la dirección de búsqueda es  $h = \alpha(i+1) - \alpha(i)$ .

#### Método 2.

Se uso el método dado por Pesaran(1980), en la estimación de modelos de regresión con ruido autoregresivos, es decir,

$$z_t = x_t \beta + u_t, \quad (B.5)$$

donde  $u_t = u_{t-1}\alpha_1 + \dots + u_{t-p}\alpha_p + e_t$ ,

$e_t$  son v.a. normales, tales que  $E\{e_k e_l\} = \delta(k-l)\sigma^2$  pero su aplicación no es directa, porque en lugar de tener las cantidades fijas  $x_t \beta$ , se tienen las cantidades estocásticas  $T_t + S_t$ .

El método 2 consiste en usar

$$h = (D^{-1}/N) \left[ \begin{array}{c} \frac{\delta L}{\delta \alpha_1}, \dots, \frac{\delta L}{\delta \alpha_p} \end{array} \right]_{\alpha = \alpha(i)}, \quad (B.6)$$

donde  $D^{-1}/N$  es la matriz asintótica de covarianzas de los estimadores máximo-verosímiles de los parámetros del proceso autoregresivo en el modelo (B.5), y aproximar a la log-verosimilitud por (B.4)..

La expresión para  $D^{-1}$  es,

$$D^{-1} = \phi \phi' - \theta \theta', \quad \text{con } \alpha_0 = 1, \text{ y}$$

$$\phi = \begin{bmatrix} \alpha_0 \\ \alpha_1 & \alpha_0 \\ \vdots & \vdots & \ddots \\ \alpha_{p-1} & \alpha_{p-2} & \cdots & \alpha_0 \end{bmatrix} \quad \theta = \begin{bmatrix} \alpha_p \\ \alpha_{p-1} & \alpha_p \\ \vdots & \vdots & \ddots \\ \alpha_1 & \alpha_2 & \cdots & \alpha_p \end{bmatrix}$$

Al probar el programa, se observó que los 2 métodos si proporcionan direcciones que incrementan la verosimilitud, pero el método 2 converge en un menor número de iteraciones de manera significativa. Otro aspecto que se tuvo en cuenta, fue ignorar las primeras  $d$  innovaciones porque son muy grandes y no representan a verdaderas innovaciones.

Uno de los procesos mas costosos, es la actualización de las matrices de covarianzas, por lo que algunos autores han propuesto, que en el momento que se considere, las matrices de covarianza ya han convergido, ya no se actualicen.

Se implementó el método de Gersh y Kitagawa(1984), porque fue la referencia inicial de trabajo, pero otras alternativas son :

a) usar el algoritmo EM (Harvey,1989,p.188), que consiste en resolver sucesivamente la ecuación

$$E \left\{ \frac{\partial \ln L(\psi | Y)}{\partial \psi} \mid y_1, \dots, y_n \right\} = 0,$$

y parece ser un mejor método.

b) Usar un método de maximización general.

La manera de usar el programa K se explica a continuación.

El programa K se forma al conjuntar los archivos KLIB,OPMAT y LEC, el contenido de los archivos es el siguiente,

ARCHIVO	CONTENIDO
KLIB	Filtro de Kalman
OPMAT	Operaciones entre matrices
LEC	Rutinas de entrada y salida

Al ejecutarse el programa, aparece un menú con los encabezados :

- 0.- Salir del programa
- 1.- Generar Variables (Simulación)
- 2.- Ajustar un modelo
- 3.- Leer variables regresoras de un archivo
- 4.- Leer Variables a ajustar
- 5.- Borrar Variables
- 6.- Crear variables autoregresivas
- 7.- Crear variables con el número de días por periodo
- 8.- Modificar Opciones
- 9.- Observaciones faltantes.
- 10.- Dar valor inicial para simulaciones
- 11.- Rutina cero (  $y(t) = y(t) - y(0)$  )
- 12.- Rutina uno (  $y(t) = y(t)*a$  )

En casi todas las opciones la especificación de los valores de entrada es muy clara, y sólo se explica la opción 8.

Al dar la opción 8 aparecen una serie de parámetros junto con su valor actual. El número que aparece en la columna de Parámetro es el valor inicial del parámetro.

Parámetro	Significado y Posibles Valores.
Monitor Kalman (0)	(0,1) Monitoréa la búsqueda de los parámetros
Monitor Suaviza (0)	(0,1) Monitoréa el algoritmo de suavizamiento
Monitor Veros (0)	(0,1) Monitor más extenso que monitor Kalman.
Monotor Ajusta (0)	(0,1) Monitoréa las ecuaciones del filtro de Kalman.
Monitor Newton (0)	(0,1) Monitoréa el algoritmo Quasi-Newton.
Opt. Kalman	0 Búsqueda exhaustiva 1 Búsqueda no exhaustiva
# Correlaciones (20)	( 10,11,...,30).
# Pronósticos (0)	( 0,1,...,30).
# Valor grande (1e6)	$\lambda \in [0, \infty)$ para inicializar a la matriz inicial
est_mat (1)	1 Estima la matriz de covarianza, al resolver la ecuación de Riccati. 0 La matriz inicial es $\lambda I$
inovación	1 La verosimilitud se concentra en r 0 No se concentra la verosimilitud.

## Bibliografía.

Se usaron las abreviaturas,

JASA : Journal of the American Statistical Association,

J.R.S.S : Journal of the Royal Statistical Society.

Anderson, B.D.O. y Moore, J.B. (1979) : Optimal Filtering, Prentice-Hall, Englewood Cliffs.

Anderson, T.W. (1971) : The Statistical Analysis of Time Series, Wiley, New York.

Ansley, C.F., y Khon R. (1985) : "Estimation, Filtering and Smoothing in State Space Models With Incompletely Specified Initial Conditions", The Annals of Statistics, V. 13, p. 1286-1316.

Akaike, H. (1980) : " Seasonal Adjustment by a Bayesian Modeling", Journal of Time Series Analysis, V. 1, p. 1-13.

Bell, W.R. y Hillmer, S.C (1984) : "Issues involved with the Seasonal Adjustment of Economic Time Series", J. of Buss. and Econ. Stat., V. 2, p. 291-320.

Bierman, G.J. (1977) : Factorization Methods for Discrete Sequential Estimation, Academic Press, New York.

Box, G.E.P. y Muller, M.A. (1958) : "A Note on the Generation of Random Normal Deviates", Ann. Math. Stat. V. 29, p. 610.

Box, G.E.P. y Jenkins, G.M. (1976) : Time Series Analysis Forecasting and Control, San Francisco : Holden Day.

Box, G.E.P. y Tiao (1975) "Intervention Analysis with Applications to Economic an Environmental Problems", JASA, V 70, p. 70 -79.

Burman, J.P. (1980) : "Seasonal Adjustment by Signal Extraction", J.R.S.S, Ser. A, V. 143, p 321-337.

Burridge, P., y Hall A. (1987) : "Convergence of the Kalman Filter Gain for a Class of Nondetectable Signal-Extraction Problems", IEEE Transactions on Automatic Control, V. AC-31, p. 1036-1039.

Caines, P.E. (1988) : Linear Stochastics Systems, John Wiley & Sons, New York.

Cleveland, W.S. y Devlin, S.J. (1980) : "Calendar Effects in Monthly Time Series : Detection by Spectrum Analysis and Graphical Methods", JASA, V. 75, p. 487- 496.



Cleveland, W.P. y Tiao, G. C. (1978) : "Descomposition of Seasonal Time Series : A model for the Census X-11 Program", JASA, V. 71, p. 591-597.

Chan, S.W., Goodwin, G.C., y Sin K.S. (1984) : "Convergence Properties of the Riccati Difference Equation in Optimal Filtering of Nonstabilizable Systems", IEEE Transactions on Automatic Control, V. AC-29, p. 110 -118.

Chen, H.F. (1985) : Recursive Estimation and Control for Stochastic Systems, John Wiley & Sons, New York.

Conover, W.J. (1981) : Practical Nonparametric Statistics, John Wiley & Sons, New York.

D'Agostino, R.B., Stephens, M.A. (1986) : Goodness of Fit Techniques, Marcel Decker.

Daum, F.E. (1988): "New Exact Non-Linear Filters", en Bayesian Analysis of Time Series an Dynamic Models, Ed. por James C. Spall, p. 199-228.

Dagum, E.B. (1978) : "A Comparasion and Assesment of Seasonal Adjustment Methods for Employment and Unemployment Statistics", National Comission on Employment an Unemployment Statistics, Background Paper No. 5, Washington D.C., Government Printing Office, p. 1-94.

Dagum, E.B. (1983) : The X-11-ARIMA Seasonal Adjustment Method, Statistics Canada, Time Series Research and Analysis Staff, Ottawa, Catalogue 12-584E.

D.E.C.(1985) : VAX/VMS Run-Time Library Routines, Reference Manual, Maynard Massachusetts, p. RTL-72.

Duncan, D. y Horn, S. (1972) : "Linear Dynamic Recursive Estimation from the Viewpoint of the Regression Analysis", JASA, V. 57, p. 815-821.

Durbin, J. y Murphy, M.J. (1975) : "Seasonal Adjustment Based on a Mixed Additive-Multiplicative Model", J.R.S.S, Ser. A, V. 138, Parte 3, p. 385-409.

De Jong, P. (1988) : "The Likelihood for State Space Model", Biometrika, V. 75, p. 185-189.

De Souza, C.E., Givers, M.R., y Goodwin G.C. (1986): "Riccati Equations in Optimal Filtering of Nonstabilizable Systems Having Singular State Transition Matrices", IEEE Transactions on Automatic Control, V. AC-31, p. 831-838.

Eisenpress, H. (1956) : "Regression Techniques Applied to Seasonal Corrections and Adjustment for Calendar Shifts", JASA, V. 51, p. 615-621.

Gardner, G., Harvey, A.C., y Phillips, G.D.A (1980) : "An Algorithm for Exact Maximum Likelihood Estimation of Autoregressive Moving Average Models by the Kalman Filter ", Applied Statistics, V. 29, p. 311-322.

Gersh, W. y Kitagawa, G.(1984) : "A smoothness Priors State Space Modeling of Time Series with Trends and Seasonality", JASA, V. 79, p. 378-389.

Grether, D.M., y Nerlove, M. (1970) : "Some properties of Optimal Seasonal Adjustment", Econometrica, V. 38, p. 682-703.

Hannan, E.J., y Diestler, M. (1988) : The Statistical Theory of Linear Systems, John Wiley & Sons, New York.

Hausman, J.A. y Watson, M.A.(1985) : "Errors in variables and Seasonal Adjustment Procedures", JASA, V. 80, p. 531-540.

Harvey, A.C., y Phillips, G.D.A (1979) : " Maximum Likelihood Estimation of Regression Models with Autoregressive Moving Average Disturbances", Biometrika, V. 66, p. 49-58.

Harvey, A. C., y Franzini, L. (1983) : "Testing for Deterministic Trend and Seasonal Components in Time Series Models", Biometrika, V. 70, p. 673-682.

Harvey, A.C. (1989) : Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press.

Hillmer, S.C., y Tiao, G.C. (1982) : "An ARIMA-Modeling-Based Approach to Seasonal Adjustment", JASA, V. 77, p. 63-70.

Imhof (1961) : "Computing the Distribution of Quadratic Forms in Normal Variables", Biometrika, V. 48, p. 419 -426

Jones, R.H. (1985) : "Time Series Analysis With Unequally Spaced Data", en Handbook of Statistics, Vol 5, Ed. por E.J. Hannan, P.R. Krishnaiah, M. M. Rao.

Jorgenson, D.W. (1967) : "Seasonal Adjustment for Econometric Analysis", JASA, V. 62, p. 137-140.

Kailath, T. (1980) : Linear Systems, Prentice-Hall, Englewood Cliffs, New Jersey.

Kendall, M.G., Stuart, A., y Ord, J.K. (1983) : The Advanced Theory of Statistics , Vol III, Charles Griffin & Co. Ltd, Londres.

- Kenny, P.B. y Durbin, J. (1982) : "Local Trend estimation and Seasonal Adjustment of Economic and Social Time Series", J.R.S.S., Ser A, No. 145, p 1-41.
- Lothian, J. y Morry, M.(1978) : "A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Program", Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Lovell, M.C. (1966) : Alternative Axiomatizations for Seasonal Adjustment, JASA, V. 61, p. 800-802.
- La Motte, L.R., y Mc Worther, A. (1978) : "An Exact Test for the Presence of Random Walk Coefficients in a Linear Regression Model", JASA, V. 73, p. 816-820.
- Ljung, G.M. y Box, G.E.P. (1978) : "On a Measure of Lack of Fit in Time Series Models", Biometrika, V. 66, p. 67-72.
- Nerlove, M. (1964) : "Spectral Analysis of Seasonal Adjustment Procedures", Econometrica, Vol. 32, p.241-286.
- Nerlove, M., Grether, D.M. y Carvalho, J.L. (1979), Analysis of Economic Time Series : A Synthesis, New York: Academic Press.
- Nyblom, J. (1986) : "Testing for Deterministic Linear Trend in Time Series", JASA, V. 81, p. 545-549.
- Pesaran, M.H. (1980), Dynamic Regression : Theory and Algorithms, Chichester, Gr. Bret.
- Peña, D. y Guttman, I. (1988) : "Bayesian Approach to Robustifying the Kalman Filter", en Bayesian Analysis of Time Series an Dynamic Models, Ed. por James C. Spall, p. 227-254.
- Pierce, D.A. (1979) : "Signal Extraction Error in Nonstationary Time Series", The Annals of Statistics, V. 7, p. 1303-1320.
- Poole, A., West M., y Harrison J. (1988) : "Nonnormal and Nonlinear Dynamic Bayesian Modeling", en Bayesian Analysis of Time Series an Dynamic Models, Ed. por James C. Spall, p. 167-197.
- Rosenblat, H.M. (1968) : "Spectral Evaluation of BLS and Census Revised Seasonal Adjustment Procedures", JASA, V. 63, p. 472-503.
- Shiskin, J. y Eisenpress, H. (1957) : "Seasonal Adjustment by Electronic Computer Methods", JASA, V. 52, p. 415-451.
- Shiskin, J., Young, A.H., y Musgrave, J.C. (1967) : " The X-11 Variant of Census Method II Seasonal Adjustment", Technical Paper No. 15, Bureau of the Census, U.S., Departamnet of Commerce.

Sims, C.A. (1974) : "Seasonality in Regression", JASA, V. 69, p. 618-62.

Tiao, G.C. (1985) : " Autoregressive Moving Average Models, Intervention Problems and Outlier Detection in Time Series", en Handbook of Statistics, Vol 5, Ed. por E.J. Hannan, P.R. Krishnaiah, M. M. Rao.

Thomas, J.J. y Wallis, K.F.(1971) : "Seasonal Variation in Regression Analysis", JRSS, Ser. A, V. 134, p. 57-72.

Wallis, K.F. (1974) : "Seasonal Adjustment and Relation Between Variables", JASA, V. 69, p. 18-31.

Wichman, B.A. y Hill, I.D.(1982), An Efficient and Portable Pseudo-Random Number Generator, JRSS, Ser C, V.31, No. 2.

Young, A.H. (1968) : "Linear Aproximations to the Census and BLS Seasonal Adjustment Methods", JASA, V. 63, p. 445-471.

Zellner, A. (1971) : " An Introduction to Bayesian Inference in Econometrics", Wiley, New York.