

24.
18

Estimación de Parámetros de Modelos ARMA(p,q) usando Mínimos Cuadrados No-lineales

TESIS PROFESIONAL
Hortensia Góngora Treviño

U. N. A. M.

Facultad de Ciencias
A C T U A R I A



Cd. Universitaria 1986



Universidad Nacional
Autónoma de México



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

Introducción	1
Capítulo 1 CONCEPTOS BASICOS EN EL ANALISIS DE SERIES DE TIEMPO UNIVARIADAS	5
1.1 Definición.	5
1.2 Objetivos del Análisis en Series de Tiempo.	6
1.2.1 Descripción.	6
1.2.2 Explicación.	7
1.2.3 Simulación.	7
1.2.4 Pronóstico.	7
1.2.5 Control.	8
1.3 Principales Componentes de una Serie de Tiempo.	8
1.3.1 Tendencia.	9
1.3.2 Estacionalidad.	9
1.3.3 Ciclos.	9
1.3.4 Residuales o Ruido.	10

1.4 Algunos de los Distintos Tipos de Modelos en Series de Tiempo Univariadas.	10
1.4.1 Método de Descomposición.	11
1.4.2 Modelo de Box y Jenkins.	12
1.5 Etapas del Análisis de Series de Tiempo Univariadas.	16
1.6 Procesos Estocásticos.	19
1.7 Estacionariedad.	23
1.7.1 Estacionariedad Estricta.	24
1.7.2 Estacionariedad débil o de Segundo Orden.	25
1.8 Ergodicidad.	28
1.9 Descripción y Propiedades de Modelos ARMA(p,q).	30
1.10 Procesos Autorregresivos.	35
1.11 Procesos de Promedio Móvil.	36
1.12 Procesos Mixtos Autorregresivos y de Promedios Móviles.	38
Capítulo 2 ESTIMACION DE PARAMETROS DEL MODELO ARMA(p,q) EN EL DOMINIO DEL TIEMPO	40
2.1 Métodos de Estimación.	40

2.1.1 Verosimilitud No Condicional.	42
2.1.2 Enfoque No Condicional.	50
2.2 Verosimilitud Condicional.	57
2.2.1 Enfoque Condicional.	60
2.3 No Linealidad en los Parámetros.	63
2.4 Método General para obtener Estimadores Iniciales de los Parámetros de un Proceso ARMA(p,q).	70
2.5 El Método de Mínimos Cuadrados No Lineales a través de las Autocovarianzas.	78
2.5.1 Estimadores de los Parámetros de un Modelo ARMA(p,q) a través de las Autocovarianzas.	79
2.5.2 Propiedades de los Estimadores de las Autocovarianzas.	84
2.5.3 Ajuste de las Autocovarianzas por Mínimos Cuadrados No Lineales.	89

Capítulo 3 ALGORITMO DE MINIMOS CUADRADOS NO
LINEALES PARA LA

ESTIMACION DE PARAMETROS DE MODELOS ARMA (p,q)	99
3.1 Introducción.	99
3.2 Relaciones Básicas.	102
3.3 Implementación del Algoritmo.	120
3.3.1 Estimador de las Autocovarianzas.	120
3.3.2 Estimador inicial del vector de parámetros .	122
3.3.3 Determinación del vector de autocovarianzas ().	125
3.3.4 Determinación del Jacobiano del vector de Autocovarianzas ()	135
Capítulo 4 RESULTADOS	185
4.1 Resultados Adicionales.	198
Chapter 5 CONCLUSIONES	203
Apéndice A ANTECEDENTES ESTADISTICOS	205
A.1 Procesos Gaussianos.	205
A.2 Función Generadora de Autocovarianzas y Función de Densidad Espectral de un Proceso ARMA(p,q).	206

Apéndice B PROGRAMA	211
Apéndice C RUTINAS DE PAQUETERIA	226
Referencias	243

INTRODUCCION

La mayoría de los métodos estadísticos, han sido diseñados para tratar con observaciones independientes. Sin embargo, en muchas áreas de aplicación tales como, economía, metereología, medicina y otras, las observaciones ocurren en forma de series de tiempo, es decir, como una sucesión de valores ordenados por algún parámetro, generalmente el tiempo. Aquí, las observaciones son dependientes y la naturaleza de esta dependencia es de importancia considerable. El Análisis de Series de Tiempo comprende, a las técnicas que se utilizan en el estudio de este tipo de series de observaciones dependientes.

Hay dos caminos complementarios, más que alternativos, de llevar a cabo el análisis de una serie de tiempo. El primero, está enfocado a estudiar las relaciones que existen entre las observaciones de la serie, en el dominio del tiempo y el segundo, conocido como Análisis Espectral, estudia los movimientos cíclicos de la serie, en el dominio de las frecuencias. Este último, aunque también muy importante, no será discutido en esta tesis.

Existen diferentes objetivos y fases en el análisis de series de tiempo. En el capítulo 1 se hace una revisión muy general, de los conceptos básicos, necesarios para el estudio de series de tiempo univariadas.

Entre las distintas fases del análisis de series de tiempo, una de las más importantes es la construcción de un modelo matemático para la serie. Una de las técnicas más utilizadas para este fin, es la de Box y Jenkins (1970). Esta técnica consiste, en representar a una serie de tiempo, después de haber aplicado algunas transformaciones y/o tomado diferencias apropiadas de la misma, por medio de un modelo con el menor número posible de parámetros. Tales modelos son conocidos como modelos ARMA(p,q) (autorregresivos y de promedios móviles). El procedimiento para construir tales modelos, así como su descripción y propiedades, se explican también en el capítulo 1.

Para construir modelos ARMA(p,q), se necesitan estimar los coeficientes o parámetros que aparecen en dicho modelo. Se han hecho estudios extensivos en la literatura con respecto a este problema, el cual durante mucho tiempo, ha generado una serie de controversias, en cuanto a las propiedades de los estimadores obtenidos por distintos métodos, y en cuanto a la eficiencia computacional para obtenerlos.

El capítulo 2, se dedica exclusivamente al problema de

estimación de parámetros de modelos ARMA, en el dominio del tiempo. En la sección 2.1, se presentan dos de los métodos de estimación más utilizados: Máxima Verosimilitud y Máxima Verosimilitud Condicional. Alrededor de estos métodos, se han desarrollado otros que vienen a ser modificaciones de los primeros, todos con el propósito común de aproximarse lo más posible a la Función de Verosimilitud Exacta. En esta misma sección, se hace una presentación de algunos problemas adicionales, derivados del de estimación, como lo son, no-linealidad en los parámetros y la obtención de estimadores iniciales de los mismos.

El interés del presente trabajo, está centrado en el problema de estimación de parámetros de modelos $ARMA(p,q)$, en el dominio del tiempo. Su objetivo es presentar un método de estimación, que utiliza las autocovarianzas teóricas y las muestrales. El método trata de aproximar las autocovarianzas muestrales, con las autocovarianzas teóricas, expresadas éstas en términos de los parámetros a estimar. La búsqueda de los valores que deben asignarse a los estimadores de los parámetros, resulta de aplicar el criterio de Mínimos Cuadrados. El procedimiento que resulta, se propone como una alternativa a los métodos Máximo Verosímiles.

Con base en algunos experimentos de simulación, se compara el método de estimación estudiado en este trabajo, con el de David J. Pack (1974), que utiliza una aproximación a la función de verosimilitud, utilizando el enfoque

condicional propuesto por Box y Jenkins. Finalmente se discuten los resultados observados, en términos del Error Cuadrático Medio de los estimadores.

El trabajo descrito en los dos párrafos anteriores, se trata en distintas partes de la tesis. En el capítulo 2, sección 2.2, se hace una presentación formal del método de estimación de parámetros de modelos ARMA(p,q), a través de las autocovarianzas del proceso. También en esa sección, se discuten algunas de las propiedades de los estimadores de las autocovarianzas utilizadas, así como, las propiedades que se espera, posean los estimadores de los parámetros obtenidos por ese método. En el capítulo 3, se presenta el algoritmo desarrollado para este fin, así como, todos los detalles de su implementación en una computadora Burroughs 7800. En el capítulo 4, se presentan los resultados de los experimentos de simulación, realizados para algunos modelos ARMA(p,q). Los hechos relevantes de las simulaciones, se discuten en el capítulo 5.

Finalmente, se incluyen tres apéndices. El apéndice A contiene algunos antecedentes estadísticos, que por su naturaleza, no fueron incluidos en el texto y a los cuales se hace referencia en los capítulos 2 y 3. En el apéndice B, se presenta un listado del programa que se elaboró para el presente trabajo. Por último, en el apéndice C se incluye la descripción de las subrutinas de paquetería, que se utilizaron como apoyo computacional.

Capítulo 1

CONCEPTOS BASICOS EN EL ANALISIS DE SERIES DE TIEMPO UNIVARIADAS

Este capítulo pretende hacer una revisión muy general de los conceptos básicos, necesarios para el análisis de series de tiempo univariadas.

1.1 Definición.

Una serie de tiempo se puede definir como un conjunto de observaciones de una variable cuantitativa y no determinística, denotada por X , tomadas en periodos de tiempo igualmente espaciados. Esta variable X se puede referir a temperatura de un proceso químico, número (anual) de manchas solares, número (mensual) de pasajeros en una

aereolínea, venta (mensual) de automóviles, en fin un sinnúmero de fenómenos que son de interés en diversas áreas del quehacer humano. Los períodos de tiempo pueden ser minutos, días, meses, trimestres, años, etc., dependiendo del tipo de fenómeno, o de los fines del investigador.

1.2 Objetivos del Análisis en Series de Tiempo.

Los objetivos o metas que se persiguen al analizar una serie de tiempo, no siempre son los mismos, sino que dependen, de las varias razones que pueda tener un investigador para llevar a cabo el estudio de una serie de tiempo. Estas razones generalmente, van a determinar el tipo de análisis que se decida aplicar. Los objetivos del análisis se pueden clasificar de la siguiente manera:

1.2.1 Descripción.

En ocasiones el estudio de una serie de tiempo, consiste en describir el comportamiento de un fenómeno, durante un cierto período de tiempo, resumir sus propiedades y destacar sus principales características, para lo cual, un análisis descriptivo de la serie de tiempo es suficiente.

1.2.2 Explicación.

En ocasiones el interés principal, puede estar enfocado a explicar el comportamiento de una serie de tiempo, a partir del comportamiento de otras variables, para tener un mejor conocimiento del mecanismo que generó dicha serie de tiempo. Por ejemplo, en el área de Economía puede desearse explicar el comportamiento de la producción, a partir del capital y del trabajo.

1.2.3 Simulación.

El interés, bien puede estar concentrado en contar con un modelo matemático, que sea capaz de simular un fenómeno con las mismas características del fenómeno real, representado a través de la serie de tiempo original.

1.2.4 Pronóstico.

A partir de 1.2.1 o 1.2.2, se puede estar interesado en predecir el comportamiento futuro del fenómeno bajo estudio, esto se presenta muy a menudo en series de tiempo económicas o industriales.

1.2.5 Control.

Este objetivo esta muy relacionado con 1.2.2 y con 1.2.4, ya que, cuando se conoce qué variables y cómo influyen en el comportamiento de cierto fenómeno, o después de conocer cuál va a ser el comportamiento de un fenómeno en el futuro, se puede intentar tomar medidas correctivas a tiempo, que impidan la ocurrencia de un evento no deseado.

Es importante destacar que en todos los objetivos mencionados anteriormente, se necesita de un modelo matemático que describa de la mejor manera posible el mecanismo que dio origen a los datos observados.

1.3 Principales Componentes de una Serie de Tiempo.

En ciertas áreas de aplicación, en particular en economía, se tiene el concepto de que una serie de tiempo es una mezcla de cuatro componentes: la tendencia, la estacionalidad, los ciclos y los residuales o componente de ruido.

1.3.1 Tendencia.

Se dice que una serie de tiempo presenta una tendencia, cuando muestra un movimiento sistemático prolongado, que bien puede ser creciente, decreciente, o ambos. En otras palabras, se puede decir que existe una tendencia en una serie, cuando ésta presenta cambios prolongados en la media.

1.3.2 Estacionalidad.

Estacionalidad, se puede definir como un cambio o variación sistemática en la serie de tiempo, que se repite cada s períodos de tiempo, es decir, como si el comportamiento de una parte de la serie fuese similar cada s períodos de tiempo.

1.3.3 Ciclos.

Este componente es uno de los más difíciles de aislar y de explicar. Se define como una variación cíclica de período no identificado y mayor de un año. En ocasiones, el período asociado a esta variación cíclica es más largo que el período de observación de la serie y suele confundirse con una componente de tendencia.

1.3.4 Residuales o Ruido.

Los residuales se definen como las fluctuaciones que quedan en la serie, después de haber separado de ésta, las componentes de tendencia y variación estacional. En el mejor de los casos, estos residuales pueden ser completamente aleatorios, en el otro, pueden comportarse como una serie que presente movimientos oscilatorios suaves. Generalmente, estos residuales no son puramente aleatorios y tienen un cierto comportamiento sistemático, en el sentido de que no son independientes, y el valor del residual al tiempo t ϵ_t , depende de alguno o algunos residuales anteriores.

1.4 Algunos de los Distintos Tipos de Modelos en Series de Tiempo Univariadas.

El propósito de esta sección, es presentar de una manera un tanto informal, dos de las técnicas más utilizadas en modelar series de tiempo, la primera de ellas conocida tradicionalmente como "Método de Descomposición", y la segunda el método muy conocido y usado de "Box y Jenkins".

1.4.1 Método de Descomposición.

La filosofía básica de este modelo, consiste en considerar a una serie de tiempo Z_t , como el resultado de la suma de tres componentes:

$$Z_t = T_t + S_t + I_t$$

donde, T_t es el valor de la tendencia en la serie, S_t es el valor de la estacionalidad, e I_t es el residual de la serie, mejor conocido como ruido. Estos componentes fueron explicados brevemente en la sección anterior.

El método consiste de una técnica, por medio de la cual se pueda separar y estimar, cada ~~uno~~ de los tres componentes que conforman la serie, para después, extrapolar cada uno de ellos hacia el futuro, sumarlos y obtener el pronóstico deseado de la serie Z_t , por ejemplo, al tiempo Z_{t+L} .

Se han desarrollado varios métodos, para poder separar cada componente en la serie de la mejor manera posible. El lector interesado en tales métodos, puede consultar Chatfield (1974), Kendall (1966) y, Nerlove, Grether y Carvalho (1979), éstos últimos trabajan mucho con estos modelos, a los que llaman Modelos de Componentes no Observables. Estos métodos no serán tratados aquí, ya que el

interés principal de la presente tesis, esta relacionado con el modelaje matemático propuesto por Box y Jenkins (1970).

1.4.2 Modelo de Box y Jenkins.

El modelo Mixto Autorregresivo y de Promedio Móvil ARMA(p,q), especialmente, fue diseñado para poder describir el comportamiento de los residuales o ruido, de una serie de tiempo Z_t .

El método consiste en obtener series que sean aproximadamente estacionarias, es decir, que aproximadamente posean una media y una varianza constantes. (Más adelante en la sección 1.7, se describe con más detalle el concepto de estacionariedad.)

En la realidad es muy difícil, si no es que imposible, encontrarse con series de tiempo estacionarias, por lo que en algunos casos puede ser necesario aplicar transformaciones estabilizadoras de varianza a los datos, como por ejemplo, la transformación logarítmica. Box y Jenkins proponen el uso de transformaciones simplificadoras, cuyo fin, es obtener series aproximadamente estacionarias a las que se ajuste un modelo ARMA(p,q), estacionario e invertible, en una forma parsimoniosa, es decir, con un número reducido de parámetros. (En la sección 1.9 se describen con detalle estos modelos, así como sus propiedades.)

Las transformaciones simplificadoras son:

$$a) (1 - R)^D = \nabla^D$$

$$b) (1 - R^S)^D = \nabla_S^D$$

donde :

R es el operador de atraso, definido como

$$R^\tau Z_t = Z_{t-\tau} \quad ; \quad \tau = 1, 2, \dots$$

y

$$R^0 Z_t = Z_t$$

Este operador, tiene la propiedad de que puede ser manipulado como cualquier cantidad algebraica, y juega un papel importante en el desarrollo analítico de series de tiempo (ver Fuller (1976), cap. 2, sec. 2.4).

d es un entero positivo que indica el número de diferencias de la serie Z_t , como se verá más adelante.

S es un entero positivo que está relacionado con el período de una serie estacional.

D es un entero positivo que juega un papel similar al

de d.

∇ es el primer operador diferencial, tal que $\nabla Z_t = Z_t - Z_{t-1}$, y en términos del operador de atraso R, es igual a $(1-R)Z_t$.

La transformación en (a), se aplica a series de tiempo que tienen un comportamiento no estacionario, en el sentido de que no poseen una media constante. La idea de aplicar esta transformación es eliminar la componente de tendencia de la serie original, no estacionaria Z_t , así, al aplicar

$$(1 - R)^d Z_t = X_t$$

se espera que la nueva serie X_t , igual a la d-ésima diferencia de Z_t , sea estacionaria, y por lo tanto, pueda ser explicada por un modelo estacionario ARMA.

La transformación en (b) se aplica a series de tiempo que tienen un comportamiento estacional, es decir, cuya característica fundamental, es que observaciones separadas s instantes de tiempo son similares. Se espera que el operador $R^s Z_t = Z_{t-s}$, juegue un papel importante en el análisis de este tipo de series. Adicionalmente, puede suceder que la nueva serie $Z_t, Z_{t-s}, Z_{t-2s}, \dots$ presente un comportamiento no estacionario, y de ahí, que una transformación del tipo

$$(1 - R^s)^D Z_t = X_t$$

sea útil para acercarse a lo que podría ser una serie estacionaria, de la cual, se han separado las componentes de tendencia y estacionalidad.

Los modelos que tratan con series del tipo $\nabla^d Z_t$ y $\nabla_s^D Z_t$, son conocidos como procesos Mixtos Integrados Autorregresivos y de Promedios Móviles ARIMA.

Resumiendo, en un proceso no estacionario, por ejemplo un ARIMA(p,d,q), la d-ésima diferencia de las observaciones es modelada por un proceso estacionario ARMA(p,q).

En adelante, cuando se haga referencia a una serie de tiempo X_t , se supondrá que dicha serie es estacionaria, es decir, que es el resultado de haber aplicado algunas de las transformaciones descritas anteriormente a la serie Z_t . Y en consecuencia, se hará referencia a procesos ARMA que modelen en forma adecuada a la serie X_t , más que a procesos ARIMA.

El lector interesado en conocer más sobre modelos ARIMA y su aplicación, puede consultar Box y Jenkins (1970), o Nelson (1973), quien hace una descripción sencilla de los mismos.

1.5 Etapas del Análisis de Series de Tiempo Univariadas.

Como se mencionó en la sección 1.3, las etapas del análisis en series de tiempo, dependen de los objetivos que pueda tener el investigador, interesado en llevar a cabo el estudio de una serie de tiempo. Ahí se describieron los distintos objetivos que puede haber, y fue claro, que en todos ellos es necesario obtener un modelo que explique el comportamiento de la serie. Aquí surgen un par de preguntas: ¿El procedimiento para obtener un modelo, es el mismo en todos los casos?, ¿Cómo se obtiene dicho modelo?. La respuesta a la primer pregunta es NO, ya que como se vio en la sección 1.4, no hay un procedimiento universal para modelar a una serie de tiempo.

Se puede ver simplificada la decisión, en cuanto a qué técnica de modelaje es la más apropiada, si se decide con anterioridad, cuál de los siguientes dos objetivos es de interés:

- a) Encontrar un modelo que ajuste bien a los datos.
- b) Encontrar un modelo que haga buenos pronósticos de la serie.

En términos muy generales, se puede decir que el modelo de descomposición puede ser apropiado en el caso (a), pero puede no funcionar igual de bien en el caso (b). El modelo de Box y Jenkins en cambio, parece funcionar mejor en el caso (b).

El procedimiento para pronosticar con un modelo ARMA, consiste en encontrar primero una serie aproximadamente estacionaria, como se explicó en la sección 1.4 anterior, después, analizar las propiedades de esta nueva serie transformada X_T , y en base a éstas, escoger un modelo apropiado ARMA, a esta etapa del análisis Box y Jenkins la llaman Identificación. Este modelo es entonces ajustado y sus parámetros estimados, etapa a la cual Box y Jenkins llaman Estimación. En el capítulo 2 se trata este problema específico, el cual, ha generado durante mucho tiempo una serie de controversias en cuanto a las propiedades muestrales de los estimadores, y la eficiencia computacional para obtenerlos. Esto ha propiciado que muchos investigadores se interesen en el desarrollo de nuevos métodos de estimación, que cubran las deficiencias en otros. De hecho, el interés principal de esta tesis, es el desarrollo de un método alternativo de estimación de parámetros de modelos ARMA, como se verá en el capítulo 2. Una vez que los parámetros han sido estimados, se continua con un análisis de los residuales, etapa a la cual Box y Jenklins llaman Diagnóstico. Si el análisis anterior resulta satisfactorio, entonces, el modelo ARMA se utiliza para

hacer pronósticos, etapa a la cual Box y Jenkins por razones obvias llaman Pronóstico. Si este no es el caso, entonces, se procede a regresar a la etapa de Identificación, para buscar otro modelo ARMA que pueda ser más adecuado y se repite todo el proceso.

Una de las ventajas de usar modelos ARMA para pronósticos, es que se puede evaluar explícitamente el error cuadrático medio de predicción. (Ver Harvey (1981), cap. 1).

En lo que resta de este capítulo, se hace una presentación un poco más teórica, de algunos conceptos que se manejarán con mucha frecuencia en los siguientes capítulos, y que son esenciales para entender el análisis de series de tiempo.

Dado que la teoría desarrollada para modelar series de tiempo, considera a dichos modelos como procesos estocásticos, y a las series de tiempo como realizaciones de estos procesos, creo conveniente hacer una breve presentación sobre los mismos, con objeto de dar un panorama general sobre la clase especial a la cual pertenecen los modelos ARMA.

1.6 Procesos Estocásticos.

En un proceso estocástico, cada observación es una variable aleatoria y estas observaciones evolucionan a través del tiempo de acuerdo a ciertas leyes probabilísticas. Así, un proceso estocástico se puede definir como un conjunto de variables aleatorias, las cuales tienen un orden en el tiempo, es decir, $\{Z(t) ; t \in T\}$, donde T es el conjunto de instantes en el tiempo en el cual se define el proceso. Cuando T abarca todo un intervalo de valores (usualmente $-\infty < t < \infty$), entonces se dice que $\{Z(t) ; t \in T\}$ es un proceso estocástico continuo. Si T es un conjunto de instantes aislados (usualmente $t = 0, \pm 1, \pm 2, \dots$), entonces se dice que $\{Z_t ; t \in T\}$ es un proceso estocástico discreto.

Por simplicidad, la presentación siguiente se hará, suponiendo que el conjunto de instantes T es discreto, aunque la presentación en el caso de T continuo sería muy similar.

Una clase especial de los procesos estocásticos, son los procesos estacionarios, a la cual pertenecen los modelos ARMA descritos más adelante en la sección 1.9. Con objeto de

introducir los conceptos de estacionariedad y ergodicidad, así como su importancia en el análisis de series de tiempo, se comienza por definir la distribución probabilística de un proceso estocástico, y por analizar los problemas que surgen al tratar de describir el primer y segundo momentos del mismo, conocidos como la media, la varianza y la función de autocovarianzas.

A cada proceso estocástico $\{Z_t ; t \in T\}$, donde $T = 0, \pm 1, \pm 2, \dots$, le corresponde una distribución probabilística que queda determinada por la familia de todas las distribuciones conjuntas, de todos los subconjuntos finitos de variables aleatorias pertenecientes al proceso. Así la distribución del proceso incluye a las distribuciones marginales

$$\dots F(Z_{-2}), F(Z_{-1}), F(Z_0), F(Z_1), F(Z_2), \dots$$

también incluye a las distribuciones conjuntas entre pares de variables

$$\dots F(Z_{-1}, Z_0), F(Z_{-1}, Z_1), F(Z_0, Z_1), \dots$$

etc.

Sea $\{Z_t ; t \in T\}$ un proceso estocástico, para cualquier conjunto de instantes $T = \{t_1, t_2, \dots, t_N\}$, y cada entero N .

La media del proceso al tiempo t , se define como

$$\mu_t = E(Z_t) \quad ; \quad \forall t \in T$$

La función de autocovarianzas, se define como

$$\gamma(s, t) = \text{cov}(Z_s, Z_t) = E[(Z_s - \mu_s)(Z_t - \mu_t)] \quad ;$$

$$\forall s, t \in T$$

En particular, la varianza del proceso al tiempo t , se define como

$$\sigma_t^2 = \text{var}(Z_t) = \text{cov}(Z_t, Z_t) = E(Z_t - \mu_t)^2 \quad ; \quad \forall t \in T$$

Así para las variables aleatorias Z_{t_1} , Z_{t_2} , ..., Z_{t_N} , se tiene el vector de medias de dimensión N

$$(\mu_{t_1}, \mu_{t_2}, \dots, \mu_{t_N}) = (E(Z_{t_1}), E(Z_{t_2}), \dots, E(Z_{t_N})) \quad (1)$$

y la matriz simétrica de varianzas y covarianzas, de dimensión $N \times N$

$$\begin{pmatrix} \sigma_{t_1}^2 & \gamma(t_1, t_2) & \dots & \gamma(t_{N-1}, t_N) \\ & \sigma_{t_2}^2 & \dots & \cdot \\ & & \cdot & \cdot \\ & & & \sigma_{t_N}^2 \end{pmatrix} \quad (2)$$

Esta matriz como se mencionó anteriormente, es simétrica y definida semi positiva.

En la mayoría de los problemas estadísticos, se está interesado en estimar las propiedades de una población, a partir de una muestra representativa de la misma. En este caso, si se tuviese una muestra del conjunto de variables aleatorias Z_{t_1}, \dots, Z_{t_N} , se podría estimar, por ejemplo, el vector de medias en (1) como

$$\left(\hat{\mu}_{t_1}, \dots, \hat{\mu}_{t_N} \right) = \left(m^{-1} \sum_{i=1}^m Z_{t_1}^{(i)}, \dots, m^{-1} \sum_{i=1}^m Z_{t_N}^{(i)} \right) \quad (3)$$

donde, m sería el tamaño de la muestra y $Z_{t_i}^{(j)}$ sería la j -ésima observación de la variable aleatoria asociada al instante t_i . En otras palabras, si se pudiesen observar varias realizaciones de un proceso en el conjunto de instantes T , se podrían hacer inferencias acerca de las propiedades del proceso, del cual provienen las realizaciones.

En la práctica, sin embargo, lo común es observar una sola serie de tiempo durante un cierto período de tiempo, en estas circunstancias, no es posible hacer inferencias significativas acerca del modelo probabilístico, del cual proviene la serie de tiempo observada, a menos que se impongan ciertas restricciones al modelo o proceso generador de la serie. Esta necesidad lleva a los conceptos de estacionariedad y ergodicidad, descritos en las siguientes secciones.

1.7 Estacionariedad.

Estacionariedad, en el más amplio sentido de la palabra, se puede entender como algo que es estable.

Como se mencionó en la sección anterior, cuando se cuenta con una sola realización de un proceso, es necesario

que éste cumpla ciertas restricciones, para facilitar el análisis de la realización o serie de tiempo. Estacionariedad en sí, se puede ver como una condición que se pide al proceso generador, con este fin.

Existen dos tipos de estacionariedad :

1.7.1 Estacionariedad Estricta.

Un proceso estocástico $\{X_t ; t \in T\}$ se dice que es estacionario, si para cualquier conjunto de instantes $T = \{t_1, t_2, \dots, t_N\}$, cada N y cada desplazamiento t_0 de tiempo, se cumple que la distribución conjunta de las variables aleatorias X_{t_1}, \dots, X_{t_N} , es invariante con respecto al desplazamiento de tiempo

$$F(X_{t_1}, \dots, X_{t_N}) = F(X_{t_1+t_0}, \dots, X_{t_N+t_0}) ; \forall t_1, \dots, t_N, N, t_0 \quad (4)$$

Es decir, que no importa el punto particular en el tiempo que se elija, la distribución conjunta de las variables aleatorias X_{t_1}, \dots, X_{t_N} , va a ser siempre la misma que la de las variables aleatorias $X_{t_1+t_0}, \dots, X_{t_N+t_0}$.

Como una consecuencia de (4), se tiene que

$$F(X_t) = F(X_{t-t_0}) \quad ; \quad \forall t \in T$$

$$F(X_s, X_t) = F(X_{s-t_0}, X_{t-t_0}) \quad ; \quad \forall s, t \in T$$

(5)

y como consecuencia de (5), se tiene que

$$E(X_t) = E(X_{t-t_0}) \quad ; \quad \forall t \in T, t_0$$

$$\text{var}(X_t) = \text{var}(X_{t-t_0}) \quad ; \quad \forall t \in T, t_0$$

$$\text{cov}(X_s, X_t) = \text{cov}(X_{s-t_0}, X_{t-t_0}) \quad ; \quad \forall s, t, \in T, t_0$$

(6)

1.7.2 Estacionariedad débil o de Segundo Orden.

Este tipo de estacionariedad es mucho menos estricta que la anterior, ya que sólo se imponen condiciones al primer y segundo momentos del proceso. Así, se dice que un proceso $\{X_t ; t \in T\}$ es estacionario de segundo orden, si para cualquier conjunto de instantes $T = \{t_1, \dots, t_N\}$, cada N y cada desplazamiento t_0 , se satisfacen las siguientes condiciones:

a) Que la esperanza de las variables aleatorias X_t y $X_{t+\tau_0}$ exista y sea la misma para toda $t \in T$ y cada τ_0 , es decir,

$$E(X_t) = E(X_{t+\tau_0}) = \mu \quad ; \quad \forall t \in T, \tau_0 \quad (7)$$

Intuitivamente, esta propiedad se puede ver como si, a través del tiempo, la serie estuviera oscilando alrededor de un valor medio μ .

b) Que la autocovarianza entre dos variables aleatorias, sólo dependa de la diferencia que exista entre los dos instantes de tiempo, asociados a las variables y no a los instantes en particular.

$$\begin{aligned} \text{cov}(X_t, X_{t-j}) &= \text{cov}(X_{t+\tau_0}, X_{t+\tau_0-j}) \\ &= \text{cov}(\bar{X}_{t+\tau_0}, \bar{X}_{t+\tau_0-j}) \\ &= \gamma_j \quad ; \quad \forall t \in T, \tau_0, j \end{aligned} \quad (8)$$

Comúnmente en series de tiempo, se conoce a γ_j como la autocovarianza de orden j , el prefijo auto se debe a que γ_j es la covarianza entre diferentes variables aleatorias del mismo proceso, separadas j instantes de tiempo.

Obviamente por (5) y (6), se cumple que si un proceso es estrictamente estacionario, entonces, también es estacionario de segundo orden, sin embargo, el inverso no se cumple, a menos que la distribución conjunta de las variables aleatorias sea normal N-variada, ya que, en este caso el proceso quedaría completamente caracterizado por su primer y segundo momentos.

Ahora, si las condiciones (7) y (8) se satisfacen para toda $t \in T$, las expresiones en (1) y (2) se reducen considerablemente, ya que, el vector de medias para las variables aleatorias X_{t_1}, \dots, X_{t_N} , es

$$[E(X_{t_1}), \dots, E(X_{t_N})] = (\mu, \dots, \mu) \quad (9)$$

teniendo que estimar solamente un parámetro en lugar de N, como era el caso en (1).

La matriz simétrica de varianzas y covarianzas, es

$$\begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{N-1} \\ & \gamma_0 & \dots & \gamma_{N-2} \\ & & \ddots & \vdots \\ & & & \gamma_0 \end{pmatrix}$$

(10)

teniendo que estimar solamente N autocovarianzas en vez de $N(N+1)/2$, como era el caso en (2).

En adelante, cuando se hable de estacionariedad se supondrá estacionariedad de segundo orden, cuando este no sea el caso, se harán las especificaciones correspondientes.

1.8 Ergodicidad.

Al igual que con estacionariedad, ergodicidad es otra

propiedad que se pide al proceso generador de la serie de tiempo, para facilitar su análisis.

Intuitivamente, pedir que un proceso estocástico sea ergódico, es pedir que las variables aleatorias u observaciones de dicho proceso, considerablemente distantes en el tiempo, casi no esten correlacionadas. Así por ejemplo, al promediar una realización (serie) a través del tiempo, cada nuevo elemento en el promedio, aporta información útil y nueva. Lo que se pretende obtener con esta propiedad, son estimadores consistentes de los parámetros de interés del proceso, es decir, estimadores cuya varianza tienda a cero, conforme N tiende a infinito.

Aquí, no se explicará ergodicidad en términos matemáticos, ya que esto requiere de un desarrollo bastante complicado, innecesario para los objetivos de este capítulo (el lector interesado puede ver Hannan (1970, p. 201)).

En la práctica, no es posible probar si un proceso es o no ergódico, ya que sólo se observa una sola realización del mismo, es decir, una sola observación en cada variable aleatoria, y por lo tanto, no es posible probar si están o no correlacionadas. En adelante, se supondrá que todos los procesos considerados son ergódicos.

Por lo tanto, si la realización observada (serie de tiempo) proviene de un proceso estacionario y ergódico, ahora sí es posible hacer inferencias sobre el mismo (modelo

probabilístico), a partir, de la información proporcionada sólo por esa realización. Por ejemplo, el vector de medias en (9) se podría estimar como

$$(\mu, \dots, \mu) = \mu(1, \dots, 1) = \frac{1}{N} \sum_{t=\tau_1}^{\tau_N} x_t(1, \dots, 1)$$

donde x_t es la observación al tiempo t . Este estimador es equivalente al propuesto en (3), que se obtendría al observar varias realizaciones del proceso. En el primero, la media se estima como un promedio de las observaciones a través del tiempo; en el segundo, las medias se estiman como promedios de las observaciones en instantes particulares de tiempo. Obviamente, estos estimadores son equivalentes, si y sólo si, el proceso es ergódico y estacionario al menos de segundo orden, ya que entonces, la media, varianza y autocovarianzas del proceso no dependen del tiempo.

1.9 Descripción y Propiedades de Modelos ARMA(p,q).

Los modelos matemáticos más comúnmente usados en la práctica, para modelar series de tiempo estacionarias, son conocidos como AR, MA, y ARMA. Estos modelos pertenecen a la

clase general de procesos estocásticos lineales discretos, los cuales, se pueden expresar en dos formas equivalentes:

i) Dados, un proceso $\{X_t ; t \in T\}$, con $T = \{0, \pm 1, \pm 2, \dots\}$ y para el cual se supondrá, sin pérdida de generalidad, que tiene media igual a cero y varianza finita σ_x^2 , y un proceso $\{\varepsilon_t ; t \in T\}$ de variables aleatorias no correlacionadas, con media igual a cero y varianza finita σ_ε^2 , conocido como un proceso de ruido blanco. Entonces, el valor del proceso $\{X_t\}$ al tiempo t se puede expresar como:

$$X_t = \varepsilon_t + u_1 \varepsilon_{t-1} + u_2 \varepsilon_{t-2} + \dots \quad (11)$$

donde $\{u_i ; i=1, 2, \dots\}$ es una sucesión de parámetros fijos.

(11) se puede también expresar como

$$X_t = U(R)\varepsilon_t \quad (12)$$

donde $U(R) = \sum_{i=0}^{\infty} u_i \cdot R^i$, con $u_0 = 1$ y R el operador de atraso,

definido en la sección 1.4.

$U(R)$ se puede ver como la función de transferencia que transforma el proceso de ruido blanco $\{\xi_t\}$, en el valor de el proceso $\{X_t\}$ al tiempo t .

Para que el proceso en (12) sea estacionario, es necesario que su primer momento exista y no dependa del tiempo, y que su segundo momento exista y sólo dependa del período de tiempo entre dos variables cualesquiera del mismo proceso. Para lo cual, es necesario que se satisfagan las siguientes condiciones:

a) Para que la media del proceso $\{X_t\}$ exista, es necesario que la serie

$$\sum_{i=0}^{\infty} u_i < \infty$$

converja.

b) Para que la varianza del proceso $\{X_t\}$ sea finita, es necesario que la serie

$$\sum_{i=0}^{\infty} u_i^2 < \infty$$

converja.

c) Y para que la autocovarianza de orden j sea finita, es necesario que la serie

$$\sum_{i=0}^{\infty} u_i u_{i+j} < \infty$$

también converja.

Existe una sola condición impuesta a la serie de potencias $U(Z)$, donde Z toma valores complejos, la cual, comprende a todas las anteriores y garantiza la estacionariedad del proceso en (12). Esta condición es que $U(Z)$ converja en una región conteniendo al círculo unitario complejo, es decir, que converja para $|Z| \leq 1$. (Ver Grenander y Rosenblatt (1957)).

ii) Dados, un proceso $\{X_t\}$, y un proceso $\{\varepsilon_t\}$ como en (i), entonces, el valor del proceso $\{X_t\}$ al tiempo t , se puede expresar como

$$X_t = \varepsilon_t - v_1 X_{t-1} - v_2 X_{t-2} - \dots \quad (13)$$

donde v_i ; $i=1,2,\dots$ es una sucesión de parámetros fijos.

(13) se puede expresar también como

$$V(R) X_t = \varepsilon_t \quad \Rightarrow \quad X_t = [V(R)]^{-1} \varepsilon_t \quad (14)$$

donde $V(R) = \sum_{i=0}^{\infty} v_i R^i$, con $v_0 = 1$ y R el operador de atraso.

$[V(R)]^{-1}$ se puede ver como la función de transferencia que transforma al proceso de ruido blanco $\{\varepsilon_t\}$, en el valor del proceso $\{X_t\}$ al tiempo t .

De los procesos en (12) y (14), se puede ver que

$$U(Z) = 1/V(Z)$$

de ahí que las condiciones que debe cumplir $V(Z)$, donde Z toma valores complejos, para que el proceso en (14) sea estacionario, dependen de las condiciones dadas anteriormente para $U(Z)$. En este caso se necesita que los ceros de la ecuación característica $V(Z) = 0$, estén todos fuera de el círculo unitario complejo.

En adelante, cuando se haga referencia a un proceso $\{X_t\}$, se supondrá un proceso lineal discreto $\{X_t\}$, con $t = 0, \pm 1, \dots$, media cero y varianza finita σ_x^2 , y cuando se haga referencia a un proceso $\{\varepsilon_t\}$, se supondrá un proceso $\{\varepsilon_t\}$, con $t = 0, \pm 1, \dots$, de variables aleatorias no

correlacionadas con media cero y varianza finita σ_{ϵ}^2 .

1.10 Procesos Autorregresivos.

Un proceso autorregresivo de orden p , comúnmente denotado por $AR(p)$, es un proceso lineal discreto como en (13), pero donde sólo existe un número finito p de parámetros v_i distintos de cero. Sea $-v_i = a_i$, para $i = 1, \dots, p$, entonces un proceso $AR(p)$ queda expresado como

$$X_t = \epsilon_t + a_1 X_{t-1} + \dots + a_p X_{t-p} \quad (15)$$

X_t expresado como en (15) se puede observar como una variable aleatoria, la cual, esta linealmente relacionada con p valores pasados del mismo proceso, más un residual ó discrepancia aleatoria ϵ_t , semejando la relación observada en una ecuación de regresión, de ahí el nombre de autorregresivo.

Expresando a X_t en términos de la función de transferencia del proceso, se tiene

$$X_t = [A(R)]^{-1} \varepsilon_t \quad (16)$$

donde $A(R) = \sum_{i=0}^p a_i R^i$, con $a_0 = 1$ y R el operador de atraso.

Nuevamente, como se analizó en 1.9, para que el proceso en (16) sea estacionario se debe cumplir que $[A(Z)]^{-1}$, donde Z toma valores complejos, converja en una región conteniendo al círculo unitario complejo, ó equivalentemente, que los ceros de la ecuación característica $A(Z) = 0$ estén todos fuera del círculo unitario complejo.

1.11 Procesos de Promedio Móvil.

Un proceso de promedio móvil de orden q , comúnmente denotado por $MA(q)$, es un proceso lineal discreto como en (11), pero en donde, los parámetros u_i se anulan para toda i después de un cierto entero q . Sea $u_i = -b_i$, para $i = 1, \dots, q$, entonces, un proceso $MA(q)$ queda expresado como

$$X_t = \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q} \quad (17)$$

X_t expresado como en (17), no se puede observar como un promedio móvil tal cual, ya que los pesos b_i , para $i = 1, \dots, q$ no suman 1, pero sí se puede ver como un múltiplo de un promedio móvil, de q valores pasados del proceso de ruido blanco ξ_t .

Haciendo uso de la función de transferencia del proceso, X_t se puede expresar también como

$$X_t = B(R) \xi_t \quad (18)$$

donde $B(R) = \sum_{j=0}^q b_j R^j$, con $b_0 = 1$ y R el operador de atraso.

En este caso en que $B(Z)$, donde Z toma valores complejos, es finito, no hay restricciones de convergencia para que el proceso X_t sea estacionario. Sin embargo, (18) también se puede expresar como

$$[B(R)]^{-1} X_t = \xi_t \quad (19)$$

Comparando (19) con (14) se debe cumplir que $V(Z) = 1/B(Z)$, expandiendo en fracciones parciales esta última igualdad

$$V(Z) = 1/B(Z) = \sum_{i=1}^q \sum_{j=1}^{m_i} \frac{C_{ij}}{(1 - r_i Z)^j}$$

donde, $1/r_i$ son los ceros de la ecuación característica $B(Z) = 0$ y m_i es la multiplicidad del cero $1/r_i$. Así, para que $V(Z)$ converja en una región conteniendo al círculo unitario complejo es necesario que $|r_i| < 1$, es decir, que los ceros $1/r_i$ de $B(Z) = 0$ estén todos fuera del círculo unitario complejo. Si ésto se cumple se dice que el proceso en (19) es invertible.

Un proceso autorregresivo finito se puede expresar como un proceso de promedio móvil infinito, tal como se muestra en (16), si el proceso es estacionario. Un proceso de promedio móvil finito se puede expresar como un proceso autorregresivo infinito, tal como se muestra en (19), si el proceso es invertible.

1.12 Procesos Mixtos Autorregresivos y de Promedios Móviles.

Un proceso mixto autorregresivo y de promedio móvil, comúnmente denotado como ARMA(p,q), es como su nombre lo indica, un proceso doblemente finito, constituido por

términos autorregresivos y de promedios móviles

$$X_t = \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q} + a_1 X_{t-1} + \dots + a_p X_{t-p} \quad (20)$$

Este proceso también se puede expresar como

$$A(R) X_t = B(R) \varepsilon_t \quad (21)$$

donde, como antes, $A(R)$ y $B(R)$ son polinomios en el operador de atraso R , de ordenes p y q respectivamente.

Para que un proceso como (21) sea estacionario, se requiere que los ceros de la ecuación característica $A(Z) = 0$, estén todos fuera del círculo unitario complejo, y para que sea invertible, se requiere que los ceros de la ecuación característica $B(Z) = 0$ estén todos fuera de el círculo unitario complejo.

Capítulo 2

ESTIMACION DE PARAMETROS DEL MODELO ARMA(p,q) EN EL DOMINIO DEL TIEMPO

2.1 Métodos de Estimación.

Una vez que se ha elegido un modelo tentativo que ajuste a la serie de tiempo observada $x = (x_1, \dots, x_n)$, el siguiente paso es el de obtener estimadores de los parámetros involucrados en dicho modelo.

Supóngase que la serie de tiempo observada x , fue generada por un proceso gaussiano (ver apéndice A), estacionario e invertible ARMA(p,q)

$$A(R) X_t = B(R)\varepsilon_t$$

(1)

donde, $A(R)$ es un polinomio de orden p que satisface la condición de estacionariedad y $B(R)$ es un polinomio de orden q que satisface la condición de invertibilidad. Se supone además, que la sucesión de variables aleatorias X_1, \dots, X_N tienen media cero y matriz de varianzas y covarianzas Γ , con su (i, j) -ésimo elemento igual a la $\text{cov}(X_i, X_j) = \gamma_{|i-j|}$, es decir,

$$\Gamma = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{N-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{N-2} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \gamma_{N-1} & \gamma_{N-2} & \dots & \gamma_0 \end{pmatrix} \quad (2)$$

donde γ_k , la autocovarianza de orden k del proceso, es una función de los parámetros del modelo en (1), y $\{\xi_t\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas como normales, con media cero y varianza finita σ_ξ^2 .

En una situación tal, el método de estimación de máxima verosimilitud, se puede aplicar para estimar el vector de parámetros desconocido $\theta = (a_1, \dots, a_p, b_1, \dots, b_q, \sigma_\epsilon^2)$ en el modelo (1).

2.1.1 Verosimilitud No Condicional.

Dada la serie de tiempo observada x , y el problema de estimar el vector de parámetros θ , la función de verosimilitud $L(\theta; x)$ para el modelo en (1) expresado como

$$X_t = a_1 X_{t-1} + \dots + a_p X_{t-p} + \epsilon_t - b_1 \epsilon_{t-1} - \dots - b_q \epsilon_{t-q} \quad (3)$$

es

$$L(\theta; x) = \left(\frac{1}{2\pi}\right)^{\frac{N}{2}} |\Gamma|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} x^T \Gamma^{-1} x \right\} \quad (4)$$

y la función logaritmo de $L(\theta; x)$ es igual a

$$l(\theta; x) = -\frac{N}{2} \log(2\pi) + \frac{1}{2} \log |\Gamma| - \frac{1}{2} \{x^T \Gamma^{-1} x\} \quad (5)$$

el vector $\hat{\theta}$ buscado es el que maximiza $l(\theta ; x)$, es decir, que bajo condiciones de regularidad satisfaga la ecuación

$$\frac{dl(\theta ; x)}{d\theta} = \frac{1}{2} \frac{d \log |\Gamma|}{d\theta} - \frac{1}{2} \frac{d}{d\theta} (x^T \Gamma x) = 0 \quad (6)$$

donde, $\frac{dl(\theta ; x)}{d\theta}$ es un vector de dimensión $(p+q+1) \times 1$, de derivadas parciales de $l(\theta ; x)$ con respecto a cada uno de los elementos de θ .

Encontrar una expresión analítica general para el vector en (6), daría como resultado un desarrollo algebraico sumamente complicado por las siguientes razones:

a) La matriz Γ depende del vector de parámetros θ a través de las autocovarianzas $\gamma_{|i-j|}$ y éstas a su vez dependen de θ en forma no lineal.

b) La dimensión de la matriz Γ depende del tamaño de muestra observado N .

Primero, habría que encontrar expresiones analíticas

para las autocovarianzas en términos de los parámetros del modelo, para después, expresar a la matriz inversa de Γ y a su determinante en términos de éstas en (5). Una vez que se hubiese encontrado una expresión analítica general para $l(\theta; x)$ en (5), faltaría encontrar la derivada de esta función con respecto al vector de parámetros θ , lo cual, daría como resultado un vector de derivadas parciales con respecto a cada uno de los elementos de θ .

Es obvio que resolver el problema de estimación por este camino, parece ser una hazaña casi imposible de realizar. En este sentido, existen algunos trabajos publicados donde se da solución a algunos modelos en particular, más que una solución al modelo general ARMA(p,q). Entre estos trabajos se encuentra el de Tiao y Ali (1971), quienes dan expresiones explícitas para los elementos de la matriz inversa de covarianzas, a partir de la serie de tiempo observada, así como para su determinante en un proceso ARMA(1,1), ellos también dan un procedimiento recursivo para calcular la función de verosimilitud $L(\theta; x)$. A pesar de que solamente son tres los parámetros por estimar para este proceso, las expresiones que obtienen son complicadas.

Shaman (1973), hace una discusión sobre la inversión de la matriz de covarianzas muestral, para el caso más general ARMA(p,q), y llega a una expresión para los elementos de dicha matriz. También encuentra expresiones explícitas para la matriz inversa de covarianzas muestral, en un proceso

ARMA(0,2), sin embargo, no obtiene expresiones para su determinante. Las expresiones obtenidas por él son complicadas y necesitan de cálculos muy laboriosos.

Existen otros trabajos publicados donde se da solución al modelo general ARMA(p,q), Newbold (1974) da una expresión explícita para la función de verosimilitud de un proceso ARMA(p,q), en la que no es necesario el cálculo de la inversa de la matriz de autocovarianzas. Su método sigue la misma línea que el desarrollado por Box y Jenkins, el cual se verá más adelante bajo el nombre de enfoque no condicional, pero donde las esperanzas de ϵ_t condicionadas a la serie de tiempo observada x , se obtienen en forma explícita para cualquier vector θ dado, y donde el exponente de la función de verosimilitud, consiste de una suma finita de esperanzas condicionales, por lo que no es necesario usar una aproximación a esta suma, como en el caso del enfoque no condicional. Este procedimiento sin embargo, necesita del cálculo de matrices e inversas de matrices a lo largo del camino, que si bien tienen una estructura más sencilla que la de covarianzas, no dejan de ser largos y algo complicados.

En 1984 en un reporte técnico, Wincek y Reinsel desarrollan un método numérico basado en una transformación de los residuales, a partir de la cual, derivan un procedimiento recursivo para el cálculo eficiente de la función de verosimilitud, así como para las derivadas asociadas. Un estudio más detallado de este método, así

como la programación de un algoritmo, que permitiese hacer comparaciones con los estimadores obtenidos por otros métodos, podría ser conveniente, ya que éste parece ser el trabajo más completo y reciente en cuanto a la obtención de la función de verosimilitud exacta.

Un camino alternativo para obtener la función de verosimilitud $L(\theta; x)$, es haciendo uso de la condición de estacionariedad del modelo ARMA(p,q) en (1). Como se mencionó en el capítulo 1, cuando un proceso ARMA(p,q) es estacionario, éste se puede expresar como un proceso MA infinito, o sea que cada X_t depende de todas las ε 's anteriores hasta ε_t , y en consecuencia la información que sobre el proceso proporcionan las variables X_1, \dots, X_N , esta contenida en la información proporcionada por las variables $\varepsilon_{\infty}, \dots, \varepsilon_N$. Por lo tanto, el modelo en la ecuación (1) se puede expresar como

$$X_t = [A(R)]^{-1} B(R) \varepsilon_t = \sum_{\kappa=0}^{\infty} \psi_{\kappa} \varepsilon_{t-\kappa} \quad (7)$$

donde $\{\psi_{\kappa}\}$ son parámetros fijos que dependen del vector de parámetros desconocido θ .

La función de verosimilitud no condicional, asociada

con la serie de tiempo observada x , generada por el proceso ARMA(p,q) en (7) es (ver Box y Jenkins (1970) pag. 269-273):

$$L(\theta; x) = (2\pi)^{-\frac{N}{2}} (\sigma_\epsilon^2)^{-\frac{N}{2}} |M|^{-\frac{N}{2}} \exp\left\{-\frac{S(\theta^*)}{2\sigma_\epsilon^2}\right\} \quad (8)$$

donde $\sigma_\epsilon^2 \cdot M^{-1} = \Gamma$ definida en (2)

$$\theta^* = (a_1, \dots, a_p, b_1, \dots, b_q)$$

$$S(\theta^*) = \sum_{t=-\infty}^N E(\xi_t / x, \theta^*)^2 \quad (9)$$

y $E(\xi_t / x, \theta^*)$ es la esperanza de ξ_t condicionada a la serie observada x y a un vector de parámetros dado θ^* .

Box y Jenkins sugieren aproximar $S(\theta^*)$ en (9) por una suma finita, argumentando, que dado el carácter estacionario del operador autorregresivo $\hat{A}(R)$, se espera que las esperanzas de ξ_t condicionadas a la serie observada x decrezcan en magnitud, cuando t tiende a menos infinito, y en consecuencia puedan ser consideradas cero después de un

cierto punto en el tiempo, digamos $t = 1-Q$, para Q suficientemente grande. Por lo que (9) se puede aproximar por

$$S(\theta^*) \approx \sum_{t=1-Q}^N (E(\epsilon_t/x, \theta^*))^2 \quad (10)$$

La función logaritmo de $L(\theta; x)$, sin una constante y substituyendo (10) es

$$l(\theta; x) \approx -\frac{N}{2} \log(\sigma_\epsilon^2) + \frac{1}{2} \log |M| - \frac{1}{2\sigma_\epsilon^2} \sum_{t=1-Q}^N (E(\epsilon_t/x, \theta^*))^2 \quad (11)$$

El segundo término en el lado derecho de (11), es una función del vector de parámetros θ^* que no involucra a la serie de tiempo observada x , por lo que sería importante considerarla sólo para N pequeña, así, para tamaños de muestra razonables o grandes, este término puede ser omitido

y $l(\theta; x)$ en (11) estaría dominada por

$$l(\theta; x) \approx - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=1-a}^N (E(\varepsilon_t/x, \theta^*))^2 \quad (12)$$

Encontrar un estimador del vector de parámetros $\hat{\theta}$ que maximice la función en (12), es equivalente al problema de encontrar $\hat{\theta}^*$ que

$$\theta^* = (a_1, \dots, a_p, b_1, \dots, b_q) \underset{\text{MIN}}{\sum_{t=1-a}^N} (E(\varepsilon_t/x, \theta^*))^2 \quad (13)$$

y una vez que un estimador $\hat{\theta}^*$ ha sido obtenido, se obtiene el estimador de la varianza residual $\hat{\sigma}_\varepsilon^2$, derivando (11) con respecto a éste e igualando a cero, de donde

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1-a}^N (E(\varepsilon_t/x, \theta^*))^2}{N} \quad (14)$$

Se espera que los estimadores obtenidos por este camino, se aproximen lo suficiente a los estimadores máximo verosímiles.

Este desarrollo, sin embargo, presenta el problema práctico de cómo obtener las esperanzas de ξ_t condicionadas a la serie observada x . Box y Jenkins ofrecen una solución a este problema conocida como enfoque no condicional, el cual se describe a continuación.

2.1.2 Enfoque No Condicional.

Este tipo de enfoque consiste en estimar a las series de valores iniciales y desconocidos $x_I = (x_{1-p}, \dots, x_0)$ y $\xi_I = (\xi_{1-q}, \dots, \xi_0)$, necesarias para comenzar el proceso, a partir de la serie de tiempo observada x . Para llevar a cabo la estimación de estos valores iniciales, Box y Jenkins han desarrollado un método iterativo, el cual ha sido ampliamente usado y consiste en lo siguiente:

Se parte del hecho de que las dos siguientes representaciones de un proceso, tienen la misma estructura de covarianza, ya que la función generatriz de autocovarianzas es la misma para ambos modelos

$$A(R) X_t = B(R) \xi_t$$

(15)

$$A(R^{-1}) X_t = B(R^{-1}) \eta_t \quad (16)$$

donde, R en (15) es el operador de atraso, tal que $R^i X_t = X_{t-i}$, R^{-1} en (16) es el operador de adelanto, tal que $R^{-i} X_t = X_{t+i}$ y $\{\eta_t\}$ es una sucesión de variables aleatorias no correlacionadas con media cero y varianza finita σ_η^2 (las series η_t y ε_t no son necesariamente iguales aunque la varianza σ_η^2 de η_t es igual a la varianza σ_ε^2 de ε_t). Y por lo tanto, ambas representaciones describen igualmente bien a un proceso, (15) en términos del pasado, y (16) en términos del futuro del mismo.

Se asigna a los valores desconocidos de ε 's y η 's, valores iguales a sus esperanzas condicionadas en la serie de tiempo observada x .

En (15) se observa que los valores de $\varepsilon_{t+1}, \varepsilon_{t+2}, \dots$ no dependen de X_t, X_{t-1}, \dots , por lo que,

$$E(\varepsilon_t / x_1, \dots, x_N) = 0 \quad \text{para } t > N \quad (17)$$

y en (16) se observa que los valores de $\eta_{t-1}, \eta_{t-2}, \dots$ no dependen de X_t, X_{t+1}, \dots , por lo que,

$$E(\eta_t / x_1, \dots, x_N) = 0 \quad \text{para } t < 1 \quad (18)$$

por último, $E(X_t/x) = x_t$ para $t = 1, \dots, N$ (19)

El cálculo comienza con el pronóstico hacia atrás de la serie $\hat{X}_{1-p}, \dots, \hat{X}_0$, dado un vector inicial de parámetros θ^0 , para lo cual se hace uso de las esperanzas condicionales de la ecuación en (16), es decir,

$$\begin{aligned} E(X_t/x) - a_1 E(X_{t+1}/x) - \dots - a_p E(X_{t+p}/x) = \\ = E(\eta_t/x) - b_1 E(\eta_{t+1}/x) - \dots - b_q E(\eta_{t+q}/x) \end{aligned} \quad (20)$$

estimando valores para $\eta_{N-p}, \eta_{N-p-1}, \dots, \eta_1$. Aquí Box y Jenkins hacen la aproximación adicional

$$E(\eta_{N-p+j}/x) = 0 \quad \text{para } j = 1, \dots, q \quad (21)$$

argumentando, que para series de tiempo observadas suficientemente grandes, el efecto introducido por esta aproximación, va desapareciendo conforme nos acercamos al principio del periodo de tiempo observado. Por lo tanto, se genera la serie de valores $\hat{\eta}_{n-p}, \hat{\eta}_{n-p-1}, \dots, \hat{\eta}_1$ como:

$$E(\eta_{n-p} / x) = \hat{\eta}_{n-p} = x_{n-p} - a_1 x_{n-p+1} - \dots - a_p x_n$$

$$E(\eta_{n-p-1} / x) = \hat{\eta}_{n-p-1} = x_{n-p-1} - a_1 x_{n-p} - \dots - a_p x_{n-1} + b_1 \hat{\eta}_{n-p}$$

·
·
·

$$E(\eta_1 / x) = \hat{\eta}_1 = x_1 - a_1 x_2 - \dots - a_p x_{p+1} + b_1 \hat{\eta}_2 + \dots + b_q \hat{\eta}_{q+1}$$

y posteriormente se comienza el pronóstico hacia atrás de la serie $\hat{x}_{1-p}, \dots, \hat{x}_1, \hat{x}_0$ como:

$$E(X_0/x) = \hat{X}_0 = a_1 x_1 + \dots + a_p x_p - b_1 \hat{\eta}_1 - \dots - b_q \hat{\eta}_q$$

$$E(X_{-1}/x) = \hat{X}_{-1} = a_1 \hat{X}_0 + a_2 x_1 + \dots + a_p x_{p-1} - b_2 \hat{\eta}_1 - \dots - b_q \hat{\eta}_{q-1}$$

.

.

.

$$E(X_{1-p}/x) = \hat{X}_{1-p} = a_1 \hat{X}_{2-p} + a_2 \hat{X}_{3-p} + \dots + a_p x_1 - b_q \hat{\eta}_{1-p+q}$$

Observe que la sucesión $\{\eta_t\}$, no es otra, que la sucesión de errores de pronóstico a un paso hacia atrás con mínimo error cuadrático medio.

Box y Jenkins sugieren continuar con el pronóstico hacia atrás, hasta encontrar un valor \hat{X}_{1-q} aproximadamente cero, en el cual, se espera que la esperanza de ξ_t condicionada en la serie de tiempo observada x sea también aproximadamente cero, ya que el proceso es estacionario. Y en consecuencia, sea válido aplicar en este punto en el tiempo $(1-Q)$, la aproximación

$$E(\mathcal{E}_t / x_1, \dots, x_N) \approx 0 \quad \text{para } t < 1-Q \quad (22)$$

que permite aproximar la suma de cuadrados en (9) por la suma en (10).

El proceso termina con la estimación de valores para la serie $\hat{\mathcal{E}}_{1-Q}, \dots, \hat{\mathcal{E}}_0$, a partir de las esperanzas condicionales de la ecuación en (15), es decir,

$$\begin{aligned} E(X_t/x) - a_1 E(X_{t-1}/x) - \dots - a_p E(X_{t-p}/x) = \\ = E(\mathcal{E}_t/x) - b_1 E(\mathcal{E}_{t-1}/x) - \dots - b_q E(\mathcal{E}_{t-q}/x) \end{aligned} \quad (23)$$

La serie de valores $\hat{\mathcal{E}}_{1-a+1}, \dots, \hat{\mathcal{E}}_0$ se genera como:

$$E(\mathcal{E}_{1-a+1}/x) = \hat{\mathcal{E}}_{1-a+1} = \hat{X}_{1-a+1}$$

$$E(\mathcal{E}_{1-a+2}/x) = \hat{\mathcal{E}}_{1-a+2} = \hat{X}_{1-a+2} - a_1 \hat{X}_{1-a+1} + b_1 \hat{\mathcal{E}}_{1-a+1}$$

·
·
·

$$E(\mathcal{E}_0/x) = \hat{\mathcal{E}}_0 = \hat{X}_0 - a_1 \hat{X}_{-1} - \dots - a_p \hat{X}_{-p} + b_1 \hat{\mathcal{E}}_{-1} + \dots + b_q \hat{\mathcal{E}}_{-q}$$

En este momento del proceso, se termina con una iteración.

Es conveniente notar que si el tamaño de muestra observado es pequeño, puede ser necesario un número mayor de iteraciones, para obtener estimadores de $\hat{\mathcal{E}}_{1-a+1}, \dots, \hat{\mathcal{E}}_0$, que no dependan mucho de las aproximaciones hechas en (17), (18), (21) y (22).

Si un número mayor de iteraciones fuese necesario, entonces, se tendrían que obtener los estimadores $\hat{X}_{n+1}, \hat{X}_{n+2}, \dots$, a partir de la ecuación en (23), para después, obtener estimadores mejores a los dados inicialmente en la ecuación (21), a partir de la ecuación en (20).

El procedimiento continúa, hasta que la función a minimizar converja, es decir, hasta que la suma de cuadrados en (10) o (13) converja.

2.2 Verosimilitud Condicional.

Otra forma de resolver el problema de estimación, la cual es más sencilla aunque menos eficiente que la anterior, consiste en lo siguiente:

Conceptualmente, dada la serie de tiempo observada x , generada por el proceso ARMA(p,q) en (1), el cual, se puede expresar en términos de ξ_t como

$$\xi_t = x_t - a_1 x_{t-1} - \dots - a_p x_{t-p} + b_1 \xi_{t-1} + \dots + b_q \xi_{t-q} \quad (24)$$

y un vector de parámetros dado

$$\theta^* = (a_1, \dots, a_p, b_1, \dots, b_q),$$

puede ser posible obtener el conjunto de valores no observables ξ_1, \dots, ξ_n , a partir de (24), si las series de

valores iniciales $x_1 = (x_{1-p}, \dots, x_0)$ y $\xi_1 = (\xi_{1-q}, \dots, \xi_0)$ anteriores a la serie de tiempo observada x , fuesen conocidas.

Dado que la sucesión de variables aleatorias e independientes ξ_1, \dots, ξ_N , se distribuyen como normales con media cero y varianza finita σ_ξ^2 , su función de distribución conjunta estaría dada por

$$f(\xi_1, \dots, \xi_N) = (2\pi)^{-\frac{N}{2}} (\sigma_\xi^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma_\xi^2} \sum_{t=1}^N \xi_t^2 \right\}$$

y la función de verosimilitud asociada con el vector de parámetros $\theta = (a_1, \dots, a_p, b_1, \dots, b_q, \sigma_\xi^2)$, condicionada a la elección de las series x_1 y ξ_1 , sería

$$L(\theta ; x/x_1, \xi_1) = (2\pi)^{-\frac{N}{2}} (\sigma_\xi^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma_\xi^2} \sum_{t=1}^N \xi_t^2(\theta^*/x_1, \xi_1, x) \right\} \quad (25)$$

La función logaritmo de (25), sin una constante es

$$l(\theta ; x/x_1, \xi_1) \approx -\frac{N}{2} \log (\sigma_\xi^2) - \frac{1}{2\sigma_\xi^2} \sum_{t=1}^N \xi_t^2(\theta^*/x_1, \xi_1, x) \quad (26)$$

Encontrar un estimador $\hat{\theta}$ del vector de parámetros θ , que maximice la función en (26), es equivalente al problema de encontrar $\hat{\theta}^*$ que

$$\text{MIN}_{\hat{\theta}} \sum_{t=1}^N \xi_t^2(\theta^*/x_t, \varepsilon_t, x) \quad (27)$$

Y una vez que estos estimadores han sido obtenidos, se obtiene el estimador $\hat{\sigma}_\varepsilon^2$ de la varianza residual σ_ε^2 , derivando (26) con respecto a éste e igualando a cero, de donde

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^N \xi_t^2(\theta^*/x_t, \varepsilon_t, x)}{N} \quad (28)$$

por la propiedad de invarianza de los estimadores máximo verosímiles.

Sin embargo, este método de estimación presenta el

problema práctico, que las series de valores iniciales x_1 y ϵ_1 no son observables. Al respecto Box y Jenkins proponen una solución conocida como enfoque condicional, el cual consiste, en tratar a las dos series de valores iniciales, como constantes conocidas y asignarles un valor igual a su esperanza no condicional. Este método se describe en detalle a continuación.

2.2.1 Enfoque Condicional.

A partir de la serie de tiempo observada x , y de la representación del modelo en (24)

$$\epsilon_t = X_t - a_1 X_{t-1} - \dots - a_p X_{t-p} + b_1 \epsilon_{t-1} + \dots + b_q \epsilon_{t-q}$$

se pueden obtener los valores de la serie $\epsilon_1, \dots, \epsilon_N$, dado un vector inicial de parámetros θ^0 y las series de valores iniciales desconocidos x_1 y ϵ_1 . Box y Jenkins sugieren dar a estos valores desconocidos, un valor igual a su esperanza no condicional, es decir,

$$E(\epsilon_t) = 0 \quad \text{para } t = 0, \dots, 1-q \quad (29)$$

$$E(X_t) = 0 \quad \text{para } t = 0, \dots, 1-p$$

esto permite generar la serie $\varepsilon_1, \dots, \varepsilon_N$, para un vector de parámetros dado θ^0 , como

$$\varepsilon_1 = x_1$$

$$\varepsilon_2 = x_2 - a_1 x_1 + b_1 \varepsilon_1$$

.

.

.

$$\varepsilon_N = x_N - a_1 x_{N-1} - \dots - a_p x_{N-p} + b_1 \varepsilon_{N-1} + \dots + b_q \varepsilon_{N-q}$$

Sin embargo, si alguna de las raíces del polinomio autorregresivo $A(Z)$, está muy cercana al círculo unitario, la asignación de las condiciones iniciales dadas por (29), pueden ocasionar valores de las primeras ε_t 's con una fuerte dependencia en dichas condiciones, ya que existiría una alta correlación entre valores sucesivos del proceso. En consecuencia, si la serie extendida $x_e = (x_1, x_2, \varepsilon_1)$ fuese usada para calcular pronósticos para la serie observada x o para valores futuros del proceso, se obtendrían grandes errores de pronóstico.

Si se sospecha el comportamiento anterior para el operador autorregresivo, Box y Jenkins sugieren comenzar el cálculo de la serie de valores ε_t , al tiempo $t = p+1$, y así,

a partir de la serie de tiempo observada x y de un vector de parámetros θ^0 dado, se obtienen

$$\varepsilon_{p+1} = x_{p+1} - a_1 x_p - \dots - a_p x_1$$

$$\varepsilon_{p+2} = x_{p+2} - a_1 x_{p+1} - \dots - a_p x_2 + b_1 \varepsilon_{p+1}$$

.

.

.

$$\varepsilon_N = x_N - a_1 x_{N-1} - \dots - a_p x_{N-p} + b_1 \varepsilon_{N-1} + \dots + b_q \varepsilon_{N-q}$$

donde, $\varepsilon_{p+1-j} = E(\varepsilon_{p+1-j}) = 0$; para $j = 1, \dots, q$.

Es conveniente observar que en el caso de muestras pequeñas, esta última aproximación genera un número pequeño de valores ε_t , peor aún en el caso de procesos estacionales, donde el número de valores ε_t por calcular, sería igual a $N-p-ps$, donde ps es el orden del operador estacional.

Algunos otros autores como comenta Denis J. Aigner (1971), han desarrollado otros métodos iterativos para calcular los valores iniciales desconocidos. Entre ellos, Phillips, A. W. (1966) utiliza un método de pronóstico "parcial" hacia atrás, mientras que Pagan, A. (1970) utiliza

un procedimiento recursivo algo diferente, el cual, también suprime algunas observaciones a lo largo del camino. Basado en experimentos de Monte Carlo, Pagan observó que no había bases suficientes, para preferir usar alguno de los métodos desarrollados por él y por Phillips, al método condicional que es mucho más simple.

Existen otros problemas no mencionados hasta ahora, derivados del de estimación. Uno de ellos, se relaciona con el hecho de que en general ϵ_t es una función no lineal en los parámetros por estimar, cuando se le expresa a ésta en términos de la serie observada x , dado que la sucesión $\{\psi_k\}$ en (7), son parámetros fijos que dependen del vector de parámetros desconocido θ , por lo que, se hace necesario el uso de técnicas de regresión no lineal. El otro problema, se deriva del anterior y consiste en la buena elección de un vector inicial de parámetros θ^0 , necesario en cualquier método de estimación de parámetros en modelos no lineales.

2.3 No Linealidad en los Parámetros.

Una vez que se ha tomado alguna decisión, en cuanto a cómo resolver el problema de valores iniciales desconocidos, se está en posibilidad de decidir, cómo resolver el problema de no linealidad en los parámetros.

Existen varios métodos de estimación de parámetros en sistemas no lineales, ampliamente usados en varios campos de investigación, por lo que para muchos de ellos, ya existen algoritmos programados en computadora, que permiten obtener, en forma rápida y eficiente, los estimadores deseados. De entre éstos, el método de linealización o series de Taylor, generalmente se usa en la estimación de parámetros de modelos tales como (24). Este tipo de método utiliza los resultados de la técnica de mínimos cuadrados lineales, en una secuencia de pasos, y consiste en lo siguiente:

Se define un nuevo vector de parámetros, diferente del vector θ que se ha venido utilizando, como

$$C' = (C_1, C_2, \dots, C_{p+q}) = (a_1, \dots, a_p, b_1, \dots, b_q)$$

Sea $C'_0 = (C_1, 0, C_2, 0, \dots, C_{p+q}, 0)$, un vector de estimadores iniciales para los parámetros, obtenido a partir de un procedimiento que se describirá en detalle más adelante. Se espera que este estimador C'_0 , sea mejorado en cada una de las siguientes iteraciones.

Dado que la variable aleatoria ξ_t , depende de los parámetros en forma no lineal, se considera la expansión en series de Taylor de ξ_t alrededor del punto C'_0 , hasta el término que contiene a las primeras derivadas, es decir,

$$\xi_t \approx \xi_{t,0} - \sum_{i=1}^{p+q} \left[\frac{\partial \xi_t}{\partial C_i} \right]_{C=C_0} (C_i - C_{i,0}) \quad (30)$$

donde $\xi_{t,0}$, es el valor de ξ_t , que se obtiene al substituir C_0 en el modelo en (24).

Sea $Z_{t,i} = - \frac{\partial \xi_t}{\partial C_i} \Big|_{C=C_0}$ y $d_i^0 = C_i - C_{i,0}$

entonces, la ecuación en (30), queda como

$$\xi_t \approx \xi_{t,0} - \sum_{i=1}^{p+q} d_i^0 Z_{t,i}$$

Esta última ecuación, es una aproximación a la ecuación de mínimos cuadrados

$$\xi_{t,0} = \sum_{i=1}^{p+q} d_i^0 Z_{t,i} + \xi_t$$

donde d_i^0 ; con $i = 1, \dots, p+q$, se estiman aplicando la teoría

de mínimos cuadrados lineales, como

$$\hat{d}_0 = (Z'Z)^{-1} Z'\epsilon_0$$

donde,

$$Z = \begin{pmatrix} Z_{1,1} & Z_{1,2} & \dots & Z_{1,p+q} \\ Z_{2,1} & Z_{2,2} & \dots & Z_{2,p+q} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N,1} & Z_{N,2} & \dots & Z_{N,p+q} \end{pmatrix}; \quad \epsilon_0 = \begin{pmatrix} \epsilon_{1,0} \\ \epsilon_{2,0} \\ \vdots \\ \epsilon_{N,0} \end{pmatrix}; \quad \hat{d}_0 = \begin{pmatrix} \hat{d}_{1,0} \\ \hat{d}_{2,0} \\ \vdots \\ \hat{d}_{p+q,0} \end{pmatrix}$$

Por lo tanto, \hat{d}_0 minimiza la suma de cuadrados

$$\sum_{t=1}^N \left\{ \epsilon_{t,0} - \sum_{i=1}^{p+q} d_i^0 z_{t,i} \right\}^2$$

Para mejorar el estimador inicial C_0 , se utiliza al estimador \hat{d}_0 , como

$$C^{(1)}_0 = C_0 + \hat{d}_0 = C_0 + \widehat{C - C_0}$$

Se repiten todos los cálculos anteriores, pero ahora, reemplazando C_0 por $C^{(j)}$. El procedimiento iterativo continúa, hasta que la solución converge, es decir, hasta que se cumpla que

$$\left| \frac{C^{(j+1)}_{i,o} - C^{(j)}_{i,o}}{C^{(j)}_{i,o}} \right| < \delta \quad ; \quad \text{para } i = 1, \dots, p+q$$

para dos iteraciones sucesivas j y $j+1$, donde δ , es una precisión preestablecida.

Una vez que el procedimiento iterativo termina, se obtiene el estimador de la varianza residual σ_{ϵ}^2 , como

$$\hat{\sigma}_{\epsilon}^2 = \frac{\sum_{t=1}^N \epsilon_{t,o}^2}{N-p-q} \quad (31)$$

así como, el estimador de la matriz de varianzas y covarianzas V , de los parámetros estimados

$$\hat{V} = \hat{\sigma}_t^2 (Z'Z)^{-1}$$

donde $\xi_{t,0}$ y Z , son el vector y la matriz, obtenidos en la última iteración. Estos resultados se pueden extender, para el caso en que se decida utilizar un enfoque no condicional, simplemente reemplazando, ξ_t y $\xi_{t,0}$ por $E(\xi_t/x)$ y $E(\xi_{t,0}/x)$, y el rango de variación de t , sería de $1-Q$ a N , en vez de 1 a N . El estimador de la varianza residual, para este caso, sería

$$\hat{\sigma}_t^2 = \sum_{t=1-Q}^N \frac{E(\xi_{t,0}/x)^2}{N-p-q} \quad (32)$$

Este método, sin embargo, presenta el problema práctico de obtener una expresión analítica para el gradiente de ξ_t en (30), por lo que puede ser necesario obtenerlo numéricamente, mediante una aproximación que use diferencias finitas.

El gradiente de ξ_t , esta dado por

$$\left[\frac{\partial \xi_t}{\partial C_i} \right]_{C=C_0} = [Z_{t,i}]_{C=C_0}$$

Sea

$$Z_{t,i} \approx \frac{\epsilon_t(C_1, 0, \dots, C_i, 0 + \delta_i, \dots, C_{p+q}, 0) - \epsilon_t(C_1, 0, \dots, C_i, 0, \dots, C_{p+q}, 0)}{\delta_i}$$

donde, $i = 1, \dots, p+q$ y δ_i es suficientemente pequeño, tal que, el límite cuando δ_i tiende a cero es por definición la derivada parcial $Z_{t,i}$.

Para una descripción general, sobre los métodos alternativos de estimación de parámetros en modelos no lineales, se puede consultar Draper y Smith (1981), segunda edición, capítulo 10, y para una descripción más detallada, Yonathan Bard (1974), Nonlinear Parameter Estimation.

En todos los métodos de estimación de parámetros en modelos no lineales, es necesario contar con un vector de valores iniciales para los parámetros, la elección de este vector determina la diferencia entre éxito o fracaso, en cuanto a la localización de un máximo o un mínimo de la función objetivo, según sea el caso, o entre la rapidez o lentitud de la convergencia a la solución buscada. De ahí que sea conveniente, una buena elección del vector inicial de parámetros

$$\hat{\theta}^0 = (a_1^0, \dots, a_p^0, b_1^0, \dots, b_q^0, \sigma_{\epsilon}^{02})$$

Box y Jenkins proponen un método, para estimar valores iniciales de los parámetros en un modelo ARMA(p,q), a partir de las autocovarianzas muestrales del mismo, el cual, se describe en la siguiente sección.

2.4 Método General para obtener Estimadores Iniciales de los Parámetros de un Proceso ARMA(p,q).

Sea $\{X_t\}$ un proceso ARMA(p,q) estacionario e invertible, con media cero, tal que

$$X_t - a_1 X_{t-1} - \dots - a_p X_{t-p} = \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q}$$

$$A(R) X_t = B(R) \varepsilon_t$$

Para obtener los estimadores iniciales de los parámetros de dicho modelo, es necesario conocer la sucesión de autocovarianzas muestrales $\hat{\gamma}_0, \dots, \hat{\gamma}_{p+q}$ obtenidas a partir de la serie observada $x = (x_1, \dots, x_N)$ como

$$\hat{\gamma}_k = \frac{1}{N} \sum_{t=1}^{N-k} x_t x_{t+k}$$

así como los ordenes p y q del modelo.

La estimación se lleva a cabo en tres pasos, los cuales se explican a continuación.

a) Estimadores iniciales de los parámetros autorregresivos.

Dado el modelo
$$X_t = \sum_{i=1}^p a_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q b_j \varepsilon_{t-j}$$

se tiene que:

$$\begin{aligned} \gamma_{X_t}^k &= E(X_{t-k} X_t) = \sum_{i=1}^p a_i E(X_{t-k} X_{t-i}) + E(X_{t-k} \varepsilon_t) - \sum_{j=1}^q b_j E(X_{t-k} \varepsilon_{t-j}) \\ &= \sum_{i=1}^p a_i \gamma_{X_t}^{k-i} + \gamma_{X_t}^k(k) - \sum_{j=1}^q b_j \gamma_{X_t}^k(k-j) \end{aligned} \quad (33)$$

donde
$$\gamma_{X_t}^k(k) = E(X_{t-k} \varepsilon_t) = \begin{cases} 0 & \text{si } k > 0 \\ \neq 0 & \text{si } k \leq 0 \end{cases}$$

por lo tanto,

$$Y_k = \sum_{i=1}^p a_i Y_{k-i} \quad \text{para } k \geq q+1 \quad (34)$$

Estas ecuaciones se conocen como las ecuaciones de Yule Walker, y se pueden expresar en forma equivalente como

$$A(R) Y_k = 0 \quad \text{para } k \geq q+1 .$$

Substituyendo los estimadores de $Y_{q-p+1}, \dots, Y_{q+1}, \dots, Y_{q+p}$ en las ecuaciones en (34), se obtiene el siguiente sistema

$$\hat{\gamma}_{q+1} = a_1 \hat{\gamma}_q + a_2 \hat{\gamma}_{q-1} + \dots + a_p \hat{\gamma}_{q+1-p}$$

·
·
·

$$\hat{\gamma}_{q+p} = a_1 \hat{\gamma}_{q+p-1} + a_2 \hat{\gamma}_{q+p-2} + \dots + a_p \hat{\gamma}_q$$

este sistema se puede expresar en forma matricial como

$$\begin{pmatrix} \hat{\gamma}_q & \hat{\gamma}_{q-1} & \dots & \hat{\gamma}_{q+1-p} \\ \hat{\gamma}_{q+1} & \hat{\gamma}_q & \dots & \hat{\gamma}_{q-p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \hat{\gamma}_{q+p} & \hat{\gamma}_{q+p-1} & \dots & \hat{\gamma}_q \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} \hat{\gamma}_{q+1} \\ \hat{\gamma}_{q+2} \\ \cdot \\ \cdot \\ \cdot \\ \hat{\gamma}_{q+p} \end{pmatrix} \quad (35)$$

$$A X = B$$

y resolviendo para X , se obtienen estimadores iniciales de los parámetros a_1, \dots, a_p de la parte autorregresiva del modelo.

b) Dado que en (33) ν_0, \dots, ν_q dependen de la elección de los parámetros a_1, \dots, a_p , y b_1, \dots, b_q , el paso siguiente es definir una nueva serie X_t^* tal que

$$A(R) * X_t = X_t^* \approx B(R) \varepsilon_t$$

donde los coeficientes de $A^*(Z)$ son $\hat{a}_1, \dots, \hat{a}_p$ obtenidos de (35).

Tratando a X_t^* como un proceso MA(q) aproximado, se pretende encontrar las primeras $q+1$ autocovarianzas $\hat{\nu}_0^*, \dots, \hat{\nu}_q^*$ para este proceso, y a partir de sus propiedades, obtener estimadores iniciales de los parámetros de promedios móviles b_1, \dots, b_q .

Lo primero que se necesita, es expresar a las autocovarianzas estimadas $\hat{\nu}_0^*, \dots, \hat{\nu}_q^*$ del proceso X_t^* , en términos de las autocovarianzas estimadas del proceso X_t obtenidas de (35).

Se sabe que (ver Box y Jenkins (1970), capítulo 6):

$$\hat{\nu}_k^* = \sum_{i=0}^p \hat{a}_i^* \hat{\nu}_k + \sum_{i=1}^p (\hat{a}_0 \hat{a}_i + \hat{a}_1 \hat{a}_{i+1} + \dots + \hat{a}_{p-i} \hat{a}_p) d_k \quad (37)$$

donde,
$$d_k = \hat{\nu}_{k+1} + \hat{\nu}_{k-1}$$

$$\hat{\alpha}_0 = -1$$

$$k = 0, \dots, q$$

c) Estimadores Iniciales de los Parámetros de Promedios Móviles.

Dadas las autocovarianzas estimadas $\hat{\nu}_0^*, \dots, \hat{\nu}_q^*$ del proceso $X_t^* \approx B(R)\xi_t$, obtenidas en (37), se pueden obtener estimadores iniciales de los parámetros de promedios móviles, por alguno de los dos siguientes métodos iterativos.

c.1) Método de Convergencia Lineal.

Las autocovarianzas para un proceso MA(q) se definen como

$$\nu_k^* = E(X_{t-k}^* X_t^*) = \sigma_\xi^2 \sum_{j=0}^{q-|k|} b_j b_{j+k} \quad \text{con } b_0 = -1$$

$$\text{para } 0 \leq k \leq q$$

$$\text{y } \nu_k^* = 0 \quad \text{para } k > q$$

Substituyendo $\hat{\nu}_k^*$ por ν_k^* se tiene que

$$\hat{\psi}_0^* = \sigma_\epsilon^2 (1 + b_1^2 + \dots + b_q^2) \quad \text{de donde}$$

$$\sigma_\epsilon^2 = \frac{\hat{\psi}_0^*}{1 + b_1^2 + \dots + b_q^2} \quad \text{y} \quad (38)$$

$$\hat{\psi}_\kappa^* = \sigma_\epsilon^2 (-b_\kappa + b_1 b_{\kappa+1} + \dots + b_{q-|\kappa|} b_q) \quad \text{de donde}$$

$$b_\kappa = - \left(\frac{\hat{\psi}_\kappa^*}{\sigma_\epsilon^2} - b_1 b_{\kappa+1} - \dots - b_{q-|\kappa|} b_q \right) \quad (39)$$

Como la relación entre los parámetros no es lineal, se tienen que resolver las ecuaciones (38) y (39) para σ_ϵ^2 y b_κ en forma iterativa, obteniendo estimadores iniciales de los parámetros $\sigma_\epsilon^2, b_q, b_{q-1}, \dots, b_1$, en este mismo orden.

Al iniciar el proceso iterativo se les da el valor inicial de cero a los parámetros b_1, \dots, b_q .

c.2) Método de Convergencia Cuadrática.

Wilson (1969) propuso un método usando el algoritmo de Newton-Raphson para obtener estimadores iniciales de los parámetros de promedios móviles b_1, \dots, b_q , y de la varianza residual σ_ϵ^2 , el cual, tiene propiedades de convergencia superiores al método anterior en c.1.

Sea $\tau' = (\tau_0, \tau_1, \dots, \tau_q)$ donde

$$\tau_0^2 = \sigma_\epsilon^2 \quad \text{y} \quad b_j = -\frac{\tau_j}{\tau_0} \quad j=1, \dots, q$$

(40)

Si $\tau^{(i)}$ es el estimador de τ obtenido en la i -ésima iteración, entonces el nuevo valor de $\tau^{(i+1)}$ en la $(i+1)$ -ésima iteración se obtiene como

$$\tau^{(i+1)} = \tau^{(i)} - (T^{(i)})^{-1} f^{(i)}$$

donde $f' = (f_0, \dots, f_q)$; $f_i = \sum_{j=1}^q \tau_j \tau_{i+j} - \hat{\beta}_i^*$ y

$$T = \begin{pmatrix} \tau_0 & \tau_1 & \dots & \tau_{q-2} & \tau_{q-1} & \tau_q \\ \tau_1 & \tau_2 & \dots & \tau_{q-1} & \tau_q & 0 \\ \tau_2 & \tau_3 & \dots & \tau_q & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \tau_q & 0 & \dots & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} \tau_0 & \tau_1 & \dots & \tau_{q-2} & \tau_{q-1} & \tau_q \\ 0 & \tau_0 & \dots & \tau_{q-3} & \tau_{q-2} & \tau_{q-1} \\ 0 & 0 & \dots & \tau_{q-4} & \tau_{q-3} & \tau_{q-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \tau_0 \end{pmatrix}$$

Conociendo el valor de τ en cada iteración, se pueden obtener los valores de los parámetros en la ecuación (40). Al iniciar el proceso iterativo se pueden dar los siguientes valores iniciales a los parámetros

$$a_t^2 = \hat{\gamma}_0^* \quad y \quad b_1 = b_2 = \dots = b_q = 0$$

el procedimiento continúa hasta que se obtiene una convergencia satisfactoria.

2.5 El Método de Mínimos Cuadrados No Lineales a través de las Autocovarianzas.

Un método alternativo de estimación a los descritos en 2.1 se desarrolla aquí. El método consiste en obtener estimadores del vector de parámetros θ para un modelo como en la ecuación (1), a través del ajuste de las autocovarianzas teóricas, expresadas en términos de los parámetros por estimar, a las autocovarianzas muestrales por mínimos cuadrados no lineales.

En la sección 2.2.1 se describen las propiedades que se deben pedir a un modelo ARMA(p,q) para asegurar que la identificación de dicho modelo a través de las

autocovarianzas sea única.

En la sección 2.2.2 se describen las propiedades que poseen los estimadores de las autocovarianzas utilizados.

En la sección 2.2.3 se describe el ajuste de las autocovarianzas teóricas, expresadas en términos de los parámetros, a los estimadores de las autocovarianzas por el método de mínimos cuadrados no lineales, así como sus propiedades.

2.5.1 Estimadores de los Parámetros de un Modelo ARMA(p,q) a través de las Autocovarianzas.

Se sabe que para un modelo ARMA(p,q) dado, éste posee una estructura de covarianza única, sin embargo dada una estructura de covarianza, existen varios modelos que pueden estar asociados a ella. Aparentemente este hecho hace difícil la posible identificación de un modelo a partir del conocimiento de la función de covarianza, pero, como se verá en esta sección, se puede pedir que el modelo satisfaga ciertas condiciones, de tal forma que la función de covarianza pueda usarse en forma lógica para identificarlo.

a) Invertibilidad.

Suponga que el proceso X_t , cuya función generatriz de autocovarianzas es $G_x(Z)$, se representa por el modelo lineal

y estacionario

$$A(R)X_t = B(R)\varepsilon_t \quad (41)$$

Este modelo también se puede expresar en términos de las raíces de los polinomios $A(Z)$ y $B(Z)$, como

$$\prod_{i=1}^p (1 - \alpha_i Z) X_t = \prod_{j=1}^q (1 - \beta_j Z) \varepsilon_t \quad (42)$$

donde α_i^{-1} , con $i = 1, \dots, p$ son las raíces del polinomio característico $A(Z) = 0$, las cuales se encuentran todas fuera del círculo unitario complejo, y β_j^{-1} son las raíces del polinomio característico $B(Z) = 0$.

Se sabe que la función generatriz de autocovarianzas de un proceso tal; es (ver apéndice A)

$$\begin{aligned} G_x(Z) &= \frac{\sigma_\varepsilon^2 B(Z)B(Z^{-1})}{A(Z)A(Z^{-1})} = \\ &= \sigma_\varepsilon^2 \frac{\prod_{j=1}^q (1 - \beta_j Z) \prod_{j=1}^q (1 - \beta_j Z^{-1})}{\prod_{i=1}^p (1 - \alpha_i Z) \prod_{i=1}^p (1 - \alpha_i Z^{-1})} = \end{aligned} \quad (43)$$

$$= \sigma_{\varepsilon}^2 \frac{\prod_{j=1}^q \beta_j^2 (1 - \beta_j^{-1} Z)(1 - \beta_j^{-1} Z^{-1})}{\prod_{i=1}^p (1 - \alpha_i Z)(1 - \alpha_i Z^{-1})}$$

o sea, que los dos modelos siguientes

$$\prod_{i=1}^p (1 - \alpha_i Z) X_t = \prod_{j=1}^q (1 - \beta_j Z) \varepsilon_t$$

$$\prod_{i=1}^p (1 - \alpha_i Z) X_t = \prod_{j=1}^q k_j (1 - \beta_j^{-1} Z) \varepsilon_t$$

tienen la misma función generatriz de autocovarianza $G_x(Z)$, donde k_j ; $j = 1, \dots, q$, son constantes escogidas de manera conveniente.

De manera general, los modelos

$$\prod_{i=1}^p (1 - \alpha_i Z) X_t = k \prod_{j=1}^q (1 - \beta_j^{\pm 1} Z) \varepsilon_t \quad (44)$$

poseen la misma $G_x(Z)$, produciéndose así una elección

múltiple de los parámetros de promedios móviles.

Analizando el comportamiento de las raíces β_j ; $j = 1, \dots, q$ en (44), se observa que si alguna β_j es real y se encuentra dentro del círculo unitario complejo, entonces su inversa β_j^{-1} se encontrará fuera, y en el caso de alguna raíz compleja dentro del círculo unitario complejo, su inversa β_k^{-1} y la raíz inversa conjugada $\bar{\beta}_k^{-1}$ se encontrarán fuera. Por lo tanto, si adicional a la condición de estacionariedad del modelo lineal en (41), se pide la condición de invertibilidad, se estará garantizando que el modelo lineal que se obtenga, tendrá todas sus raíces β_j^{-1} ; $j = 1, \dots, q$ en (42), fuera del círculo unitario complejo, y así la elección de los parámetros de promedios móviles será única.

b) Representación de X_t en términos del Pasado Histórico del Proceso.

A partir de (43) se puede observar que los siguientes modelos

$$\prod_{i=1}^p (1 - \alpha_i Z^{-1}) X_t = \prod_{j=1}^q (1 - \beta_j Z^{-1}) \epsilon_t$$

$$\prod_{i=1}^p (1 - \alpha_i Z^{-1}) X_t = \prod_{j=1}^q (1 - \beta_j Z) \epsilon_t$$

$$\prod_{i=1}^p (1 - \alpha_i Z) X_t = \prod_{j=1}^q (1 - \beta_j Z^{-1}) \varepsilon_t$$

incluyendo al modelo en (42), están asociados a la misma función generatriz de autocovarianzas $G_x(Z)$. De manera general, se puede decir que todos los modelos de la forma

$$\prod_{i=1}^p (1 - \alpha_i Z^{\pm 1}) X_t = \prod_{j=1}^q (1 - \beta_j Z^{\pm 1}) \varepsilon_t$$

poseen la misma $G_x(Z)$, sin embargo el modelo en (42) es el único que expresa el valor del proceso al tiempo t , X_t , exclusivamente en términos del pasado histórico del mismo, ya que la variable X_t depende de las variables X_{t-p}, \dots, X_{t-1} y $\varepsilon_{t-q}, \dots, \varepsilon_t$. Por lo que, si a las condiciones de estacionariedad e invertibilidad del modelo lineal en (41), se agrega la condición de que el valor del proceso al tiempo t , quede expresado sólo en términos del pasado histórico del mismo, se sabrá con certeza, al momento de identificar un modelo a partir de las autocovarianzas, a cuál de entre todos se está refiriendo.

c) No Factores en común.

Una forma aún más general del modelo en (42), podría ser

$$(1 - CZ) \prod_{i=1}^p (1 - \alpha_i Z) X_t = (1 - CZ) \prod_{j=1}^q (1 - \beta_j Z) \epsilon_t$$

en donde $(1-CZ)$ puede ser cualquier factor. Toda esta familia de procesos lineales, estacionarios e invertibles esta asociada a la misma $Gx(Z)$. Este hecho no debiese causar problemas, al momento de identificar a un modelo a partir de las autocovarianzas, ya que se debe optar siempre por el modelo más sencillo, sin embargo, es necesario a la hora de ajustar un modelo, estar seguros de que no existan factores en común entre los operadores.

En conclusión, si $\{X_t\}$ es un proceso estacionario e invertible, el cual no tiene factores en común, y donde el valor del proceso al tiempo t se expresa exclusivamente en términos del pasado histórico del mismo, entonces este proceso es determinado en forma única por su estructura de covarianza.

2.5.2 Propiedades de los Estimadores de las Autocovarianzas.

La información proporcionada por una serie de tiempo observada, la cual se supone fue generada por un proceso ARMA(p,q) estacionario, invertible y gaussiano (ver apéndice A), con media igual a cero, se puede resumir en la sucesión de autocovarianzas muestrales.

Dada una serie de tiempo observada $x = (x_1, \dots, x_N)$ del proceso $\{X_t\}$, la autocovarianza de orden k del proceso se define como

$$\text{cov}(X_t, X_{t+k}) = E(X_t X_{t+k}) = \gamma_k \quad \text{para } k = 0, \dots, N-1$$

Los dos tipos de estimadores de γ_k más usuales son:

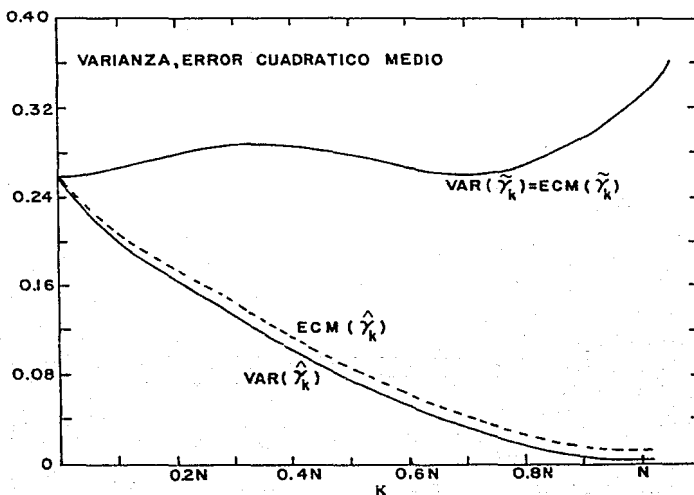
$$\begin{aligned} \tilde{\gamma}_k &= \frac{1}{N-k} \sum_{t=1}^{N-k} X_t \cdot X_{t+k} && \text{para } k = 0, \dots, N-1 \\ \hat{\gamma}_k &= \frac{1}{N} \sum_{t=1}^{N-k} X_t \cdot X_{t+k} && (45) \end{aligned}$$

y las esperanzas de estos estimadores son

$$E(\tilde{\gamma}_k) = \begin{cases} \gamma_k & \text{para } |k| \leq N \\ 0 & \text{para } |k| > N \end{cases} \quad (46)$$

$$E(\hat{\gamma}_k) = \begin{cases} \frac{N-|k|}{N} \gamma_k = \left(1 - \frac{|k|}{N}\right) \gamma_k & \text{para } |k| \leq N \\ 0 & \text{para } |k| > N \end{cases}$$

Este último tiene la desventaja de ser sesgado, sin embargo tiene una propiedad superior a la de $\tilde{\gamma}_k$. Jenkins y Watts (1968) comparan el error cuadrático medio para estos dos estimadores, con un proceso autorregresivo de primer orden, y observan que el error cuadrático medio de $\tilde{\gamma}_k$, es consistentemente más grande que el de $\hat{\gamma}_k$, como se observa en la gráfica 2.1.



Gráfica 2.1: Varianzas y errores cuadráticos medios de los estimadores de la función de autocovarianzas para un proceso continuo de primer orden.

Así mismo, observan que la varianza de \hat{Y}_k tiende a cero cuando k tiende a N , mientras que la varianza de \tilde{Y}_k tiende a infinito.

Las varianzas y covarianzas para estos estimadores son (ver Jenkins y Watts (1968))

$$\text{Cov}(\hat{Y}_k, \hat{Y}_s) = \frac{1}{N^2} \sum_{r=-(N-k)}^{N-s} \Phi(r) \{ Y_r Y_{r+s-k} + Y_{r+s} Y_{r-k} \} \quad (47)$$

$$\text{Cov}(\tilde{Y}_k, \tilde{Y}_s) = \frac{1}{(N-k)(N-s)} \sum_{r=-(N-k)}^{N-s} \Phi(r) \{ Y_r Y_{r+s-k} + Y_{r+s} Y_{r-k} \}$$

donde

$$\Phi(r) = \begin{cases} N-s-r & \text{para } r \geq 0 \\ N-s & \text{para } -(s-k) \leq r \leq 0 \\ N-k+r & \text{para } -(N-k) \leq r \leq -(s-k) \end{cases}$$

y cuando $k = s$

$$\text{Var}(\hat{p}_k) = \frac{1}{N^2} \sum_{r=-(N-k)}^{N-k} (N-k-|r|) \{ p_r p_r + p_{r+k} p_{r-k} \} \quad (48)$$

$$\text{Var}(\tilde{p}_k) = \frac{1}{(N-k)^2} \sum_{r=-(N-k)}^{N-k} (N-k-|r|) \{ p_r p_r + p_{r+k} p_{r-k} \}$$

A partir de (46), se puede observar que ambos estimadores son insesgados cuando N tiende a infinito, y a partir de (48), que sus varianzas tienden a cero cuando N tiende a infinito, por lo que ambos estimadores son consistentes asintóticamente.

En cuanto a su distribución asintótica, Anderson (1971) demuestra que, para n fija, los vectores

$$\sqrt{N} (\hat{p}_0 - p_0), \dots, \sqrt{N} (\hat{p}_n - p_n)$$

$$\sqrt{N} (\tilde{p}_0 - p_0), \dots, \sqrt{N} (\tilde{p}_n - p_n)$$

y

tienen una distribución normal n -variada, con vector de medias cero y matriz de varianzas y covarianzas Ω , cuyos elementos estan dados por

$$\begin{aligned}
\Omega_{\kappa, \xi} &= \lim_{N \rightarrow \infty} N \text{Cov}(\hat{Y}_{\kappa}, \hat{Y}_{\xi}) = \lim_{N \rightarrow \infty} N \text{Cov}(\tilde{Y}_{\kappa}, \tilde{Y}_{\xi}) \\
&= \sum_{r=-\infty}^{\infty} (\gamma_{r+\xi} \gamma_{r-\xi} + \gamma_{r-\xi} \gamma_{r+\xi})
\end{aligned}
\tag{49}$$

Aunque asintóticamente ambos estimadores se comportan igual, se utilizó el estimador sesgado \hat{Y}_{κ} , ya que como se mencionó antes, éste tiene un error cuadrático medio menor.

2.5.3 Ajuste de las Autocovarianzas por Mínimos Cuadrados No Lineales.

Como se mencionó en las dos secciones anteriores, cuando se cuenta con una serie de tiempo observada $x = (x_1, \dots, x_N)$, la cual se supone fue generada por un proceso gaussiano ARMA(p,q), estacionario e invertible, con media cero, donde el valor del proceso al tiempo t se expresa como

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q b_j \varepsilon_{t-j}
\tag{50}$$

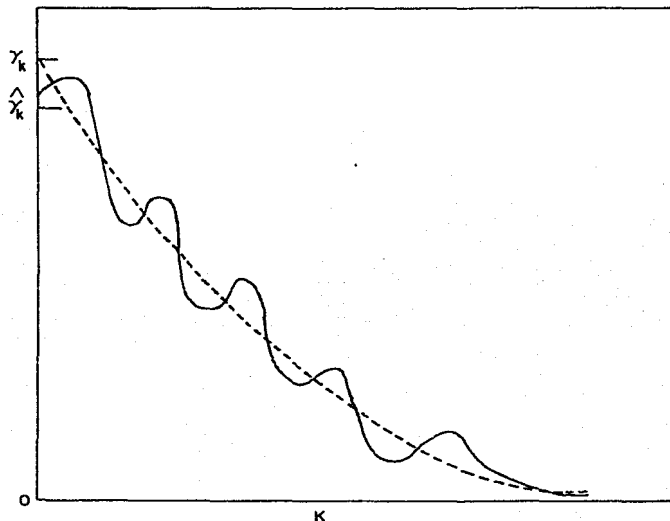
es posible resumir la información proporcionada por élla en la sucesión de autocovarianzas muestrales

$$\hat{\gamma} = (\hat{\gamma}_0, \dots, \hat{\gamma}_k)$$

donde k es un entero positivo menor que $N-1$, fijado de manera conveniente, y al cual nos referiremos en el capítulo 3 como lag. También se mencionó que a partir de la estructura de covarianza del proceso, se puede identificar de forma única al modelo, es decir que a través de las autocovarianzas parece posible estimar el vector de parámetros de interés

$$\theta = (a_1, \dots, a_p, b_1, \dots, b_q, \sigma_t^2)$$

En la práctica, se espera que la sucesión de autocovarianzas muestrales tengan un comportamiento parecido al de la sucesión de autocovarianzas teóricas para un proceso en particular. Por ejemplo, en la gráfica 2.2 se observa el comportamiento típico de la sucesión de autocovarianzas teóricas para un proceso $AR(1)$, cuyo parámetro a_1 es positivo, así como el comportamiento esperado de la sucesión de autocovarianzas muestrales para el mismo proceso.



Gráfica 2.2: Autocovarianzas teóricas de un proceso AR(1), y comportamiento esperado de sus autocovarianzas muestrales.

Como se puede observar, lo que se espera es que los valores de la sucesión de autocovarianzas muestrales, fluctuen alrededor de los valores de la sucesión de autocovarianzas teóricas.

Por otro lado se sabe (ver apéndice A), que la función generatriz de autocovarianzas de un proceso ARMA(p,q), relaciona sus autocovarianzas teóricas con el vector de parámetros θ , mediante la igualdad

$$G_x(Z) = \sum_{\kappa=-\infty}^{\infty} \gamma_{\kappa} Z^{\kappa} = \frac{\sigma_{\epsilon}^2 B(Z)B(Z^{-1})}{A(Z)A(Z^{-1})} \quad (51)$$

De esta igualdad se puede observar que las autocovarianzas teóricas son funciones no lineales del vector de parámetros θ .

Un método posible de estimar el vector de parámetros θ , puede ser mediante el ajuste de la sucesión de autocovarianzas teóricas, expresadas en términos de los elementos de θ , a la sucesión de autocovarianzas muestrales, como

$$\hat{\gamma}_0 = E(\hat{\gamma}_0) + \eta_0 = \gamma_0(\theta) + \eta_0$$

$$\hat{\gamma}_1 = E(\hat{\gamma}_1) + \eta_1 = \frac{N-1}{N} \gamma_1(\theta) + \eta_1$$

$$\begin{array}{ccc} \cdot & & \cdot \\ \vdots & & \vdots \\ \cdot & & \cdot \end{array}$$

$$\hat{\gamma}_{\kappa} = E(\hat{\gamma}_{\kappa}) + \eta_{\kappa} = \frac{N-|k|}{N} \gamma_{\kappa}(\theta) + \eta_{\kappa}$$

(52)

por mínimos cuadrados no lineales, es decir, el problema ahora sería

$$\text{MIN}_{\theta} \left(\hat{\gamma} - C \gamma(\theta) \right)^T I \left(\hat{\gamma} - C \gamma(\theta) \right) \quad (53)$$

donde $\hat{\gamma}' = (\hat{\gamma}_0, \dots, \hat{\gamma}_k)$

$$\gamma'(\theta) = (\gamma_0(\theta), \dots, \gamma_k(\theta))$$

$$\gamma_k(\theta) = f(a_1, \dots, a_p, b_1, \dots, b_q, \sigma_t^2)$$

$$C' = \left(1, \frac{N-1}{N}, \dots, \frac{N-|k|}{N} \right)$$

$$k < N-1$$

Es importante observar que en (52) se tiene un sistema sobredeterminado de k ecuaciones con $p+q+1$ incógnitas, en donde los segundos momentos muestrales (centrales), que son sólo funciones de la serie de tiempo observada x , se igualan a sus correspondientes segundos momentos poblacionales (centrales), más un error aleatorio, semejante al sistema que se obtendría al utilizar el método de momentos generalizado para estimar al vector de parámetros θ . Se sabe que los estimadores obtenidos por este último método, tienen la propiedad de ser consistentes, por lo que se espera que

el estimador $\hat{\theta}$ de θ , que se obtenga al minimizar la forma cuadrática en (53), después de haber realizado el ajuste en (52), tenga la propiedad de ser un estimador consistente de θ .

El método propuesto sin embargo, presenta el problema práctico, nada sencillo de resolver, de obtener la sucesión de autocovarianzas teóricas $\gamma(\theta)$, expresadas en términos de los parámetros en (53). Se sabe por (51) que son funciones no lineales de los parámetros de interés, pero se desconoce la forma analítica de estas funciones para el modelo general ARMA(p,q) en (50). En el capítulo 3, se propone y desarrolla un método numérico, que parte de la relación entre las autocovarianzas teóricas y los parámetros del modelo en (51), a través del cual se llega a valores numéricos de las autocovarianzas teóricas, dado un vector de parámetros θ , resolviendo en forma recursiva una serie de sistemas de ecuaciones en diferencias.

Una vez que se cuenta con un algoritmo que de solución al problema mencionado en el párrafo anterior, el método propuesto de estimación no presenta mayor problema, ya que dada una serie de tiempo observada x , y habiendo identificado por algún camino apropiado el modelo ARMA(p,q) que la generó, se obtendría la sucesión de autocovarianzas muestrales $\hat{\gamma}_0, \dots, \hat{\gamma}_k$, a partir de la segunda ecuación en (45), asignando un valor apropiado para k (algunos autores como Box y Jenkins sugieren, en la práctica dar un valor a k igual a $N/4$, donde N es el tamaño de muestra.).

Posteriormente, se obtendrían los valores numéricos de la sucesión de autocovarianzas teóricas, dado un vector inicial de parámetros θ . Finalmente por medio de algún algoritmo de optimización se resolvería el problema en (53).

En el capítulo 3, se describe en detalle el planteamiento e implementación del algoritmo desarrollado para este método de estimación.

Es importante hacer notar que este método de estimación, no toma en cuenta la información proporcionada por la estructura de covarianzas de las autocovarianzas estimadas. Un método el cual si considera esta información, es el de mínimos cuadrados no lineales generalizados, donde el problema a resolver sería

$$\text{MIN}_{\theta} (\hat{\gamma} - C \gamma (\theta))^T \Omega^{-1} (\hat{\gamma} - C \gamma (\theta))$$

donde Ω , es la matriz de varianzas y covarianzas de las autocovarianzas, cuyos elementos son funciones no lineales

de los parámetros, por lo que lo más que se podría intentar resolver sería

$$\underset{\theta}{\text{MIN}} \quad (\hat{\gamma} - C \gamma(\theta))^T \Omega^{-1}(\theta) (\hat{\gamma} - C \gamma(\theta)) \quad (54)$$

o

$$\underset{\theta}{\text{MIN}} \quad (\hat{\gamma} - C \gamma(\theta))^T \hat{\Omega}^{-1} (\hat{\gamma} - C \gamma(\theta)) \quad (55)$$

donde la forma de los elementos de la matriz de varianzas y covarianzas $\Omega(\theta)$ en (54) es desconocida, y los elementos del estimador de la matriz de varianzas y covarianzas $\hat{\Omega}$ en (55), son éstos en la primera ecuación de (47). Obviamente, el camino a seguir por este método, parece ser altamente complejo, aunque se esperaría, que al introducir más información sobre los parámetros, por considerar la forma cuadrática en (54) o (55) para la estimación, se pudiesen obtener estimadores más precisos de los mismos. En (1984) A. Bustos y M. Ahsanullah compararon ambos métodos de estimación, el de mínimos cuadrados no lineales, y el de mínimos cuadrados no lineales generalizados para un proceso AR(1); los resultados que observaron basados en un proceso de simulación, y en términos del error cuadrático medio de los estimadores, fue que el método de mínimos cuadrados no lineales, produjo mejores estimadores del parámetro

autorregresivo a_1 , para algunos valores particulares de este parámetro dentro de la región de estacionariedad, y varios tamaños de muestra, sin embargo, el estimador obtenido para la varianza residual σ_ϵ^2 fue el peor en todos los casos considerados.

En la sección 2.3, cuando se describieron las propiedades de las autocovarianzas muestrales, se vio que la sucesión

$$\sqrt{N} (\hat{\gamma}_0 - \gamma_0), \dots, \sqrt{N} (\hat{\gamma}_k - \gamma_k) \quad \text{para } k \text{ fija}$$

se distribuye asintóticamente como normal k -variada, con vector de medias cero y matriz de varianzas y covarianzas $(\Omega_{k,s})$ dada en (39), y en consecuencia la sucesión de errores aleatorios

$$\sqrt{N} \eta_0, \sqrt{N} \eta_1, \dots, \sqrt{N} \eta_k$$

tendrá la misma distribución asintótica, donde η_0, \dots, η_k se definieron al ajustar las autocovarianzas teóricas, a las autocovarianzas estimadas en (42).

Finalmente, es importante mencionar algunas de las ventajas que se observan en el método propuesto en esta última sección, con respecto a los métodos descritos en la sección 2.1 anterior. Primero, no es necesario asignar

valores iniciales a las series x_t y ε_t , definidas en sección 2.1, como sucede en el enfoque condicional, evitando así los grandes errores de pronóstico que se obtendrían, al utilizar dicho enfoque cuando alguna de las raíces del polinomio $A(Z)$, está muy cercana al círculo unitario complejo. Segundo, no es necesario estimar las series x_t y ε_t , como sucede en el enfoque no condicional, evitando así el largo proceso iterativo que se presentaría al utilizar dicho enfoque, cuando el tamaño de la serie de tiempo observada fuese pequeño. Tercero, en todo momento se trabaja con vectores y matrices de dimensión k , la cual en la práctica, siempre es mucho menor que el tamaño de muestra N , facilitando así las operaciones, y mejorando la eficiencia computacional de los otros métodos. Cuarto, se llega a expresiones numéricas "exactas" (no son exactas en el sentido de que se obtienen para un vector de parámetros θ^* , distinto del parámetro verdadero) de las autocovarianzas teóricas, así como de las derivadas parciales de las mismas, para un vector θ dado, a través de un método sencillo y fácil de programar en una computadora, como se verá en el capítulo 3.

Capítulo 3

ALGORITMO DE MINIMOS CUADRADOS NO LINEALES PARA LA ESTIMACION DE PARAMETROS DE MODELOS ARMA (p,q)

3.1 Introducción.

En este capítulo se describe la forma en que se llevó a cabo la estimación del vector de parámetros

$$\theta = (a_1, \dots, a_p, b_1, \dots, b_q, \sigma^2)$$

del modelo ARMA(p,q)

$$A(R) X_t = B(R) \xi_t$$

Inicialmente, en 3.2, se presentan las relaciones

básicas que hacen posible llegar a las autocovarianzas, así como a las derivadas de las autocovarianzas del modelo, a partir de los parámetros del mismo.

Después, en 3.3, se describe en forma detallada, cómo fue que se implementó el algoritmo en una computadora B-7800.

Este algoritmo consta principalmente de dos bloques:

En el primero de ellos, se obtienen estimadores de las autocovarianzas así como, de los parámetros iniciales del modelo. Para poder hacer ésto, se simula una muestra de datos, dados los ordenes p y q del modelo.

En este bloque se utilizaron las siguientes subrutinas del paquete IMSL:

FTGEN genera una serie de tiempo a partir de un modelo ARMA(p, q) dado.

FTAUTO obtiene autocovarianzas estimadas a partir de una serie de tiempo estacionaria.

FTARPS estima parámetros autorregresivos iniciales de un modelo ARMA.

FTMPS estima parámetros iniciales de promedios móviles de un modelo ARMA.

En el segundo de estos bloques, se obtiene la función de autocovarianza valuada en el vector de parámetros θ así como, su jacobiano valuado en θ . En sí, este bloque es el más importante, ya que aquí se implementaron las relaciones presentadas en 3.2.

Aquí también se utilizaron las siguientes subrutinas de apoyo:

Paquete de Análisis Numérico:

DECOMP calcula la descomposición triangular de una matriz cuadrada real.

SOLVE obtiene la solución de un sistema de ecuaciones lineales, a partir de la descomposición triangular de la matriz coeficiente.

Paquete IMSL:

ZRPOLY obtiene raíces de un polinomio con coeficientes reales.

Por último la estimación de los parámetros del modelo, se lleva a cabo con la información proporcionada por los dos bloques anteriores, así como la subrutina de apoyo LMDER1 del paquete de análisis numérico, para encontrar los parámetros que minimizan la función

$$(\hat{P} - CV(\theta))^T I (\hat{P} - CV(\theta)).$$

Existen otras subrutinas que se desarrollaron, y que se explican dentro de esta misma sección.

En el apéndice C, se puede encontrar una descripción detallada de cómo funcionan las subrutinas de apoyo que se emplearon.

Un listado del programa que se elaboró se encuentra en el apéndice B.

3.2 Relaciones Básicas.

En esta sección se da solución analítica a un problema relacionado con los procesos ARMA(p,q). El problema consiste en lo siguiente:

Como se mencionó en la sección 2.2, el objetivo es estimar el vector de parámetros desconocido θ , a partir de las autocovarianzas γ 's del proceso, y para hacerlo se pretende encontrar el vector de parámetros $\hat{\theta}$ que minimice

$$(\hat{P} - CV(\theta))^T I (\hat{P} - CV(\theta))$$

(1)

En consecuencia, se tienen que determinar el vector de autocovarianzas $\gamma(\theta)$ a partir de un valor dado del vector θ , y el vector de autocovarianzas estimadas del proceso $\hat{\gamma}$.

El procedimiento para determinar $\hat{\gamma}$ se describe en 3.3, y el problema de cómo determinar $\gamma(\theta)$, es precisamente al que se da solución en esta sección.

Se sabe (ver apéndice A), que la Función Generatriz de Autocovarianzas, a la que en adelante se denotará por FGA, de un proceso ARMA(p,q), relaciona las autocovarianzas de un proceso con los parámetros del mismo, mediante la igualdad

$$G_x(Z) = \sum_{k=-\infty}^{\infty} \gamma_k Z^k = \frac{\sigma_e^2 B(Z)B(Z^{-1})}{A(Z)A(Z^{-1})} \quad (2)$$

y dado que el proceso es estacionario e invertible se puede lograr tal factorización única de $G_x(Z)$ (ver apéndice A). Igualando coeficientes de Z^k en (2) se obtienen expresiones para las autocovarianzas en términos de los parámetros. Sin embargo, la relación que guardan las autocovarianzas con los parámetros en (2), es una relación no lineal, por lo que la forma de los coeficientes de Z^k en la parte derecha de la igualdad (2) no es inmediata. Por lo tanto, es necesario hacer un análisis más profundo de esta igualdad.

Sea $\{X_t\}$ un proceso ARMA(p,q) con media conocida e igual a cero, satisfaciendo condiciones de estacionariedad e invertibilidad y expresado como

$$A(R) X_t = B(R) \mathcal{E}_t$$

La FGA de un proceso tal satisface las siguientes igualdades

$$G_x(Z) = \sum_{k=-\infty}^{\infty} \gamma_k Z^k = \frac{\sigma_t^2 B(Z)B(Z^{-1})}{A(Z)A(Z^{-1})} \quad (3)$$

donde A y B son polinomios de grados p y q respectivamente, de la forma

$$A(Z) = \sum_{j=0}^p a_j Z^j$$

con $a_0 = 1$ y $a_p \neq 0$

$$A(Z^{-1}) = \sum_{j=0}^p a_j Z^{-j} \quad (4.1)$$

$$B(Z) = \sum_{i=0}^q b_i Z^i$$

con $b_0 = 1$ y $b_q \neq 0$

$$B(Z^{-1}) = \sum_{i=0}^q b_i Z^{-i}$$

(4.2)

con coeficientes reales (a_j, b_i) ; $j = 0, \dots, p$, $i = 0, \dots, q$.

Se supone que $A(Z)$ tiene todos sus ceros fuera del círculo unitario, y en consecuencia $A(Z^{-1})$ tendrá todos sus ceros dentro. Como el proceso es estacionario $A(Z)$ no tiene ceros sobre el círculo unitario, por lo tanto la función $G_x(Z)$ tendrá polos dentro del círculo unitario para $Z = r_j$, donde r_j es algún cero de $A(Z^{-1})$, y polos fuera del círculo unitario para $Z = 1/r_j$, donde $1/r_j$ es algún cero de $A(Z)$, y ningún polo sobre el círculo unitario.

Para quitar los polos de $G_x(Z)$ que se encuentran fuera del círculo unitario, se define la siguiente función:

Sea $G_l(Z) = A(Z) G_x(Z)$

$$G_l(Z) = \sum_{j=0}^p a_j Z^j \sum_{k=-\infty}^{\infty} \gamma_k Z^k = \sum_{k=-\infty}^{\infty} \sum_{j=0}^p a_j \gamma_k Z^{k+j}$$

Sea $u = k+j$ y $v = j$

$$\begin{aligned}
 G1(Z) &= \sum_{u=0}^{\infty} \sum_{v=0}^p \gamma_{u-v} a_v Z^u + \sum_{u=-\infty}^{-1} \sum_{v=0}^p \gamma_{u-v} a_v Z^u \\
 &= \sum_{u=-\infty}^{\infty} Z^u \sum_{v=0}^p \gamma_{u-v} a_v = \sum_{u=-\infty}^{\infty} S_u Z^u
 \end{aligned}$$

donde, S_u es un polinomio en γ de la forma

$$S_u = \sum_{v=0}^p \gamma_{u-v} a_v \tag{5}$$

La nueva función $G1(Z)$ sólo tiene polos dentro del círculo unitario, es decir, $G1(Z)$ satisface

$$G1(Z) = \frac{\sigma_z^2 B(Z)B(Z^{-1})}{A(Z^{-1})}$$

y por lo tanto, su expansión como serie de potencias en Z y

Z^{-1} es infinita solamente en las potencias negativas, por lo que S_u en (5) es igual a cero para $u > q$.

Para remover los polos se define una nueva función $G_2(Z)$ que satisface

$$G_2(Z) = A(Z^{-1}) G_1(Z)$$

$$G_2(Z) = \sum_{j=0}^p a_j Z^{-j} \sum_{u=-\infty}^{\infty} S_u Z^u = \sum_{u=-\infty}^{\infty} \sum_{j=0}^p a_j S_u Z^{u-j}$$

Sea $k = u-j$ y $v = j$

(6)

$$G_2(Z) = \sum_{k=-\infty}^{\infty} Z^k \sum_{v=0}^p a_v S_{k+v} = \sum_{k=-\infty}^{\infty} d_k Z^k$$

donde, d_k es un polinomio en S de la forma

$$d_k = \sum_{v=0}^p a_v S_{k+v}$$

(7)

Esta nueva función $G_2(Z)$ no tiene polos, ya que satisface

$$G_2(Z) = \sigma_t^2 B(Z)B(Z^{-1}) \quad (8)$$

y por lo tanto, su expansión en series de potencias en Z y Z^{-1} es finita, ya que

$$\begin{aligned} \sigma_t^2 B(Z)B(Z^{-1}) &= \sigma_t^2 \sum_{i=0}^q b_i Z^i \sum_{j=0}^q b_j Z^{-j} \\ &= \sigma_t^2 \sum_{i=0}^q \sum_{j=0}^q b_i b_j Z^{i-j} \end{aligned}$$

sea $u = i-j$ y $v = j$

$$\begin{aligned} &= \sigma_t^2 \sum_{u=-q}^{-1} \sum_{v=0}^{q+u} b_v b_{v-u} Z^u + \sigma_t^2 \sum_{u=0}^q \sum_{v=0}^{q-u} b_{v+u} b_v Z^u \\ &= \sigma_t^2 \sum_{u=-q}^q Z^u \sum_{v=0}^{q-|u|} b_v b_{v+|u|} = \sigma_t^2 \sum_{u=-q}^q C_u Z^u \quad (9) \end{aligned}$$

donde C_u son los coeficientes de $B(Z)B(Z^{-1})$.

Suponga que θ es conocido, es decir, que se conocen los coeficientes de los polinomios $A(Z)$ y $B(Z)$, así como la varianza del proceso $\{\epsilon_t\}$, σ_ϵ^2 . Entonces, se puede llegar a expresiones numéricas para las autocovarianzas, dados éstos parámetros conocidos; de la siguiente manera.

A partir de (9) se pueden encontrar los coeficientes C_u como

$$C_u = \begin{cases} \sigma_\epsilon^2 \sum_{v=0}^{q-|u|} b_v b_{v+|u|} & \text{para } |u| \leq q \\ 0 & \text{para } |u| > q \end{cases} \quad (10)$$

con $C_u = C_{-u}$ para u distinto de cero.

Observe que de (6), (8) y (9) se tiene la siguiente relación entre los coeficientes C_u y d_k

$$\sum_{k=-\infty}^{\infty} d_k Z^k = \sum_{u=-q}^q C_u Z^u$$

por lo que

$$d_k = \begin{cases} C_u & \text{para } |k| \leq q \\ 0 & \text{para } |k| > q \end{cases} \quad (11)$$

El sistema de ecuaciones en (7) ahora esta dado como

$$\sum_{v=0}^p a_v S_{k+v} = \begin{cases} d_k & \text{para } |k| \leq q \\ 0 & \text{para } |k| > q \end{cases} \quad (12)$$

y resolviendo este sistema se encuentran los coeficientes desconocidos S_{k+v} .

Finalmente, se resuelve el sistema de ecuaciones en (5) como

$$S_u = \begin{cases} \sum_{v=0}^p \gamma_{u-v} a_v & \text{para } -\infty < u \leq q \\ 0 & \text{para } u > q \end{cases} \quad (13)$$

y se encuentran los coeficientes γ_{u-v} , los cuales son las autocovarianzas del proceso.

Por lo tanto dado un vector de parámetros θ , y resolviendo los sistemas de ecuaciones en diferencias (10), (12) y (13), se puede determinar el vector de autocovarianzas $V(\theta)$ del proceso.

Para minimizar la expresión en (1), se hizo uso de la subrutina LMDER1 del paquete de subrutinas de Análisis Numérico. Esta subrutina minimiza la suma de cuadrados de M funciones no lineales en N variables, utilizando una modificación del algoritmo de Levenberg-Marquardt. La subrutina necesita conocer la función a minimizar, así como su jacobiano. Hasta ahora se conocen los sistemas de ecuaciones que llevan a la función, pero aún se desconoce cómo llegar al jacobiano de dicha función.

El jacobiano que se desea obtener tiene la siguiente forma

$$V(\theta) = (\nu_0(\theta), \dots, \nu_k(\theta))$$

$$\theta = (a_1, \dots, a_p, b_1, \dots, b_q, \sigma_c^2)$$

$$J_y = \begin{pmatrix} \frac{\partial \gamma_0(\theta)}{\partial a_1} & \dots & \frac{\partial \gamma_0(\theta)}{\partial b_1} & \dots & \frac{\partial \gamma_0(\theta)}{\partial \sigma_\epsilon^2} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \gamma_k(\theta)}{\partial a_1} & \dots & \frac{\partial \gamma_k(\theta)}{\partial b_1} & \dots & \frac{\partial \gamma_k(\theta)}{\partial \sigma_\epsilon^2} \end{pmatrix}_{(k+1) \times (p+q+1)} \quad (14)$$

Para determinar cada uno de los elementos de este jacobiano, se derivan las expresiones en (10), (12) y (13) con respecto a cada uno de los elementos de θ . Este procedimiento, da como resultado una serie de sistemas de ecuaciones en diferencias para las derivadas, muy similares a los que se obtuvieron para la función de autocovarianzas. Estos sistemas de ecuaciones se obtienen a continuación; primero, para los parámetros autorregresivos; segundo, para los parámetros de promedios móviles, y por último, para la varianza del proceso $\{S_t\}$.

Derivando (10) con respecto a a_i ; $i = 1, \dots, p$ se obtiene

$$\frac{\partial C_u}{\partial a_i} = \frac{\partial \sigma_\epsilon^2}{\partial a_i} \sum_{v=0}^{q-|u|} b_v b_{v+|u|} + \sigma_\epsilon^2 \frac{\partial}{\partial a_i} \sum_{v=0}^{q-|u|} b_v b_{v+|u|} = 0$$

y por (11) se tiene que $\frac{\partial}{\partial a_i} d_k = 0$.

Derivando (12) con respecto a a_i ; $i = 1, \dots, p$ se obtiene

$$\sum_{v=0}^p \frac{\partial a_v}{\partial a_i} S_{k+v} + a_v \frac{\partial S_{k+v}}{\partial a_i} = \frac{\partial d_k}{\partial a_i}$$

$$S_{k+i} + \sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial a_i} = 0 \quad \text{para toda } k.$$

Derivando (13) con respecto a a_i ; $i = 1, \dots, p$ se obtiene

$$\sum_{v=0}^p \frac{\partial a_v}{\partial a_i} \gamma_{u+v} + a_v \frac{\partial \gamma_{u+v}}{\partial a_i} = \frac{\partial S_u}{\partial a_i}$$

por lo tanto, la secuencia de p sistemas de ecuaciones que se tienen que resolver, para encontrar las derivadas de las autocovarianzas con respecto a los parámetros autorregresivos, son:

para $i = 1, \dots, p$

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial a_i} = \begin{cases} -S_{k+i} & \text{para } |k+i| \leq q \\ 0 & \text{en otro caso} \end{cases} \quad (15)$$

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial a_i} = \begin{cases} \frac{\partial S_u}{\partial a_i} - \gamma_{u-i} & \text{para } u \leq q \\ -\gamma_{u-i} & \text{en otro caso.} \end{cases} \quad (16)$$

Derivando (10) con respecto a b_j ; $j = 1, \dots, q$ se obtiene

$$\begin{aligned} \frac{\partial C_u}{\partial b_j} &= \frac{\partial \sigma_\epsilon^2}{\partial b_j} \sum_{v=0}^{q-|u|} b_v b_{v+|u|} + \sigma_\epsilon^2 \frac{\partial}{\partial b_j} \sum_{v=0}^{q-|u|} b_v b_{v+|u|} \\ &= \sigma_\epsilon^2 \sum_{v=0}^{q-|u|} \frac{\partial b_v}{\partial b_j} b_{v+|u|} + b_v \frac{\partial b_{v+|u|}}{\partial b_j} \end{aligned}$$

$$\frac{\partial C_u}{\partial b_j} = \begin{cases} \sigma_\varepsilon^2 (b_{j+|u|} + b_{j-|u|}) & \text{si } 0 \leq j+|u| \leq q \text{ y } 0 \leq j-|u| \leq q \\ \sigma_\varepsilon^2 b_{j+|u|} & \text{si } j-|u| < 0 \text{ y } j+|u| \leq q \\ \sigma_\varepsilon^2 b_{j-|u|} & \text{si } j-|u| \geq 0 \text{ y } j+|u| > q \\ 0 & \text{si } j-|u| < 0 \text{ y } j+|u| > q \end{cases}$$

$|u| \leq q$

y por (11) se tiene que $\frac{\partial d_k}{\partial b_j} = \frac{\partial C_u}{\partial b_j}$ para $|k| \leq q$.

Derivando (12) con respecto a b_j ; $j = 1, \dots, q$

$$\sum_{v=0}^p \frac{\partial a_v}{\partial b_j} S_{k+v} + a_v \frac{\partial S_{k+v}}{\partial b_j} = \frac{\partial d_k}{\partial b_j} \quad \text{de donde}$$

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial b_j} = \begin{cases} \frac{\partial d_k}{\partial b_j} & \text{para } |k| \leq q \\ 0 & \text{en otro caso.} \end{cases}$$

Derivando (13) con respecto a b_j ; $j = 1, \dots, q$

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial b_j} = \begin{cases} \frac{\partial S_u}{\partial b_j} & \text{para } u \leq q \\ 0 & \text{en otro caso.} \end{cases}$$

En consecuencia, los q sistemas de ecuaciones que se

tienen que resolver en forma recursiva, para obtener las derivadas de las autocovarianzas, con respecto a cada uno de los parámetros de promedios móviles son:

$$\frac{\partial du}{\partial b_j} = \left\{ \begin{array}{ll} \sigma_e^2 (b_{j+|u|} + b_{j-|u|}) & \text{si } \begin{array}{l} 0 \leq j+|u| \leq q \\ 0 \leq j-|u| \leq q \end{array} \text{ y} \\ \sigma_e^2 b_{j+|u|} & \text{si } \begin{array}{l} j-|u| < 0 \\ j+|u| \leq q \end{array} \text{ y} \\ \sigma_e^2 b_{j-|u|} & \text{si } \begin{array}{l} j-|u| \geq 0 \\ j+|u| > q \end{array} \text{ y} \\ 0 & \text{si } \begin{array}{l} j-|u| < 0 \\ j+|u| > q \end{array} \text{ y} \end{array} \right.$$

(17)

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial b_j} = \begin{cases} \frac{\partial d_k}{\partial b_j} & \text{para } |k| \leq q \\ 0 & \text{en otro caso} \end{cases} \quad (18)$$

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial b_j} = \begin{cases} \frac{\partial S_u}{\partial b_j} & \text{para } u \leq q \\ 0 & \text{en otro caso.} \end{cases} \quad (19)$$

Derivando (10) con respecto a la varianza residual σ_e^2

$$\frac{\partial C_u}{\partial \sigma_e^2} = \begin{cases} \sum_{v=0}^{q-|u|} b_v b_{v+|u|} = du / \sigma_e^2 & \text{para } |u| \leq q \\ 0 & \text{e.o.c.} \end{cases}$$

por (11) se tiene que:

$$\frac{\partial d_k}{\partial \sigma_e^2} = \begin{cases} \frac{\partial Cu}{\partial \sigma_e^2} = du/\sigma_e^2 & \text{para } |k| \leq q \\ 0 & \text{e.o.c.} \end{cases} \quad (20)$$

Derivando (12) con respecto a la varianza residual σ_e^2

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial \sigma_e^2} = \begin{cases} du/\sigma_e^2 & \text{para } |k| \leq q \\ 0 & \text{e.o.c.} \end{cases} \quad (21)$$

Finalmente, derivando (13) con respecto a la varianza residual σ_e^2

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial \sigma_e^2} = \begin{cases} \frac{\partial Su}{\partial \sigma_e^2} & \text{para } u \leq q \\ 0 & \text{e.o.c.} \end{cases} \quad (22)$$

En consecuencia, resolviendo los tres sistemas de ecuaciones anteriores en forma recursiva, se encuentran las derivadas de las autocovarianzas con respecto a σ_e^2 .

3.3 Implementación del Algoritmo.

Esta sección, como se mencionó en la introducción de este capítulo, se divide en dos partes. En la primera de ellas, subsecciones 3.3.1 y 3.3.2, se muestra el estimador de las autocovarianzas que se utilizó, y el procedimiento para calcularlo, también aquí, se describe cómo se obtuvo el estimador inicial del vector de parámetros θ . En la segunda, subsecciones 3.3.3 y 3.3.4, se describe el procedimiento que se siguió, para obtener la función de autocovarianzas y su matriz jacobiana valuados en el vector de parámetros θ .

3.3.1 Estimador de las Autocovarianzas.

Se decidió utilizar los estimadores sesgados de las autocovarianzas, ya que, como se mencionó en el capítulo 2, éste posee un error cuadrático medio menor que los estimadores insesgados de las mismas, y asintóticamente ambos estimadores se comportan igual.

Dada una serie de tiempo observada $x = (x_1, \dots, x_N)$ del proceso $\{X_t\}$, el estimador sesgado $\hat{\gamma}$ del vector de autocovarianzas $\gamma(\theta)$ se definió como

$$\hat{\gamma} = (\hat{\gamma}_0, \dots, \hat{\gamma}_k)$$

donde

$$\hat{\gamma}_i = \frac{1}{N} \sum_{t=1}^{N-i} X_t X_{t+i} \quad \text{para } i = 0, \dots, N-1$$

Para obtener el vector de estimadores $\hat{\gamma}$, primero se simularon 300 muestras de tamaño TM, donde TM tomó los valores de 25, 100 y 300. Las series de tiempo se simularon utilizando la subrutina FTGEN del paquete IMSL, y el procedimiento que se utilizó fue el siguiente:

i) Se impusieron condiciones iniciales iguales a cero.

ii) Se simularon 100+TM observaciones y se despreciaron las primeras 100, con objeto de que las condiciones iniciales impuestas perdieran influencia. En este punto se obtuvo la primera muestra de tamaño TM.

iii) Las demás muestras, se obtuvieron dando como condiciones iniciales, las últimas observaciones de la muestra obtenida en un paso anterior, simulando 50+TM observaciones, y despreciando las primeras 50, con objeto de que las 300 muestras estuvieran poco correlacionadas. Cabe recordar que el proceso es gaussiano.

Una vez obtenidas las 300 muestras de tamaño TM, se

utilizó la subrutina FTAUTO del paquete IMSL, para calcular el vector de autocovarianzas estimadas $\hat{\gamma}$ para cada muestra, bajo el supuesto de que la media del proceso es cero.

Una descripción de cómo funcionan las subrutinas FTGEN y FTAUTO se puede encontrar en el apéndice C.

3.3.2 Estimador inicial del vector de parámetros θ .

En 3.1, donde se desarrollaron las relaciones básicas para obtener el vector de autocovarianzas $\gamma(\theta)$, se supuso que el vector de parámetros θ era conocido. En consecuencia, como se verá en la sección 3.3.3, se necesita de un vector inicial para arrancar el proceso.

Para obtener un vector inicial θ^0 , el cual se espera, no se aleje mucho del vector θ que minimiza la función en (1), se estimaron por separado los parámetros autorregresivos y los de promedios móviles, haciendo uso de la información proporcionada por el vector de autocovarianzas estimadas $\hat{\gamma}$, siguiendo el método planteado en la sección 2.1.4.

Los estimadores iniciales de los parámetros autorregresivos, se obtuvieron con la subrutina FTARPS del paquete IMSL, tomando en cuenta la información proporcionada por cada vector $\hat{\gamma}$, para cada una de las 300 muestras simuladas. Después, se verificó que estos parámetros

obtenidos, satisficieran las condiciones de estacionariedad, ésto se hizo con la subrutina REST, cuyo listado se encuentra en el apéndice B. Esta subrutina encuentra las raíces del polinomio característico:

$$1 - a_1^o Z - a_2^o Z^2 - \dots - a_p^o Z^p = 0 \quad (23)$$

donde a_1^o, \dots, a_p^o , son los parámetros iniciales obtenidos a partir de la subrutina FTARPS, y verifica que todas ellas se encuentren fuera del círculo unitario. Si alguna o algunas de las raíces del polinomio en (23), se encuentran dentro del círculo unitario, entonces toma los inversos de éstas y encuentra los nuevos coeficientes del polinomio característico, que tiene todas sus raíces fuera del círculo unitario. En esencia lo que hace REST es factorizar la función $A^*(Z) = A(Z)A(Z^{-1})$, donde $A(Z)$ tiene todas sus raíces fuera y $A(Z^{-1})$ tiene todas sus raíces dentro del círculo unitario.

Una vez que se tienen los parámetros iniciales autorregresivos, se procede a obtener los parámetros iniciales de promedios móviles, así como el parámetro inicial de la varianza del proceso $\{\epsilon_t\}$, para lo cual se utilizó la subrutina FTMPS del paquete IMSL, haciendo uso de la información proporcionada por el vector de autocovarianzas estimadas $\hat{\gamma}$ y los parámetros iniciales

autorregresivos, para cada una de las 300 muestras simuladas. Después, se verificó que estos parámetros obtenidos satisficieran las condiciones de invertibilidad, ésto se hizo con la subrutina REST mencionada anteriormente.

Una descripción de como funcionan las subrutinas FTARPS y FTMPS se puede encontrar en el apéndice C, así como la subrutina ZRPOLY, la cual se utilizó en REST para encontrar las raíces de polinomios con coeficientes reales. Un listado comentado de la subrutina REST se encuentra en el apéndice B.

Notas:

i) En todos los casos se supuso que la media del proceso era conocida e igual a cero, en base a este supuesto se simularon las series y se estimaron las autocovarianzas.

ii) En el caso en que el proceso era puramente autorregresivo, se le dio el valor inicial arbitrario de 0.5 a la varianza residual.

iii) Cuando la subrutina FTMPS no convergió para encontrar valores iniciales de los parámetros, se les dio el valor arbitrario de 0.5, tanto a los parámetros de promedios móviles como al de la varianza residual. En el caso en que el orden de la parte de promedios móviles era igual a 2, se les dio el valor arbitrario de 0.4 a los parámetros, con el

fin de evitar raíces sobre el círculo unitario.

3.3.3 Determinación del vector de autocovarianzas $\gamma(\theta)$.

Dado el vector de parámetros inicial θ , el cual se obtuvo en la sección 3.3.2, se explica a continuación la forma en que se implementaron los sistemas de ecuaciones (10) - (13) desarrollados en 3.2, para la determinación del vector de autocovarianzas $\gamma(\theta)$, en una computadora Burroughs 7800.

i) Dados los parámetros iniciales b_1^0, \dots, b_q^0 y $\sigma_\varepsilon^{2,0}$, se calcularon los valores $d_{-q}, \dots, d_0, \dots, d_q$, a partir de (10) y (11) como:

$$d_k = \begin{cases} \sigma_\varepsilon^{2,0} \sum_{v=0}^{q-|k|} b_v^0 b_{v+|k|}^0 & \text{para } |k| \leq q \\ 0 & \text{para } |k| > q \\ d_{-k} & \text{para } k \neq 0 \end{cases} \quad (24)$$

estos cálculos se efectuaron en la subrutina CALD, cuyo listado se encuentra en el apéndice B.

ii) Dados a_1^0, \dots, a_p^0 y $d_{-q}, \dots, d_0, \dots, d_q$, se resolvió el sistema de ecuaciones en diferencias (12), para encontrar valores de S_0, \dots, S_q . La relación en (12) era de la forma:

$$\sum_{v=0}^p a_v S_{k+v} = \begin{cases} d_k & \text{para } |k| \leq q \\ 0 & \text{para } |k| > q \end{cases} \quad (25)$$

en forma matricial este sistema se puede expresar como

$$A S = D$$

donde A es la matriz de parámetros autorregresivos, S es el vector de soluciones, y D es el vector de valores conocidos $d_{-q}, \dots, d_0, \dots, d_q$ obtenidos en (i).

La forma de la matriz A, así como la de los vectores S y D varía de acuerdo a los ordenes p y q del modelo ARMA, en la forma siguiente:

a) $p \geq q$

a.1) Se obtuvieron valores de S_0, \dots, S_q a partir de (12) y (13), resolviendo el siguiente sistema triangular de dimensión $(q+1) \times (q+1)$

$$\begin{pmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_q & a_{q-1} & \dots & a_0 \end{pmatrix} \begin{pmatrix} S_q \\ S_{q-1} \\ \cdot \\ \cdot \\ \cdot \\ S_0 \end{pmatrix} = \begin{pmatrix} d_q \\ d_{q-1} \\ \cdot \\ \cdot \\ \cdot \\ d_0 \end{pmatrix} \quad (26)$$

donde $a_0 = 1$ y $S_{q-i} = d_{q-i} - \sum_{j=1}^i a_j S_{q+j-i}$ para $i = 1, \dots, q$.

No se obtuvieron valores adicionales de S , ya que no fueron necesarios para los sistemas siguientes.

b) $p < q$

b.1) Se obtuvieron valores de S_{q-p}, \dots, S_q , resolviendo el siguiente sistema triangular de dimensión $(p+1) \times (p+1)$

$$\begin{pmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_p & a_{p-1} & \dots & a_0 \end{pmatrix} \begin{pmatrix} S_q \\ S_{q-1} \\ \cdot \\ \cdot \\ \cdot \\ S_{q-p} \end{pmatrix} = \begin{pmatrix} d_q \\ d_{q-1} \\ \cdot \\ \cdot \\ \cdot \\ d_{q-p} \end{pmatrix} \quad (27)$$

b.2) Dados S_{q-p}, \dots, S_q , se resolvió en forma sucesiva el siguiente sistema de dimensión $(q-p) \times (p+1)$, para obtener los valores de S_0, \dots, S_{q-p-1}

$$\begin{pmatrix} S_0 & S_1 & \dots & S_p \\ S_1 & S_2 & \dots & S_{p+1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ S_{q-p-1} & S_{q-p} & \dots & S_{q-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} d_0 \\ d_1 \\ \cdot \\ \cdot \\ \cdot \\ d_{q-p-1} \end{pmatrix} \quad (28)$$

* = valores obtenidos en b.1

Ahora se tiene un sistema de la forma $S A = D$, con S

siendo una matriz de dimensión $(q-p) \times (p+1)$, cuyos únicos elementos conocidos son los obtenidos en b.l, A es el vector de parámetros autorregresivos y D es el vector de valores conocidos d_0, \dots, d_{q-p-1} .

Los cálculos anteriores para encontrar S_0, \dots, S_q , en ambos casos (a) y (b), se efectuaron en la subrutina CALS, cuyo listado se encuentra en el apéndice B.

iii) Dadas S_0, \dots, S_q , se resolvió el sistema de ecuaciones en diferencias (13)

$$\sum_{v=0}^p a_v \gamma_{|u-v|} = \begin{cases} S_u & \text{para } u \leq q \\ 0 & \text{para } u > q \end{cases} \quad (29)$$

para obtener el vector de autocovarianzas

$$\gamma(\theta) = (\gamma_0(\theta), \gamma_1(\theta), \dots, \gamma_{LAG}(\theta)) \quad (30)$$

donde Lag es un entero positivo especificado por el usuario, y menor que $N-1$.

El proceso para obtener $\gamma(\theta)$ varía de acuerdo a los

tres casos siguientes:

a) $p > q$

a.1) Se tenía que resolver el siguiente sistema, para obtener $\gamma_0, \dots, \gamma_p$

$$\begin{pmatrix}
 \gamma_0 & \gamma_{-1} & \dots & \gamma_{-p} \\
 \gamma_1 & \gamma_0 & \dots & \gamma_{1-p} \\
 \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot \\
 \gamma_{p-1} & \gamma_p & \dots & \gamma_{-1} \\
 \gamma_p & \gamma_{p-1} & \dots & \gamma_0
 \end{pmatrix}
 \begin{pmatrix}
 a_0 \\
 a_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 a_{p-1} \\
 a_p
 \end{pmatrix}
 =
 \begin{pmatrix}
 S_0 \\
 \cdot \\
 \cdot \\
 S_q \\
 \cdot \\
 0 \\
 \cdot \\
 \cdot \\
 0
 \end{pmatrix}
 \begin{matrix}
 \left. \vphantom{\begin{pmatrix} S_0 \\ \cdot \\ \cdot \\ S_q \\ \cdot \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}} \right\} (q+1) \times 1 \\
 \left. \vphantom{\begin{pmatrix} S_0 \\ \cdot \\ \cdot \\ S_q \\ \cdot \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}} \right\} (p-q) \times 1
 \end{matrix}
 \quad (31)$$

en forma matricial: $A X = b$

donde en este caso, la matriz A es una matriz Toeplitz cuyos elementos son valores desconocidos por calcular.

Se resolvió un sistema equivalente al anterior, de dimensión $(2P+1) \times (2P+1)$, agregando p restricciones de la forma

$$\chi_i = \gamma_i \quad \text{para } i = 1, \dots, p \quad (32)$$

y el nuevo sistema que se resolvió fue

$$\begin{pmatrix} a_p & a_{p-1} & \dots & a_0 & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & a_p & a_{p-1} & \dots & a_0 \\ 1 & 0 & \dots & 0 & 0 & \dots & -1 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 1 & 0 & -1 & \dots & 0 \end{pmatrix} \begin{pmatrix} \gamma_{-p} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_0 \\ \gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_p \end{pmatrix} = \begin{pmatrix} S_0 \\ \cdot \\ \cdot \\ \cdot \\ S_q \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \begin{matrix} (q+1) \\ \\ \\ \\ \\ \\ \\ \\ \\ (2p-q) \end{matrix} \quad (33)$$

obteniendo así $\gamma_0, \dots, \gamma_p$.

a.2) si Lag era mayor que p, entonces, dados $\gamma_0, \dots, \gamma_p$ se resolvió el siguiente sistema homogéneo en forma recursiva, para obtener $\gamma_{p+1}, \dots, \gamma_{Lag}$.

$$\sum_{v=0}^p a_v \gamma_{u-v} = 0 \quad \text{donde } u > p \quad (34)$$

b) $p = q$

b.1) se resolvió un sistema parecido a (33), pero en donde los elementos del vector en la derecha cambian

$$\begin{bmatrix} a_p & a_{p-1} & \dots & a_0 & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & a_p & a_{p-1} & \dots & a_0 \\ 1 & 0 & \dots & 0 & 0 & \dots & -1 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 1 & 0 & -1 & \dots & 0 \end{bmatrix} \begin{bmatrix} \gamma_p \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \gamma_0 \\ \gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_p \end{bmatrix} = \begin{bmatrix} S_0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ S_q \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \begin{matrix} q+1 \\ \\ \\ \\ q \end{matrix} \quad (35)$$

a partir de aquí se encontró $\gamma_0, \dots, \gamma_p$.

$$\sum_{v=0}^p a_v \gamma_{u-v} = S_u \quad \text{para } p+1 \leq u \leq q. \quad (37)$$

Si Lag era mayor que p, pero menor que q, se resolvio q-Lag ecuaciones como en (37), para encontrar $\gamma_{p+1}, \dots, \gamma_{LAG}$.

c.3) Si Lag > q entonces, dados $\gamma_0, \dots, \gamma_q$, se resolvio en forma recursiva el siguiente sistema homogéneo de ecuaciones

$$\sum_{v=0}^p a_v \gamma_{u-v} = 0 \quad \text{para } q+1 \leq u \leq \text{Lag} \quad (38)$$

para encontrar $\gamma_{q+1}, \dots, \gamma_{LAG}$.

Todos los cálculos anteriores se efectúan en la subrutina CALR, cuyo listado se encuentra en el apéndice B. Esta, utiliza otras dos, del paquete de subrutinas de análisis numérico, éstas son: DECOMP, la cual calcula la descomposición triangular de la matriz cuadrada real de coeficientes en (33), (35) y (36), y SOLVE, la cual obtiene

la solución de un sistema de ecuaciones lineales, a partir de la descomposición triangular de la matriz de coeficientes. Una descripción de como funcionan éstas últimas dos subrutinas, se encuentra en el apéndice C.

3.3.4 Determinación del Jacobiano del vector de Autocovarianzas $\gamma(\theta)$

Al igual que se hizo en 3.2, aquí se describe primero cómo se implementaron los sistemas (15) y (16), para obtener las derivadas del vector de autocovarianzas $\gamma(\theta)$ con respecto a cada uno de los parámetros autorregresivos, después se describe cómo se implementaron los sistemas (17), (18) y (19), para obtener las derivadas de $\gamma(\theta)$ con respecto a cada uno de los parámetros de promedios móviles, y finalmente, se describe cómo se implementaron los sistemas (20), (21) y (22) para obtener la derivada del vector $\gamma(\theta)$ con respecto al parámetro de la varianza residual.

i) Derivadas de $\gamma(\theta)$ con respecto a los parámetros autorregresivos a_1, \dots, a_p .

Lo que interesa conocer son las primeras p columnas del jacobiano en (14), es decir:

$$\begin{array}{ccc}
 \frac{\partial \gamma_0(\theta)}{\partial a_1} & \dots & \frac{\partial \gamma_0(\theta)}{\partial a_p} \\
 \cdot & & \cdot \\
 \cdot & & \cdot \\
 \cdot & & \cdot \\
 \frac{\partial \gamma_k(\theta)}{\partial a_1} & \dots & \frac{\partial \gamma_k(\theta)}{\partial a_p}
 \end{array}$$

(39)

donde k es el máximo orden de las autocovarianzas que el usuario desea considerar, y al que se ha hecho referencia como Lag.

Dado que en 3.3.3 se calcularon valores de S_0, S_1, \dots, S_q y se conocen los parámetros iniciales a_1, \dots, a_p obtenidos en 3.3.2, el primer paso es resolver (15), para $i = 1, \dots, p$.

Recordando que (15) era igual a

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial a_i} = \begin{cases} - S_{k+i} & \text{para } |k+i| \leq q \\ 0 & \text{en otro caso.} \end{cases}$$

Este sistema de ecuaciones se puede expresar en forma

matricial como:

$$S^* A = S \quad (40)$$

donde S^* , es la matriz cuyos elementos son las $\frac{\partial S_{k+v}}{\partial a_i}$ en (15),

para los cuales se desea encontrar un valor numérico, A es el vector de parámetros autorregresivos iniciales a_0, a_1, \dots, a_p , con $a_0 = 1$, y S es un vector cuyos elementos pueden ser los valores conocidos S_0, \dots, S_q ó ceros.

La estructura de la matriz S^* y del vector S en (40), varía de acuerdo a los ordenes p y q del modelo, de la forma siguiente:

a) $p < q$

Para $0 \leq k \leq q$ e $i = 1, \dots, p$, se tiene el sistema siguiente

$$\begin{array}{c}
 S_2^* \\
 \left. \begin{array}{c}
 \frac{\partial S_0}{\partial a_i} \quad \frac{\partial S_1}{\partial a_i} \quad \dots \quad \frac{\partial S_p}{\partial a_i} \\
 \frac{\partial S_1}{\partial a_i} \quad \frac{\partial S_2}{\partial a_i} \quad \dots \quad \frac{\partial S_{p+1}}{\partial a_i} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial S_{q-p}}{\partial a_i} \quad \frac{\partial S_{q-p+1}}{\partial a_i} \quad \dots \quad \frac{\partial S_q}{\partial a_i} \\
 \frac{\partial S_{q-p+1}}{\partial a_i} \quad \frac{\partial S_{q-p+2}}{\partial a_i} \quad \dots \quad \frac{\partial S_q}{\partial a_i} \\
 \cdot \\
 \cdot \\
 \cdot
 \end{array} \right\} \\
 \\
 S_1^* \\
 \left. \begin{array}{c}
 \frac{\partial S_{q-t}}{\partial a_i} \quad \frac{\partial S_q}{\partial a_i} \quad \dots \quad 0 \\
 \frac{\partial S_q}{\partial a_i} \quad 0 \quad \dots \quad 0
 \end{array} \right\}
 \end{array}
 =
 \begin{array}{c}
 \left. \begin{array}{c}
 a_0 \\
 a_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 a_{p-1} \\
 a_p
 \end{array} \right\} \\
 \\
 \left. \begin{array}{c}
 -S_1 \\
 -S_{1+t} \\
 \cdot \\
 \cdot \\
 \cdot \\
 -S_{1+q-p} \\
 -S_{1+q-p+t} \\
 \cdot \\
 \cdot \\
 \cdot \\
 -S_{1+q-t} \\
 -S_{1+q}
 \end{array} \right\}
 \end{array}
 \begin{array}{c}
 \left. \begin{array}{c}
 q-p \\
 \\
 \\
 \\
 \\
 \\
 \\
 \\
 \\
 p+1
 \end{array} \right\}
 \end{array}
 \quad (41)$$

en forma equivalente
$$\begin{pmatrix} S_2^* \\ S_1^* \end{pmatrix} A = \begin{pmatrix} S_2 \\ S_1 \end{pmatrix}$$

donde S^* se particionó en dos submatrices; una matriz S_2^* de dimensión $(q-p) \times (p+1)$, y una matriz triangular S_1^* de dimensión $(p+1) \times (p+1)$, y el vector S se particionó en dos subvectores; un vector S_2 de dimensión $(q-p) \times 1$ y otro vector S_1 de dimensión $(p+1) \times 1$.

a.1) Se obtuvieron valores de
$$\frac{\partial S_{q-p}}{\partial a_i}, \dots, \frac{\partial S_q}{\partial a_i},$$

resolviendo el siguiente sistema triangular de dimensión $(p+1) \times (p+1)$

$$S_1^* A = S_1 \tag{42}$$

para $i = 1, \dots, p$, donde el vector S_1 , varía de la forma siguiente con respecto al valor de i

$$\begin{pmatrix} i = 1 \\ -S_{q-p+1} \\ -S_{q-p+2} \\ \cdot \\ \cdot \\ \cdot \\ -S_q \\ 0 \end{pmatrix}, \begin{pmatrix} i = 2 \\ -S_{q-p+2} \\ -S_{q-p+3} \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} i = p \\ -S_q \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{pmatrix} \quad (43)$$

a.2) Dados los valores de $\frac{\partial S_{q-p}}{\partial a_i}, \dots, \frac{\partial S_q}{\partial a_i}$,

para $i = 1, \dots, p$;

se obtuvieron los valores de $\frac{\partial S_0}{\partial a_i}, \dots, \frac{\partial S_{q-p-1}}{\partial a_i}$,

resolviendo el siguiente sistema de $(q-p)$ ecuaciones en forma recursiva

$$S_2^* A = S_2 \quad (44)$$

para $i = 1, \dots, p$, donde el vector S_2 , varía de la forma siguiente con respecto al valor de i

$$\begin{array}{cccc}
 i = 1 & i = 2 & \dots & i = p \\
 \left(\begin{array}{c} -S_1 \\ -S_2 \\ \cdot \\ \cdot \\ -S_{q-p} \end{array} \right) & \cdot \left(\begin{array}{c} -S_2 \\ -S_3 \\ \cdot \\ \cdot \\ -S_{q-p+1} \end{array} \right) & \cdot \dots \cdot & \left(\begin{array}{c} -S_p \\ -S_{p+1} \\ \cdot \\ \cdot \\ -S_{q-1} \end{array} \right)
 \end{array}
 \quad (45)$$

Finalmente, lo que se obtuvo fue la siguiente matriz:

$$\left(\begin{array}{ccc}
 \frac{\partial S_0}{\partial a_1} & \dots & \frac{\partial S_q}{\partial a_1} \\
 \cdot & & \cdot \\
 \cdot & & \cdot \\
 \cdot & & \cdot \\
 \frac{\partial S_0}{\partial a_p} & \dots & \frac{\partial S_q}{\partial a_p}
 \end{array} \right)_{p \times (q+1)} \quad (46)$$

b) $p = q$

para $0 \leq k \leq q$ se tiene el siguiente sistema triangular

$$\begin{pmatrix} \frac{\partial S_0}{\partial a_i} & \frac{\partial S_i}{\partial a_i} & \dots & \frac{\partial S_q}{\partial a_i} \\ \frac{\partial S_{q-1}}{\partial a_i} & \frac{\partial S_q}{\partial a_i} & \dots & 0 \\ \frac{\partial S_q}{\partial a_i} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_{p-1} \\ a_p \end{pmatrix} = \begin{pmatrix} -S_i \\ \vdots \\ -S_{i+q-1} \\ -S_{i+q} \end{pmatrix} \quad (47)$$

b.1) Para obtener los valores de $\frac{\partial S_0}{\partial a_i}, \dots, \frac{\partial S_q}{\partial a_i}$,

se resolvió el sistema triangular en (47) de dimensión $(q+1) \times (q+1)$, para $i = 1, \dots, p$, donde el vector en la derecha, varía con respecto al valor de i en la forma siguiente

$$\begin{array}{cccc}
 i = 1 & i = 2 & \dots & i = p \\
 \left(\begin{array}{c} -S_1 \\ -S_2 \\ \cdot \\ \cdot \\ \cdot \\ -S_q \\ 0 \end{array} \right) & , \left(\begin{array}{c} -S_2 \\ -S_3 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{array} \right) & , \dots , & \left(\begin{array}{c} -S_p = -S_q \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{array} \right)
 \end{array} \quad (48)$$

Al final se obtuvo una matriz igual a (46).

c) $p > q$

c.1) Para obtener los valores de $\frac{\partial S_0}{\partial a_i}, \dots, \frac{\partial S_q}{\partial a_i}$,

se resolvió un sistema triangular de dimensión $(q+1) \times (q+1)$, como en (47), pero ahora, sólo intervienen los parámetros autorregresivos a_0, \dots, a_q , y el vector en la derecha varía para $1 \leq i \leq q$ como en (48), y es igual al vector cero para $q+1 \leq i \leq p$. Al final se obtuvo una matriz como (46).

Para encontrar las primeras p columnas de la matriz jacobiana del vector $\gamma(\theta)$, en (39), lo que sigue es

resolver el sistema de ecuaciones en (16), el cual estaba dado como

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial a_i} = \begin{cases} \frac{\partial S_u}{\partial a_i} - \gamma_{u-i} & \text{para } u \leq q \\ -\gamma_{u-i} & \text{e.o.c.} \end{cases}$$

dados los parámetros autorregresivos a_0, a_1, \dots, a_p , con $a_0 = 1$, la matriz en (46), y el vector $\gamma(\theta)$ obtenido en 3.3.3.

La forma de resolver el sistema de ecuaciones en (16), varía de acuerdo a los ordenes p y q del modelo, de la manera siguiente:

a) $p < q$

a.1) Se tenía que resolver el siguiente sistema,

para obtener los valores de $\frac{\partial \gamma_0}{\partial a_i}, \dots, \frac{\partial \gamma_p}{\partial a_i}$.

$$\begin{array}{c}
 \left[\begin{array}{ccc}
 \frac{\partial \gamma_0}{\partial a_i} & \frac{\partial \gamma_{-1}}{\partial a_i} & \dots & \frac{\partial \gamma_{-p}}{\partial a_i} \\
 \frac{\partial \gamma_1}{\partial a_i} & \frac{\partial \gamma_0}{\partial a_i} & \dots & \frac{\partial \gamma_{1-p}}{\partial a_i} \\
 \frac{\partial \gamma_p}{\partial a_i} & \frac{\partial \gamma_{p-1}}{\partial a_i} & \dots & \frac{\partial \gamma_0}{\partial a_i}
 \end{array} \right]
 \begin{array}{c}
 a_0 \\
 a_1 \\
 \vdots \\
 a_p
 \end{array}
 =
 \begin{array}{c}
 \left[\begin{array}{c}
 \frac{\partial S_0}{\partial a_i} - \gamma_{-1} \\
 \frac{\partial S_1}{\partial a_i} - \gamma_0 \\
 \vdots \\
 \frac{\partial S_p}{\partial a_i} - \gamma_{p-1}
 \end{array} \right]
 \end{array}
 \end{array}
 \quad (49)$$

es decir, $A X = b$, en este caso la matriz A , es una matriz Toeplitz cuyos elementos son los valores desconocidos por calcular.

Al igual que se hizo en 3.3.3, se resolvió un sistema equivalente a (49) de dimensión $(2p+1) \times (2p+1)$, agregando p restricciones como en (32), pero ahora sobre las derivadas del vector $\gamma(\theta)$, es decir;

$$\frac{\partial \gamma_k}{\partial a_i} = \frac{\partial \gamma_k}{\partial a_i} \quad \text{para } k = 1, \dots, p \quad (50)$$

El nuevo sistema que se resolvió fue

$$\begin{pmatrix}
 a_p & a_{p-1} & \dots & a_0 & 0 & \dots & 0 \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 0 & 0 & \dots & a_p & a_{p-1} & \dots & a_0 \\
 1 & 0 & \dots & 0 & 0 & \dots & -1 \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 0 & 0 & \dots & 1 & 0 & -1 & \dots & 0
 \end{pmatrix}
 \begin{pmatrix}
 \frac{\partial \gamma_p}{\partial a_i} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial \gamma_0}{\partial a_i} \\
 \frac{\partial \gamma_1}{\partial a_i} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial \gamma_p}{\partial a_i}
 \end{pmatrix}
 =
 \begin{pmatrix}
 \frac{\partial S_0}{\partial a_i} - \gamma_{-i} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial S_p}{\partial a_i} - \gamma_{p-i} \\
 0 \\
 \cdot \\
 \cdot \\
 \cdot \\
 0
 \end{pmatrix}
 \begin{matrix}
 \left. \vphantom{\begin{matrix} \frac{\partial S_0}{\partial a_i} - \gamma_{-i} \\ \cdot \\ \cdot \\ \cdot \end{matrix}} \right\} p+1 \\
 \left. \vphantom{\begin{matrix} 0 \\ \cdot \\ \cdot \\ \cdot \end{matrix}} \right\} p \\
 (51)
 \end{matrix}$$

para $i = 1, \dots, p$.

Dado que $\gamma_k = \gamma_{-k}$, el vector en la derecha varía con respecto a i , en la forma siguiente

$$\begin{array}{cccc}
 i = 1 & & i = 2 & \dots & i = p \\
 \left[\begin{array}{c} \frac{\partial S_0}{\partial a_1} - \gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial S_p}{\partial a_1} - \gamma_{p-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{array} \right] & , & \left[\begin{array}{c} \frac{\partial S_0}{\partial a_2} - \gamma_2 \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial S_p}{\partial a_2} - \gamma_{p-2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{array} \right] & , \dots , & \left[\begin{array}{c} \frac{\partial S_0}{\partial a_p} - \gamma_p \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial S_p}{\partial a_p} - \gamma_0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{array} \right]
 \end{array}$$

a.2) Para encontrar los valores de $\frac{\partial \gamma_{p+1}}{\partial a_i}, \dots, \frac{\partial \gamma_q}{\partial a_i}$,

dadas $\frac{\partial \gamma_0}{\partial a_i}, \dots, \frac{\partial \gamma_p}{\partial a_i}$,

se resolvió en forma recursiva el siguiente sistema

$$\begin{pmatrix} \frac{\partial \gamma_{p+1}}{\partial a_i} & \frac{\partial \gamma_p}{\partial a_i} & \dots & \frac{\partial \gamma_1}{\partial a_i} \\ \frac{\partial \gamma_{p+2}}{\partial a_i} & \frac{\partial \gamma_{p+1}}{\partial a_i} & \dots & \frac{\partial \gamma_2}{\partial a_i} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial \gamma_q}{\partial a_i} & \frac{\partial \gamma_{q-1}}{\partial a_i} & \dots & \frac{\partial \gamma_{q-p}}{\partial a_i} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} \frac{\partial S_{p+1}}{\partial a_i} - \gamma_{p+1-i} \\ \frac{\partial S_{p+2}}{\partial a_i} - \gamma_{p+2-i} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_i} - \gamma_{q-i} \end{pmatrix} \quad (52)$$

de dimensión $(q-p) \times (p+1)$, donde el vector en la derecha, varía de acuerdo al valor de i en la forma siguiente

$i = 1$

$$\begin{bmatrix} \frac{\partial S_{p+1}}{\partial a_1} - \gamma_p \\ \frac{\partial S_{p+2}}{\partial a_1} - \gamma_{p+1} \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_1} - \gamma_{q-1} \end{bmatrix}$$

 $i = 2$

$$\begin{bmatrix} \frac{\partial S_{p+1}}{\partial a_2} - \gamma_{p-1} \\ \frac{\partial S_{p+2}}{\partial a_2} - \gamma_{p-2} \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_2} - \gamma_{q-2} \end{bmatrix}$$

 \dots $i = p$

$$\begin{bmatrix} \frac{\partial S_{p+1}}{\partial a_p} - \gamma_1 \\ \frac{\partial S_{p+2}}{\partial a_p} - \gamma_2 \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_p} - \gamma_{q-p} \end{bmatrix}$$

a.3) Si $Lag = k =$ orden de la autocovarianza requerida por el usuario en 3.3.3, es mayor que q , entonces, para encontrar

$$\frac{\partial \gamma_{q+1}}{\partial a_i}, \dots, \frac{\partial \gamma_k}{\partial a_i}, \text{ dadas } \frac{\partial \gamma_{p+1}}{\partial a_i}, \dots, \frac{\partial \gamma_q}{\partial a_i},$$

se resolvió en forma recursiva el siguiente sistema

$$\begin{pmatrix} \frac{\partial \gamma_{q+1}}{\partial a_i} & \frac{\partial \gamma_q}{\partial a_i} & \dots & \frac{\partial \gamma_{q+1-p}}{\partial a_i} \\ \frac{\partial \gamma_{q+2}}{\partial a_i} & \frac{\partial \gamma_{q+1}}{\partial a_i} & \dots & \frac{\partial \gamma_{q+2-p}}{\partial a_i} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial \gamma_k}{\partial a_i} & \frac{\partial \gamma_{k-1}}{\partial a_i} & \dots & \frac{\partial \gamma_{k-p}}{\partial a_i} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} -\gamma_{q+1-i} \\ -\gamma_{q+2-i} \\ \cdot \\ \cdot \\ -\gamma_{k-i} \end{pmatrix} \quad (53)$$

de dimensión $(k-q) \times (p+1)$, donde el vector en la derecha varía con respecto al valor de i , en la forma siguiente

$$\begin{array}{cccc}
 i = 1 & & i = 2 & \dots & i = p \\
 \left[\begin{array}{c} -\gamma_q \\ -\gamma_{q+1} \\ \cdot \\ \cdot \\ -\gamma_{k-1} \end{array} \right] & , & \left[\begin{array}{c} -\gamma_{q-1} \\ -\gamma_q \\ \cdot \\ \cdot \\ -\gamma_{k-2} \end{array} \right] & , \dots , & \left[\begin{array}{c} -\gamma_{q+i-p} \\ -\gamma_{q+z-p} \\ \cdot \\ \cdot \\ -\gamma_{k-p} \end{array} \right]
 \end{array}$$

b) $p = q$

b.1) Para obtener los valores de $\frac{\partial \gamma_0}{\partial a_i}, \dots, \frac{\partial \gamma_p}{\partial a_i}$,

se resolvió un sistema parecido a (51), pero donde el vector en la derecha, varía con respecto al valor de i en la forma siguiente

$$\begin{array}{cccc}
 i = 1 & & i = 2 & \dots & i = p \\
 \left[\begin{array}{c} \frac{\partial S_0}{\partial a_1} - \gamma_1 \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_1} - \gamma_{q-1} \\ \cdot \\ \cdot \\ 0 \\ \cdot \\ \cdot \\ 0 \end{array} \right] & \cdot & \left[\begin{array}{c} \frac{\partial S_0}{\partial a_2} - \gamma_2 \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_2} - \gamma_{q-2} \\ \cdot \\ \cdot \\ 0 \\ \cdot \\ \cdot \\ 0 \end{array} \right] & \cdot \dots \cdot & \left[\begin{array}{c} \frac{\partial S_0}{\partial a_p} - \gamma_p \\ \cdot \\ \cdot \\ \frac{\partial S_q}{\partial a_p} - \gamma_0 \\ \cdot \\ \cdot \\ 0 \\ \cdot \\ \cdot \\ 0 \end{array} \right]
 \end{array}$$

b.2) Si $\text{Lag} = k$ es mayor que p , entonces, para encontrar

$$\frac{\partial \gamma_{p+1}}{\partial a_i}, \dots, \frac{\partial \gamma_k}{\partial a_i}, \text{ dadas } \frac{\partial \gamma_0}{\partial a_i}, \dots, \frac{\partial \gamma_p}{\partial a_i},$$

se resolvió en forma recursiva, un sistema igual a (53) con $q = p$.

c) $p > q$

c.1) Para obtener los valores de $\frac{\partial \gamma_0}{\partial a_i}, \dots, \frac{\partial \gamma_p}{\partial a_i}$,

se resolvió un sistema parecido a (51), pero donde el vector en la derecha, varía con respecto al valor de i en la forma siguiente

$$\frac{\partial \gamma_{p+1}}{\partial a_i}, \dots, \frac{\partial \gamma_k}{\partial a_i}, \text{ dadas } \frac{\partial \gamma_0}{\partial a_i}, \dots, \frac{\partial \gamma_p}{\partial a_i},$$

se resolvió en forma recursiva un sistema parecido a (53), pero ahora, p toma el papel de q y viceversa, dando como resultado un sistema de dimensión $(k-p) \times (p+1)$.

ii) Derivadas de $\gamma(\theta)$ con respecto a los parámetros de promedios móviles b_1, \dots, b_q .

En este punto, lo que interesa conocer, son las columnas de la matriz jacobiana de $\gamma(\theta)$, asociadas a los parámetros de promedios móviles, es decir; de la columna p+1 a la p+q

$$\begin{array}{ccc} \frac{\partial \gamma_0(\theta)}{\partial b_1} & \dots & \frac{\partial \gamma_0(\theta)}{\partial b_q} \\ \vdots & & \vdots \\ \frac{\gamma_k(\theta)}{\partial b_1} & \dots & \frac{\gamma_k(\theta)}{\partial b_q} \end{array}$$

(54)

Para obtener los valores de los elementos en estas columnas,

es necesario resolver los sistemas (17), (18) y (19) en este orden.

Recordando que (17) era igual a:

$$\frac{\partial du}{\partial b_j} = \begin{cases} \sigma_\epsilon^2 (b_{j+|u|} + b_{j-|u|}) & \text{si } \begin{matrix} 0 \leq j+|u| \leq q \\ 0 \leq j-|u| \leq q \end{matrix} \text{ y} \\ \sigma_\epsilon^2 b_{j+|u|} & \text{si } \begin{matrix} j-|u| < 0 \\ j+|u| \leq q \end{matrix} \text{ y} \\ \sigma_\epsilon^2 b_{j-|u|} & \text{si } \begin{matrix} j-|u| \geq 0 \\ j+|u| > q \end{matrix} \text{ y} \\ 0 & \text{si } \begin{matrix} j-|u| < 0 \\ j+|u| > q \end{matrix} \text{ y} \end{cases}$$

los cálculos para resolver (17) se llevan a cabo dentro de la subrutina CALDB, cuyo listado se encuentra en el apéndice B. Los valores que le llegan, son los estimadores iniciales de los parámetros σ_ϵ^2 , b_0, b_1, \dots, b_q , con $b_0 = 1$, y regresa los valores

$$\text{de } \frac{\partial d_{-q}}{\partial b_j}, \dots, \frac{\partial d_0}{\partial b_j}, \dots, \frac{\partial d_q}{\partial b_j} ; \text{ para } j = 1, \dots, q.$$

Dado que $d-u = du$ para u distinta de cero, sólo se necesita conocer la siguiente matriz de dimensión $(q+1) \times (q+1)$

$$\begin{pmatrix} \frac{\partial d_0}{\partial b_1} & \dots & \frac{\partial d_0}{\partial b_q} \\ \frac{\partial d_q}{\partial b_1} & \dots & \frac{\partial d_q}{\partial b_q} \end{pmatrix}$$

(55)

y los estimadores iniciales de a_0, a_1, \dots, a_p , con $a_0 = 1$, para resolver el sistema de ecuaciones en (18).

Recordando que (18) era igual a:

$$\sum_{k=0}^p a_k \frac{\partial S_{u+k}}{\partial b_j} = \begin{cases} \frac{\partial d_u}{\partial b_j} & \text{para } |u| \leq q \\ 0 & \text{en otro caso} \end{cases}$$

En forma equivalente, este sistema de ecuaciones se puede expresar como

$$\begin{pmatrix} \frac{\partial S_0}{\partial b_j} & \frac{\partial S_1}{\partial b_j} & \dots & \frac{\partial S_p}{\partial b_j} \\ \frac{\partial S_1}{\partial b_j} & \frac{\partial S_2}{\partial b_j} & \dots & \frac{\partial S_{p+1}}{\partial b_j} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial S_q}{\partial b_j} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} \frac{\partial d_0}{\partial b_j} \\ \frac{\partial d_1}{\partial b_j} \\ \cdot \\ \cdot \\ \frac{\partial d_q}{\partial b_j} \end{pmatrix}$$

(56)

Al igual que en los casos anteriores, la estructura de (56) cambia con respecto a las ordenes p y q del modelo. Analizando los tres casos siguientes:
 a) $p < q$; b) $p = q$; c) $p > q$, se tiene que:

a) $p < q$

a.1) Para obtener los valores de

$$\frac{\partial S_{q-p}}{\partial b_j}, \dots, \frac{\partial S_q}{\partial b_j}, \text{ dadas } \frac{\partial d_{q-p}}{\partial b_j}, \dots, \frac{\partial d_q}{\partial b_j}.$$

se resolvió el siguiente sistema triangular, de dimensión $(p+1) \times (p+1)$.

$$\begin{pmatrix}
 \frac{\partial S_{q-p}}{\partial b_j} & \frac{\partial S_{q-p+1}}{\partial b_j} & \dots & \frac{\partial S_{q-1}}{\partial b_j} & \frac{\partial S_q}{\partial b_j} \\
 \frac{\partial S_{q-p+1}}{\partial b_j} & \frac{\partial S_{q-p+2}}{\partial b_j} & \dots & \frac{\partial S_q}{\partial b_j} & 0 \\
 \vdots & \vdots & & \vdots & \vdots \\
 \frac{\partial S_{q-1}}{\partial b_j} & \frac{\partial S_q}{\partial b_j} & \dots & 0 & 0 \\
 \frac{\partial S_q}{\partial b_j} & 0 & \dots & 0 & 0
 \end{pmatrix}
 \begin{pmatrix}
 a_0 \\
 a_1 \\
 \vdots \\
 a_{p-1} \\
 a_p
 \end{pmatrix}
 =
 \begin{pmatrix}
 \frac{\partial d_{q-p}}{\partial b_j} \\
 \frac{\partial d_{q-p+1}}{\partial b_j} \\
 \vdots \\
 \frac{\partial d_{q-1}}{\partial b_j} \\
 \frac{\partial d_q}{\partial b_j}
 \end{pmatrix}$$

para $j = 1, \dots, q$.

a.2) Para obtener los valores de $\frac{\partial S_0}{\partial b_j}, \dots, \frac{\partial S_{q-p-1}}{\partial b_j}$,

dadas $\frac{\partial S_{q-p}}{\partial b_j}, \dots, \frac{\partial S_q}{\partial b_j}$ y $\frac{\partial d_0}{\partial b_j}, \dots, \frac{\partial d_{q-p-1}}{\partial b_j}$.

se resolvió en forma recursiva, el siguiente sistema de dimensión $(q-p) \times (p+1)$

$$\begin{pmatrix} \frac{\partial S_0}{\partial b_j} & \frac{\partial S_1}{\partial b_j} & \dots & \frac{\partial S_p}{\partial b_j} \\ \frac{\partial S_1}{\partial b_j} & \frac{\partial S_2}{\partial b_j} & \dots & \frac{\partial S_{p+1}}{\partial b_j} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial S_{q-p-1}}{\partial b_j} & \frac{\partial S_{q-p}}{\partial b_j} & \dots & \frac{\partial S_{q-1}}{\partial b_j} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \frac{\partial d_0}{\partial b_j} \\ \frac{\partial d_1}{\partial b_j} \\ \vdots \\ \frac{\partial d_{q-p-1}}{\partial b_j} \end{pmatrix}$$

para $j = 1, \dots, q$.

b) $p = q$

b.1) Para obtener los valores de

$$\frac{\partial S_0}{\partial b_j}, \dots, \frac{\partial S_q}{\partial b_j}, \text{ dadas } \frac{\partial d_0}{\partial b_j}, \dots, \frac{\partial d_q}{\partial b_j},$$

se resolvió el siguiente sistema triangular

$$\begin{bmatrix} \frac{\partial S_0}{\partial b_j} & \frac{\partial S_1}{\partial b_j} & \dots & \frac{\partial S_{q-1}}{\partial b_j} & \frac{\partial S_q}{\partial b_j} \\ \frac{\partial S_1}{\partial b_j} & \frac{\partial S_2}{\partial b_j} & \dots & \frac{\partial S_q}{\partial b_j} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial S_{q-1}}{\partial b_j} & \frac{\partial S_q}{\partial b_j} & \dots & 0 & 0 \\ \frac{\partial S_q}{\partial b_j} & 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \\ a_p \end{bmatrix} = \begin{bmatrix} \frac{\partial d_0}{\partial b_j} \\ \frac{\partial d_1}{\partial b_j} \\ \vdots \\ \frac{\partial d_{q-1}}{\partial b_j} \\ \frac{\partial d_q}{\partial b_j} \end{bmatrix} \quad (57)$$

para $j = 1, \dots, q$.

c) $p > q$

c.1) Para obtener los valores de

$$\frac{\partial S_0}{\partial b_j}, \dots, \frac{\partial S_q}{\partial b_j}, \text{ dadas } \frac{\partial d_0}{\partial b_j}, \dots, \frac{\partial d_q}{\partial b_j},$$

se resolvió un sistema triangular de dimensión $(q+1) \times (q+1)$, similar a (57), pero donde sólo intervienen los parámetros autorregresivos a_0, \dots, a_q , para $j = 1, \dots, q$.

Finalmente se obtuvo la siguiente matriz:

$$\begin{pmatrix} \frac{\partial S_0}{\partial b_1} & \frac{\partial S_1}{\partial b_1} & \dots & \frac{\partial S_q}{\partial b_1} \\ \frac{\partial S_0}{\partial b_2} & \frac{\partial S_1}{\partial b_2} & \dots & \frac{\partial S_q}{\partial b_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial S_0}{\partial b_q} & \frac{\partial S_1}{\partial b_q} & \dots & \frac{\partial S_q}{\partial b_q} \end{pmatrix} \quad (58)$$

la cual, se utilizó para resolver (19) como se explica a continuación, para los tres casos a) $p < q$, b) $p = q$, y c) $p > q$.

Recordando que (19) era igual a:

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial b_j} = \begin{cases} \frac{\partial S_u}{\partial b_j} & \text{para } u \leq q \\ 0 & \text{en otro caso} \end{cases}$$

a) $p < q$

a.1) Se tenía que resolver el siguiente sistema,

para obtener los valores de $\frac{\partial \gamma_0}{\partial b_j}, \dots, \frac{\partial \gamma_p}{\partial b_j}$,

dados los estimadores de los parámetros autorregresivos, y la matriz en (58).

$$\begin{pmatrix}
 \frac{\partial \gamma_0}{\partial b_j} & \frac{\partial \gamma_1}{\partial b_j} & \dots & \frac{\partial \gamma_p}{\partial b_j} \\
 \frac{\partial \gamma_1}{\partial b_j} & \frac{\partial \gamma_0}{\partial b_j} & \dots & \frac{\partial \gamma_{1-p}}{\partial b_j} \\
 \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot \\
 \frac{\partial \gamma_p}{\partial b_j} & \frac{\partial \gamma_{p-1}}{\partial b_j} & \dots & \frac{\partial \gamma_0}{\partial b_j}
 \end{pmatrix}
 \begin{pmatrix}
 a_0 \\
 a_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 a_p
 \end{pmatrix}
 =
 \begin{pmatrix}
 \frac{\partial s_0}{\partial b_j} \\
 \frac{\partial s_1}{\partial b_j} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial s_p}{\partial b_j}
 \end{pmatrix}
 \quad (59)$$

para $j = 1, \dots, q$.

Tal como se hizo, en el caso en que se obtenían las derivadas de $\gamma(\theta)$ con respecto a los parámetros

autorregresivos, se resolvió un sistema equivalente a (59), de dimensión $(2p+1) \times (2p+1)$ de la forma siguiente

$$\begin{pmatrix}
 a_p & a_{p-1} & \dots & a_0 & 0 & \dots & 0 \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 0 & 0 & \dots & a_p & a_{p-1} & \dots & a_0 \\
 1 & 0 & \dots & 0 & 0 & \dots & -1 \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & & \cdot \\
 0 & 0 & \dots & 1 & 0 & -1 & \dots & 0
 \end{pmatrix}
 \begin{pmatrix}
 \frac{\partial \gamma_p}{\partial b_j} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial \gamma_0}{\partial b_j} \\
 \frac{\partial \gamma_1}{\partial b_j} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial \gamma_p}{\partial b_j}
 \end{pmatrix}
 =
 \begin{pmatrix}
 \frac{\partial S_0}{\partial b_j} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial S_p}{\partial b_j} \\
 0 \\
 \cdot \\
 \cdot \\
 \cdot \\
 0
 \end{pmatrix}
 \begin{matrix}
 p+1 \\
 \\
 \\
 \\
 \\
 p \\
 \\
 \\
 \\
 \end{matrix}
 \quad (60)$$

a.2) Para obtener los valores de

$$\frac{\partial \gamma_{p+1}}{\partial b_j}, \dots, \frac{\partial \gamma_q}{\partial b_j}, \text{ dadas } \frac{\partial \gamma_0}{\partial b_j}, \dots, \frac{\partial \gamma_p}{\partial b_j},$$

se resolvió el siguiente sistema de $(q-p)$ ecuaciones, en forma recursiva

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial b_j} = \frac{\partial S_u}{\partial b_j} \quad \text{para } p+1 \leq u \leq q$$

a.3) Si $\text{Lag} = k$ es mayor que q , entonces para encontrar

$$\frac{\partial \gamma_{q+1}}{\partial b_j}, \dots, \frac{\partial \gamma_k}{\partial b_j}, \text{ dadas } \frac{\partial \gamma_{q-p+1}}{\partial b_j}, \dots, \frac{\partial \gamma_q}{\partial b_j},$$

se resolvió el siguiente sistema homogéneo de $(k-q)$ ecuaciones

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial b_j} = 0 \quad \text{para } q+1 \leq u \leq k$$

b) $p = q$

b.1) Para obtener los valores de $\frac{\partial \gamma_0}{\partial b_j}, \dots, \frac{\partial \gamma_p}{\partial b_j}$,

se resolvió un sistema de dimensión $(2p+1) \times (2p+1)$ como en (60).

b.2) Si $Lag = k$ era mayor que p , entonces para encontrar

$$\frac{\partial \gamma_{p+1}}{\partial b_j}, \dots, \frac{\partial \gamma_k}{\partial b_j}, \quad \text{dadas} \quad \frac{\partial \gamma_0}{\partial b_j}, \dots, \frac{\partial \gamma_p}{\partial b_j},$$

se resolvió en forma recursiva el siguiente sistema homogéneo de $(k-p)$ ecuaciones.

$$\sum_{v=0}^p a_v \frac{\partial \chi_{u-v}}{\partial b_j} = 0 \quad \text{para } p+1 \leq u \leq k \quad (61)$$

c) $p > q$

c.1) Para obtener los valores de $\frac{\partial \chi_0}{\partial b_j}, \dots, \frac{\partial \chi_p}{\partial b_j}$,

se resolvió un sistema de dimensión $(2p+1) \times (2p+1)$ similar a (60), pero ahora, el vector en la derecha de (60) tiene como primeros $(q+1)$ elementos a

$$\frac{\partial S_0}{\partial b_j}, \dots, \frac{\partial S_q}{\partial b_j},$$

y ceros para los $(2p-q)$ elementos restantes.

c.2) Si $\text{Lag} = k$ era mayor que p , entonces para encontrar

$$\frac{\partial \gamma_{p+1}}{\partial b_j}, \dots, \frac{\partial \gamma_k}{\partial b_j}, \text{ dadas } \frac{\partial \gamma_0}{\partial b_j}, \dots, \frac{\partial \gamma_p}{\partial b_j},$$

se resolvió en forma recursiva, un sistema homogéneo de $(k-p)$ ecuaciones como en (61).

iii) Derivadas de $\gamma(\theta)$ con respecto a la varianza residual σ_ϵ^2 .

Por último, falta encontrar la columna $p+q+1$ de la matriz jacobiana del vector $\gamma(\theta)$ en (39), en donde se encuentran las derivadas de las autocovarianzas con respecto al parámetro de la varianza residual.

Los sistemas de ecuaciones en diferencias, por resolver en forma recursiva son; (20), (21), y (22) para finalmente obtener el vector

$$\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \frac{\partial \gamma_1}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_k}{\partial \sigma_\epsilon^2} \quad (62)$$

Comenzando con (20), en donde se tenía la siguiente relación

$$\frac{\partial d_k}{\partial \sigma_\epsilon^2} = \begin{cases} \frac{\partial C_u}{\partial \sigma_\epsilon^2} = du/\sigma_\epsilon^2 & \text{para } |k| \leq q \\ 0 & \text{e.o.c.} \end{cases}$$

Dado que los valores de $d_{-q}, \dots, d_0, \dots, d_q$, y un estimador inicial del parámetro σ_ϵ^2 se obtuvieron en 3.3.3, obtener

$$\frac{\partial d_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial d_q}{\partial \sigma_\epsilon^2}, \text{ no representa ningún problema.}$$

Continuando con (21), en donde se tenía

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial \sigma_\epsilon^2} = \begin{cases} \partial d_k / \partial \sigma_\epsilon^2 & \text{para } |k| \leq q \\ 0 & \text{e.o.c.,} \end{cases}$$

el procedimiento para encontrar los valores de

$$\frac{\partial S_0}{\partial \sigma_\epsilon^2} , \frac{\partial S_1}{\partial \sigma_\epsilon^2} , \dots , \frac{\partial S_q}{\partial \sigma_\epsilon^2}$$

varía como antes para los tres casos siguientes:

a) $p < q$

a.1) Dados el vector de parámetros iniciales autorregresivos a_0, a_1, \dots, a_p , con $a_0 = 1$,

y el vector $\frac{\partial d_{q-p}}{\partial \sigma_\epsilon^2} , \dots , \frac{\partial d_q}{\partial \sigma_\epsilon^2}$,

se resolvió el siguiente sistema triangular, de dimensión $(p+1) \times (p+1)$,

para obtener $\frac{\partial S_{q-p}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_q}{\partial \sigma_\epsilon^2}$.

$$\begin{pmatrix}
 \frac{\partial S_{q-p}}{\partial \sigma_\epsilon^2} & \frac{\partial S_{q-p+1}}{\partial \sigma_\epsilon^2} & \dots & \frac{\partial S_{q-1}}{\partial \sigma_\epsilon^2} & \frac{\partial S_q}{\partial \sigma_\epsilon^2} \\
 \frac{\partial S_{q-p-1}}{\partial \sigma_\epsilon^2} & \frac{\partial S_{q-p}}{\partial \sigma_\epsilon^2} & \dots & \frac{\partial S_q}{\partial \sigma_\epsilon^2} & 0 \\
 \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot \\
 \frac{\partial S_{q-1}}{\partial \sigma_\epsilon^2} & \frac{\partial S_q}{\partial \sigma_\epsilon^2} & \dots & 0 & 0 \\
 \frac{\partial S_q}{\partial \sigma_\epsilon^2} & 0 & \dots & 0 & 0
 \end{pmatrix}
 \begin{pmatrix}
 a_0 \\
 a_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 a_{p-1} \\
 a_p
 \end{pmatrix}
 =
 \begin{pmatrix}
 \frac{\partial d_{q-p}}{\partial \sigma_\epsilon^2} \\
 \frac{\partial d_{q-p+1}}{\partial \sigma_\epsilon^2} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial d_{q-1}}{\partial \sigma_\epsilon^2} \\
 \frac{\partial d_q}{\partial \sigma_\epsilon^2}
 \end{pmatrix}$$

a.2) Dados el vector $\frac{\partial d_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial d_{q-p-1}}{\partial \sigma_\epsilon^2}$, y los

valores de $\frac{\partial S_{q-p}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_{q-1}}{\partial \sigma_\epsilon^2}$ obtenidos en (a.1),

se resolvió en forma recursiva el siguiente sistema de (q-p)

ecuaciones, para obtener $\frac{\partial S_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_{q-p-1}}{\partial \sigma_\epsilon^2}$

$$\sum_{v=0}^p a_v \frac{\partial S_{k+v}}{\partial \sigma_\epsilon^2} = \frac{\partial d_k}{\partial \sigma_\epsilon^2} \quad \text{para } 0 \leq k \leq q-p-1$$

b) p = q

b.1) Dado el vector de parámetros iniciales autorregresivos a_0, a_1, \dots, a_p , con $a_0 = 1$,

y los valores de $\frac{\partial d_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial d_q}{\partial \sigma_\epsilon^2}$,

se resolvió el siguiente sistema triangular de dimensión $(p+1) \times (p+1)$,

para obtener los valores de $\frac{\partial S_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_q}{\partial \sigma_\epsilon^2}$.

$$\begin{pmatrix} \frac{\partial S_0}{\partial \sigma_\epsilon^2} & \frac{\partial S_1}{\partial \sigma_\epsilon^2} & \dots & \frac{\partial S_{q-1}}{\partial \sigma_\epsilon^2} & \frac{\partial S_q}{\partial \sigma_\epsilon^2} \\ \frac{\partial S_1}{\partial \sigma_\epsilon^2} & \frac{\partial S_2}{\partial \sigma_\epsilon^2} & \dots & \frac{\partial S_q}{\partial \sigma_\epsilon^2} & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \frac{\partial S_{q-1}}{\partial \sigma_\epsilon^2} & \frac{\partial S_q}{\partial \sigma_\epsilon^2} & \dots & 0 & 0 \\ \frac{\partial S_q}{\partial \sigma_\epsilon^2} & 0 & \dots & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_{p-1} \\ a_p \end{pmatrix} = \begin{pmatrix} \frac{\partial d_0}{\partial \sigma_\epsilon^2} \\ \frac{\partial d_1}{\partial \sigma_\epsilon^2} \\ \cdot \\ \cdot \\ \frac{\partial d_{q-1}}{\partial \sigma_\epsilon^2} \\ \frac{\partial d_q}{\partial \sigma_\epsilon^2} \end{pmatrix}$$

(63)

c) $p > q$

c.1) Dado el vector de parámetros iniciales autorregresivos a_0, a_1, \dots, a_q , con $a_0 = 1$, se resolvió un sistema triangular de dimensión $(q+1) \times (q+1)$ similar a (63),

para obtener los valores de $\frac{\partial S_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_q}{\partial \sigma_\epsilon^2}$.

Para encontrar los valores que interesa calcular, se hizo uso de (22), en donde se tenía

$$\sum_{v=0}^p a_v \frac{\partial \chi_{u-v}}{\partial \sigma_\epsilon^2} = \begin{cases} \frac{\partial S_u}{\partial \sigma_\epsilon^2} & \text{para } u \leq q \\ 0 & \text{e.o.c.} \end{cases}$$

Analizando los tres casos por separado a) $p < q$, b) $p = q$, y c) $p > q$, para obtener (62) se tiene que:

a) $p < q$

a.1) Dado el vector de parámetros iniciales autorregresivos a_0, a_1, \dots, a_p , con $a_0 = 1$,

y los valores de $\frac{\partial S_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_p}{\partial \sigma_\epsilon^2}$,

se tenía que resolver el siguiente sistema,

para encontrar los valores de $\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2}$.

$$\begin{pmatrix} \frac{\partial \gamma_0}{\partial \sigma_\epsilon^2} & & & \\ \frac{\partial \gamma_1}{\partial \sigma_\epsilon^2} & \frac{\partial \gamma_0}{\partial \sigma_\epsilon^2} & & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2} & \frac{\partial \gamma_{p-1}}{\partial \sigma_\epsilon^2} & \dots & \frac{\partial \gamma_0}{\partial \sigma_\epsilon^2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} \frac{\partial S_0}{\partial \sigma_\epsilon^2} \\ \frac{\partial S_1}{\partial \sigma_\epsilon^2} \\ \cdot \\ \cdot \\ \frac{\partial S_p}{\partial \sigma_\epsilon^2} \end{pmatrix}$$

Al igual que se ha venido haciendo para obtener el vector de autocovarianzas $\gamma(\theta)$, y las derivadas de éste con respecto a los otros parámetros, se resolvió el siguiente sistema equivalente, de dimensión $(2p+1) \times (2p+1)$.

$$\begin{pmatrix}
 a_p & a_{p-1} & \dots & a_0 & 0 & \dots & 0 \\
 \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
 0 & 0 & \dots & a_p & a_{p-1} & \dots & a_0 \\
 1 & 0 & \dots & 0 & 0 & \dots & -1 \\
 \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
 0 & 0 & \dots & 1 & 0 & -1 & \dots & 0
 \end{pmatrix}
 \begin{pmatrix}
 \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial \gamma_0}{\partial \sigma_\epsilon^2} \\
 \frac{\partial \gamma_1}{\partial \sigma_\epsilon^2} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2}
 \end{pmatrix}
 =
 \begin{pmatrix}
 \frac{\partial S_0}{\partial \sigma_\epsilon^2} \\
 \cdot \\
 \cdot \\
 \cdot \\
 \frac{\partial S_p}{\partial \sigma_\epsilon^2} \\
 0 \\
 \cdot \\
 \cdot \\
 \cdot \\
 0
 \end{pmatrix}
 \begin{matrix}
 p+1 \\
 \\
 \\
 \\
 \\
 \\
 p \\
 \\
 \\
 \end{matrix}$$

(64)

a.2) Dados los valores de $\frac{\partial S_{p+1}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_q}{\partial \sigma_\epsilon^2}$,

y de $\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2}$ obtenidos en (a.1),

se resolvió en forma recursiva, el siguiente sistema de (q-p) ecuaciones,

para obtener los valores de $\frac{\partial \gamma_{p+1}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_q}{\partial \sigma_\epsilon^2}$.

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial \sigma_\epsilon^2} = \frac{\partial S_u}{\partial \sigma_\epsilon^2} \quad \text{para } p+1 \leq u \leq q$$

a.3) Si Lag = k = número de autocovarianzas obtenidas en 3.3.3, es mayor que q, entonces,

dados los valores de $\frac{\partial \gamma_{q-p}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_q}{\partial \sigma_\epsilon^2}$,

se resolvió en forma recursiva, el siguiente sistema homogéneo de (k-q) ecuaciones

para obtener los valores de $\frac{\partial \gamma_{q+1}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_k}{\partial \sigma_\epsilon^2}$.

$$\sum_{v=0}^p a_v \frac{\partial \gamma_{u-v}}{\partial \sigma_\epsilon^2} = 0 \quad \text{para } q+1 \leq u \leq k \quad (65)$$

b) p = q

b.1) Dado el vector de parámetros iniciales autorregresivos a_0, a_1, \dots, a_p , con $a_0 = 1$,

y los valores de $\frac{\partial S_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_q}{\partial \sigma_\epsilon^2}$.

se resolvió un sistema de dimensión $(2p+1) \times (2p+1)$, similar a (64),

para obtener los valores de $\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_q}{\partial \sigma_\epsilon^2}$.

b.2) Si $\text{Lag} = k$ es mayor que q , entonces,

dados los valores de $\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_q}{\partial \sigma_\epsilon^2}$,

se resolvió en forma recursiva un sistema homogéneo de $(k-q)$ ecuaciones en diferencias como en (65),

para obtener los valores de $\frac{\partial \gamma_{q+1}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_k}{\partial \sigma_\epsilon^2}$.

c) $p > q$

c.1) Dados el vector de parámetros iniciales autorregresivos a_0, a_1, \dots, a_p , con $a_0 = 1$,

y los valores de $\frac{\partial S_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial S_q}{\partial \sigma_\epsilon^2}$.

se resolvió un sistema de dimensión $(2p+1) \times (2p+1)$, similar a (64), pero donde el vector en la derecha, ahora contiene $(2p-q)$ elementos iguales a cero,

para obtener los valores de $\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2}$.

c.2) Si $Lag = k$ era mayor que p , entonces,

dados los valores de $\frac{\partial \gamma_0}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_p}{\partial \sigma_\epsilon^2}$,

se resolvió en forma recursiva, un sistema homogéneo de $(k-p)$ ecuaciones en diferencias como en (65), para $p+1 \leq u \leq k$,

con lo que se obtuvieron los valores para $\frac{\partial \gamma_{p+1}}{\partial \sigma_\epsilon^2}, \dots, \frac{\partial \gamma_k}{\partial \sigma_\epsilon^2}$.

Todos los cálculos para los sistemas presentados en

3.3.3 y 3.3.4, se llevan a cabo en la subrutina FCN, cuyo listado se encuentra en el apéndice B.

En virtud de que los sistemas de ecuaciones en diferencias, para obtener los valores numéricos del vector de autocovarianzas $\gamma(\theta)$, así como de los elementos de su matriz jacobiana J_p , mantienen en general una estructura muy similar, que difiere principalmente en las dimensiones de los sistemas involucrados, y en los elementos conocidos o por determinar, dependiendo de la variación que exista entre los órdenes p y q del modelo ARMA que se considere, y del índice del parámetro con respecto al cual se esté derivando, fue posible crear una serie de subrutinas bastante generales, logrando con ésto un ahorro considerable de espacio en la memoria de la computadora.

Capítulo 4

RESULTADOS

Con objeto de probar qué tan bien se comportan los estimadores, obtenidos con el método de estimación de parámetros de modelos ARMA(p,q), descrito en el capítulo 2, sección 2.2; se programó en una computadora Burroughs 7800 el algoritmo descrito en el capítulo 3, sección 3.3 (un listado de dicho programa se encuentra en el apéndice B); y se compararon dichos estimadores, con los obtenidos con el método tradicional de Box y Jenkins. En ambos casos se obvió el procedimiento de identificación, indicando a los paquetes los ordenes p y q de los modelos con que se generaron los datos.

Para ambos métodos, se simularon muestras de tamaños 25, 100 y 300 para 4 modelos ARMA : (0,2), (2,0), (1,1) y (2,2). Los cuatro conjuntos de parámetros asociados a los modelos anteriores fueron: $(a_1, a_2, b_1, b_2, \sigma_c^2) = (0, 0, .11, -.14, 2), (.4, .45, 0, 0, 1),$

(.91,0,.5,0,1) y (.5,-.17,.25,.13,3). Para cada una de las 12 combinaciones entre modelos y tamaños de muestra, se simularon 300 muestras, utilizando la subrutina FTGEN del paquete IMSL, la cual genera series de tiempo a partir de un modelo ARIMA(p,d,q) dado, en todos los casos, d se hizo igual a cero. Para cada tamaño de muestra TM, se impusieron condiciones iniciales iguales a cero para generar las primeras 100 + TM observaciones, descartando las primeras 100. A partir de aquí, se utilizaron las últimas observaciones de cada muestra generada, para generar 50 + TM observaciones, descartando en cada vez las primeras 50.

Para el primer método de estimación, el número k, de autocovarianzas muestrales, a las cuales el modelo se ajustó, fue de 6 y 15 para muestras de tamaño 25 y de 25 para muestras de tamaños 100 y 300. Estas autocovarianzas muestrales se obtuvieron con la subrutina FTAUTO del paquete IMSL, bajo el supuesto de que la media del proceso $\{X_t\}$, era conocida e igual a cero.

En ambos métodos, se obtuvieron estimadores iniciales de los parámetros, a partir del método descrito en el capítulo 2, sección 2.1.4, por medio de las subrutinas FTARPS y FTMPs del paquete IMSL. Cuando ésta última no convergió, para encontrar valores aceptables de los estimadores de los parámetros de promedios móviles, se dio el valor inicial arbitrario de 0.5 a los parámetros de promedios móviles, o de 0.4 cuando el orden q era igual a 2, con objeto de evitar raíces sobre el círculo unitario. Así

mismo, se dio el valor inicial arbitrario de 0.5 al parámetro de la varianza residual, cuando el orden q era igual a cero, o cuando FTMPS no convergió. En la tabla 4.6 aparece el número de veces de entre 300, en que FTMPS no convergió, para los modelos ARMA (0,2), (1,1), (2,2). Se verificó que los parámetros iniciales obtenidos satisficieran las condiciones de estacionariedad e invertibilidad, por medio de la subrutina REST, cuyo algoritmo se explicó en la sección 3.3.2. y cuyo listado se encuentra en el apéndice B.

Con objeto de resolver el problema de optimización no lineal, para el método de Mínimos Cuadrados No Lineales a través de las Autocovarianzas, se utilizó la subrutina LMDER1 del paquete de análisis numérico (ver Calderón y Nocedal (1981), y referencias ahí), para encontrar los estimadores de los parámetros que minimizan la función

$$[y - C \cdot \varphi(\theta)]^T I [y - C \cdot \varphi(\theta)] ;$$

es decir, los estimadores de Mínimos Cuadrados No Lineales. La solución numérica del problema de optimización no lineal, necesita de valores iniciales con los cuales comenzar la búsqueda. Se utilizaron como tales valores a los estimadores iniciales de los parámetros. En todos los casos se utilizó una tolerancia de .0001 para la solución. La tabla 4.7 muestra, el número de veces en que la subrutina LMDER1 no cambió los valores iniciales impuestos, de 0.5

para el estimador del parámetro de la varianza residual, y de 0.4 para los estimadores de los parámetros de promedios móviles, en los 12 casos considerados.

Para el segundo método de estimación, se utilizó el paquete de subrutinas para análisis de modelos de series de tiempo univariadas, que utilizan una versión modificada del método de estimación de Box y Jenkins, basado en un enfoque no condicional (ver David J. Pack (1974), y referencias ahí). Al igual que en el caso anterior, se utilizaron los estimadores iniciales de los parámetros, como los valores iniciales a estas subrutinas. En todos los casos se utilizó una tolerancia de 0.01 para la solución, con objeto de hacer comparables los dos métodos.

En las tablas 4.1 a 4.4, aparecen los valores de la raíz cuadrada del error cuadrático medio de los estimadores, para muestras de tamaños 25, 100 y 300, y para ambos métodos de estimación.

En general para este modelo, los valores de la RCECM que aparecen en la tabla 4.1, son del mismo orden de magnitud. El método de MCNI presenta valores de la RCECM un poco más pequeños, en tamaño de muestra 25, para el caso de los

TABLA 4.1 RAIZ CUADRADA DEL ERROR CUADRATICO MEDIO (RCECM) DE LOS ESTIMADORES DE LOS PARAMETROS DEL MODELO ARMA (0,2).

		= .393 + i2.644		
RAICES DE B(Z) = 0 :		= .393 - i2.644		
		=	= 7.143	
PARAMETROS :		.11	-.14	$\hat{\sigma}_e^2$
		\hat{b}_1	\hat{b}_2	
k=6 ; TM=25	MCNL	.2295	.2443	.6034
	BJ	.2404	.2571	.5836
k=15 ; TM=25	MCNL	.2295	.2443	.6034
	MCNL	.1073	.1101	.2952
k=25 ; TM=100	BJ	.1028	.1035	.2970
	MCNL	.0586	.0638	.1627
k=25 ; TM=300	BJ	.0562	.0590	.1629

estimadores de los parámetros de promedios móviles b_1 y b_2 , y en tamaños de muestra 100 y 300, para el caso del estimador del parámetro de la varianza residual σ_ϵ^2 . También se observa que, los valores de la RCECM no cambian, cuando aumenta el número de autocovarianzas k , de 6 a 15, en el método de MCNL. Para ambos métodos, los valores de la RCECM disminuyen aproximadamente a la mitad, cuando aumenta el tamaño de muestra de 25 a 100, y de 100 a 300.

En general para el modelo a que se refiere la tabla 4.2, se obtuvieron valores de la RCECM más pequeños con el método de BJ, siendo más grandes las diferencias, entre ambos métodos, en tamaño de muestra 300, así como, en el caso de los estimadores del parámetro de la varianza residual σ_ϵ^2 . Por otro lado, se observa que los valores de la RCECM casi no cambian, en el método de MCNL, cuando aumenta el número de autocovarianzas k , de 6 a 15, siendo incluso menor para k igual a 6, en el caso del estimador del primer parámetro autorregresivo a_1 . Para ambos métodos, se presenta una disminución en los valores de la RCECM, cuando aumenta el tamaño de muestra, siendo más rápida esta disminución con el método de BJ.

TABLA 4.2 RAIZ CUADRADA DEL ERROR CUADRATICO MEDIO
(RCECM) DE LOS ESTIMADORES DE LOS
PARAMETROS DEL MODELO ARMA (2,0).

RAICES DE $A(Z) = 0$:		= 1.11 ;	= -2.0	
PARAMETROS :		.40 \hat{a}_1	.45 \hat{a}_2	1 $\hat{\sigma}_\epsilon^2$
k=6 ; TM=25	MCNL	.2239	.2622	.3561
	BJ	.2126	.1950	.2995
k=15 ; TM=25	MCNL	.2248	.2509	.3546
	MCNL	.1368	.1533	.3280
k=25 ; TM=100	BJ	.1030	.0980	.1402
	MCNL	.1019	.1175	.3033
k=25 ; TM=300	BJ	.0542	.0546	.0791

TABLA 4.3 RAIZ CUADRADA DEL ERROR CUADRATICO MEDIO
 (RCECM) DE LOS ESTIMADORES DE LOS
 PARAMETROS DEL MODELO ARMA (1,1).

RAIZ DE A(Z) = 0 : = 1.10

RAIZ DE B(Z) = 0 : = 2

PARAMETROS :	.91 \hat{a}_1	.5 \hat{b}_1	1 $\hat{\sigma}_\epsilon^2$
k=6 ; TM=25	MCNL .4505	.4256	.3637
	BJ .3874	.4671	.3257
k=15 ; TM=25	MCNL .4422	.4284	.3614
	MCNL .0974	.1928	.2232
k=25 ; TM=100	BJ .0830	.1363	.1447
	MCNL .0559	.1363	.1143
k=25 ; TM=300	BJ .0404	.0712	.0809

En general para este modelo, el método de BJ produjo valores de la RCECM más pequeños, exceptuando el caso del

estimador del parámetro de promedio móvil b_1 , en tamaño de muestra 25. Las diferencias entre ambos métodos, fueron del orden de 1/100. Nuevamente, como en los modelos anteriores, las diferencias en los valores de la RCECM son insignificantes, cuando el número de autocovarianzas k , aumenta de 6 a 15, en el método de MCNL. Para ambos métodos, se presenta una disminución en los valores de la RCECM, cuando aumenta el tamaño de muestra, siendo más notoria esta disminución, para el caso del estimador del parámetro autorregresivo a_1 . El más grande valor de la RCECM, para el método de MCNL, se presenta en tamaño de muestra 25, en el caso del estimador del parámetro autorregresivo a_1 , en este caso en particular, se conjugan dos hechos importantes, un tamaño de muestra pequeño y una raíz cercana al círculo unitario.

Para el modelo en la tabla 4.4, el método de MCNL presenta valores más pequeños de la RCECM, en tamaño de muestra 25, en el caso de los estimadores de los parámetros autorregresivos a_1 y a_2 , así como, para el de promedio móvil b_2 , y en tamaño de muestra 100, para el caso del estimador del parámetro de promedio móvil b_1 . En todos los demás casos BJ es menor, siendo más grandes las diferencias entre ambos métodos, en el caso de los estimadores del parámetro de la varianza residual σ_t^2 . En el caso del estimador del parámetro de promedio móvil b_1 , en tamaños de muestra 25 y 300, MCNL produjo un valor mayor de la RCECM, mientras que en tamaño de muestra 100, BJ produjo un valor mayor.

TABLA 4.4 RAIZ CUADRADA DEL ERROR CUADRATICO MEDIO
(RCECM) DE LOS ESTIMADORES DE LOS
PARAMETROS DEL MODELO ARMA (2,2).

		= 1.471 + i1.929				
RAICES DE A(Z) = 0 :		= 1.471 - i1.929				
		= 5,882				
RAICES DE B(Z) = 0 :		= 1.974 ;	= -3.897			
PARAMETROS :		.50 \hat{a}_1	-.17 \hat{a}_2	.25 \hat{b}_1	.13 \hat{b}_2	3 $\hat{\sigma}_e^2$
k=6 ; TM=25	MCNL	.6008	.4074	1.068	.3733	1.406
	BJ	.6042	.4377	.7531	.7512	1.134
k=15 ; TM=25	MCNL	.5899	.4049	.5585	.3825	1.409
	MCNL	.4995	.3272	.4536	.3624	.8430
k=25 ; TM=100	BJ	.4635	.2841	.4861	.3614	.4630
	MCNL	.3081	.2090	.2890	.2388	.4400
k=25 ; TM=300	BJ	.2556	.1759	.2616	.2046	.2480

En general, para ambos métodos, los valores de la RCECM disminuyen aproximadamente a la mitad, cuando aumenta el tamaño de muestra de 25 a 300, en el caso de los estimadores de los parámetros autorregresivos, mientras que, para el caso de los de promedio móvil, la situación no es muy clara, ya que, para el caso de b_1 , MCNL presenta una mayor disminución que BJ, y para b_2 , BJ presenta una mayor disminución que MCNL. Para el caso del estimador del parámetro de la varianza residual σ_e^2 , los valores de la RCECM disminuyen aproximadamente a una tercera parte, con el método de MCNL y aproximadamente a una cuarta parte, con el método de BJ. Cuando el número de autocovarianzas k , aumenta de 6 a 15, el método de MCNL aumenta un poco los valores de la RCECM, en el caso de los estimadores de los parámetros de promedio móvil b_2 y de la varianza residual σ_e^2 , y disminuye un poco, para el caso de los estimadores de los parámetros autorregresivos a_1 y a_2 , en el caso del de promedio móvil b_1 , se presenta una disminución de casi la mitad.

Se decidió resumir la información presentada en las tablas 4.1 a 4.4, por tamaño de muestra, para los casos en que el método de MCNL resultó ser mejor en el experimento de simulación. En la tabla 4.5 siguiente, aparecen las diferencias en la RCECM de estimación, para los casos en que MCNL tuvo una RCECM menor que BJ.

TABLA 4.5 DIFERENCIAS EN RAIZ CUADRADA DEL ERROR CUADRATICO MEDIO DE ESTIMACION, EN LOS CASOS EN QUE MCNL RESULTO MENOR QUE BJ.

TAMAÑO DE MUESTRA 25

	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	$\hat{\sigma}_\epsilon^2$
ARMA (0,2)			.0109	.0128	x
ARMA (2,0)	x	x			x
ARMA (1,1)	x		.0415		x
ARMA (2,2)	.0034	.0303	x	.3779	x

TAMAÑO DE MUESTRA 100

	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	$\hat{\sigma}_\epsilon^2$
ARMA (0,2)			x	x	.0018
ARMA (2,0)	x	x			x
ARMA (1,1)	x		x		x
ARMA (2,2)	x	x	.0325	x	x

TABLA 4.5 (CONTINUACION)

TAMAÑO DE MUESTRA 300

	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	$\hat{\sigma}_\epsilon^2$
ARMA (0,2)			x	x	.0002
ARMA (2,0)	x	x			x
ARMA (1,1)	x		x		x
ARMA (2,2)	x	x	x	x	x

x = BJ fue menor.

En la tabla anterior se observa, que el método de MCNL produjo mejores estimadores en muestras de tamaño 25, para el caso de los estimadores de los parámetros autorregresivos y de promedios móviles, ya que, en 6 de entre 10 posibilidades tuvo una RCECM menor en el experimento de simulación. En muestras de tamaños 100 y 300, y en general para el estimador del parámetro de la varianza residual, BJ fue consistentemente mejor en los experimentos de simulación.

4.1 Resultados Adicionales.

Anteriormente se mencionó, que se asignaron valores iniciales arbitrarios a los estimadores de los parámetros de promedios móviles, cuando la subrutina FTMPS no convergió para encontrar estimadores iniciales de los mismos. Se decidió contar el número de veces de entre 300 (número de muestras), en que FTMPS no convergió, para los modelos ARMA (0,2), (1,1) y (2,2). La tabla 4.6 muestra estos resultados.

El mayor porcentaje se presentó en el modelo ARMA (2,2), para tamaño de muestra 25, es decir, que en un 23% de los casos, se asignó el valor de 0.4 a los estimadores iniciales de los parámetros de promedios móviles b_1 y b_2 , y de 0.5 al de la varianza residual σ_t^2 , cuándo los valores verdaderos de los parámetros para este modelo fueron 0.25, 0.13 y 3 respectivamente. Los porcentajes no cambian, cuando k aumenta de 6 a 15, porque FTMPS utiliza solamente $p+q+1$ autocovarianzas muestrales para estimar a los parámetros (ver apéndice C).

TABLA 4.6 NUMERO DE VECES DE ENTRE 300, EN QUE FTMPS
 NO OBTIENE ESTIMADORES INICIALES DE LOS
 PARAMETROS DE PROMEDIOS MOVILES.

	TM=25		TM=100	TM=300
	k = 6	k = 15		k = 25
ARMA (0,2)	11 (4%)	11 (4%)	0 (0%)	0 (0%)
ARMA (1,1)	54 (18%)	54 (18%)	33 (11%)	5 (2%)
ARMA (2,2)	68 (23%)	68 (23%)	30 (10%)	9 (3%)

Al momento de analizar los resultados, se observó que la subrutina que se utilizó (LMDER1), para encontrar los estimadores de Mínimos Cuadrados No Lineales, no cambiaba, en algunos de los casos, el valor asignado arbitrariamente a los estimadores iniciales de los parámetros, por lo que se decidió contar el número de veces de entre 300, en que los estimadores finales fueron iguales a los iniciales. La tabla 4.7 siguiente, muestra estos resultados.

TABLA 4.7 NUMERO DE VECES DE ENTRE 300, EN QUE LMDER1 NO CAMBIA LOS VALORES INICIALES IMPUESTOS DE 0.5 PARA EL ESTIMADOR DE LA VARIANZA RESIDUAL Y DE 0.4 PARA LOS ESTIMADORES DE LA PARTE MA.

	TM=25		TM=100	TM=300
	k = 6	k = 15		k = 25
ARMA (0,2)	6 (2%)	6 (2%)	0 (0%)	0 (0%)
ARMA (2,0)	77 (26%)	83 (28%)	84 (28%)	79 (26%)
ARMA (1,1)	29 (10%)	26 (9%)	25 (8%)	5 (2%)
ARMA (2,2)	49 (16%)	47 (16%)	17 (6%)	2 (1%)

En general, los mayores porcentajes se presentan para el modelo ARMA (2,0), en el cual, siempre se asignó un valor inicial arbitrario de 0.5 al estimador inicial del parámetro de la varianza residual σ_e^2 . Curiosamente, para este mismo modelo, el método de MCNL nunca resulto ser mejor que el de BJ en el experimento de simulación, como se puede apreciar

en la tabla 4.5. El hecho de que el número de autocovarianzas k , se aumente de 6 a 15, no parece influir en el comportamiento de LMDER1, lo cual nos hace pensar que; o los valores iniciales impuestos arbitrariamente a los estimadores estaban muy cercanos a un mínimo local de la función, o que la función para algunos casos resultó ser muy plana.

Se decidió buscar los casos en que el método de MCNL produjo valores mayores de la RCECM al método de BJ, en más de 1/100. La tabla 4.8 siguiente muestra estos resultados.

TABLA 4.8 DIFERENCIAS EN LA RAIZ CUADRADA DEL ERROR CUADRATICO MEDIO DE ESTIMACION, PARA LOS CASOS EN QUE MCNL FUE MAYOR QUE BJ, EN MAS DE 1/100.

	TM=25		TM=100		TM=300	
	\hat{b}_1	$\hat{\sigma}_e^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_e^2$
ARMA (2,0)				.1878		.2242
ARMA (2,2)	.3149	.2720	.3800		.1920	

Los resultados en la tablas 4.7 y 4.8 anteriores, nos

hacen pensar que las deficiencias en el método de MCNL, se pueden deber en gran medida, a la subrutina de optimización no lineal que se utilizó, ya que como se puede apreciar en las tablas anteriores, los casos en que el método de Mínimos Cuadrados No Lineales a través de la Autocovarianzas, estimó menos bien a los parámetros, en los experimentos de simulación, coinciden con los casos, en que la subrutina LMDER1 no cambió los estimadores iniciales, en un número considerable de veces.

Capítulo 5

CONCLUSIONES

Con base en experimentos de simulación, un método de estimación de parámetros de modelos ARMA(p,q) basado en el ajuste de las autocovarianzas teóricas, expresadas en términos de los parámetros por estimar, a las autocovarianzas muestrales por Mínimos Cuadrados No Lineales (MCNL), se comparó con él de Box y Jenkins basado en un razonamiento de Máxima Verosimilitud (BJ). En términos del Error Cuadrático Medio de los estimadores, el método de MCNL parece dar mejores resultados, para los parámetros autorregresivos y de promedios móviles, en muestras pequeñas (25). Sin embargo, MCNL parece dar peores resultados, en todos los casos, para el estimador del parámetro de la varianza residual, así como, para muestras de tamaño muy grande (300). Cuando alguna de las raíces de la parte autorregresiva, esta cercana al círculo unitario, es decir, cuando algun parámetro autorregresivo esta cercano a la región de no estacionariedad, BJ es consistentemente mejor.

Cuando el número de parámetros por estimar (3) y el tamaño de muestra son pequeños (25), un número pequeño de autocovarianzas por ajustar, parece ser suficiente (6) en el método de MCNL, sin embargo, cuando el número de parámetros por estimar aumenta a 5, MCNL presenta una leve mejoría, al aumentar el número de autocovarianzas de 6 a 15, en muestras pequeñas.

La subrutina de optimización utilizada para resolver el problema de optimización no lineal (LMDER1), parece no ser la más apropiada, para el tipo de función a minimizar en el método de MCNL. Tomando en cuenta los resultados obtenidos en esta tesis, sería conveniente hacer un análisis de la forma de las funciones que se presentan, y talvés utilizar otro método de optimización no lineal que fuese más apropiado, con lo que se esperaría obtener una mejoría del método.

Apéndice A

ANTECEDENTES ESTADÍSTICOS

A.1 Procesos Gaussianos.

Se dice que un proceso $\{X_t\}$ es gaussiano, si la densidad conjunta de las variables aleatorias X_{t_1}, \dots, X_{t_N} es Normal N-variada, para cada t_1, \dots, t_N , y cada entero N.

Si el proceso $\{X_t\}$ es estacionario, entonces, éste será completamente caracterizado por su media μ , y su matriz de varianzas y covarianzas Γ , con su (i,j) -ésimo elemento igual a la $\text{cov}(X_i, X_j) = \gamma_{|i-j|}$, es decir;

$$\Gamma = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{N-2} \gamma_{N-1} \\ \gamma_{-1} & \gamma_0 & \dots & \gamma_{N-3} \gamma_{N-2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \gamma_{1-N} & \gamma_{2-N} & \dots & \gamma_{-1} \gamma_0 \end{pmatrix}_{N \times N} \quad (1)$$

y la función de densidad conjunta del vector de variables aleatorias $X = (X_{t_1}, \dots, X_{t_N})'$, esta definida como:

$$f(X) = (2\pi)^{-N/2} |\Gamma|^{-1/2} \exp \left\{ -1/2 (X-\mu)' \Gamma^{-1} (X-\mu) \right\} \quad (2)$$

A.2 Función Generadora de Autocovarianzas y

Función de Densidad Espectral de un Proceso ARMA(p,q).

Sean $\{X_t\}$ un proceso gaussiano, estacionario e invertible, y $\{\varepsilon_t\}$ un proceso de ruido blanco, tales que:

$$\begin{aligned}
 X_t - a_1 X_{t-1} - \dots - a_p X_{t-p} &= \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q} \\
 A(R) X_t &= B(R) \varepsilon_t
 \end{aligned}
 \tag{3}$$

donde $A(R)$ y $B(R)$, son polinomios en R , de ordenes p y q respectivamente.

Para encontrar la función de densidad espectral del proceso $\{X_t\}$, $f_x(\lambda)$, se definen:

i) La función de densidad espectral $f_\varepsilon(\lambda)$, del proceso de ruido blanco $\{\varepsilon_t\}$, se define como

$$f_\varepsilon(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \quad \text{para } -\pi < \lambda < \pi$$

ii) Sea $\{Y_t\}$ un proceso, tal que $Y_t = B(R) \varepsilon_t$, entonces, su función de densidad espectral, se define como

$$f_Y(\lambda) = |B(e^{i\lambda})|^2 f_\varepsilon(\lambda) = \frac{\sigma_\varepsilon^2 |B(e^{i\lambda})|^2}{2\pi} \quad -\pi < \lambda < \pi$$

(4)

De (3), se tiene también que el proceso $\{Y_t\}$, es tal que $Y_t = A(R) X_t$, y su función de densidad espectral también se puede expresar como

$$f_Y(\lambda) = |A(e^{i\lambda})|^2 f_X(\lambda) \quad -\pi < \lambda < \pi \quad (5)$$

Por lo tanto, la función de densidad espectral del proceso $\{X_t\}$, se puede obtener a partir de (4) y (5), como

$$f_X(\lambda) = \frac{\sigma_\varepsilon^2 |B(e^{i\lambda})|^2}{2\pi |A(e^{i\lambda})|^2} \quad -\pi < \lambda < \pi \quad (6)$$

Por otro lado, la función de densidad espectral del proceso $\{X_t\}$, se define como un múltiplo de la función generadora de autocovarianzas $G_X(Z)$ del proceso, sobre el círculo unitario, es decir

$$f_X(\lambda) = \frac{1}{2\pi} G_X(e^{i\lambda}) \quad -\pi < \lambda < \pi \quad (7)$$

donde $G_x(e^{i\lambda})$, es la función generadora de autocovarianzas del proceso $\{X_t\}$, definida como

$$G_x(Z) = \sum_{j=-\infty}^{\infty} \gamma_j Z^j \quad (8)$$

donde Z es una variable compleja.

A partir de (6) y (7), se tiene que

$$2\pi f_x(\lambda) = G_x(e^{i\lambda}) = \frac{\sigma_e^2 |B(e^{i\lambda})|^2}{|A(e^{i\lambda})|^2} \quad -\pi < \lambda < \pi$$

Sea $Z = e^{i\lambda}$, entonces, la función generadora de autocovarianzas del proceso $\{X_t\}$ sobre el círculo unitario, se puede expresar como

$$G_x(Z) = \frac{\sigma_e^2 |B(Z)|^2}{|A(Z)|^2} \quad (9)$$

Como se observa en (8), $G_x(Z)$ es simétrica, además los polinomios $A(Z)$ y $B(Z)$ no tienen ceros sobre el círculo unitario, así a cada cero de $G_x(Z)$ en (9), fuera del círculo unitario, correspondera uno dentro. Separando estos ceros,

mediante una factorización adecuada de los polinomios $A(Z)$ y $B(Z)$, $G_x(Z)$ en (9) se puede expresar en forma única como:

$$G_x(Z) = \frac{\sigma_e^2 B(Z)B(Z^{-1})}{A(Z)A(Z^{-1})} \quad (10)$$

donde los polinomios $A(Z)$ y $B(Z)$, tiene todos sus ceros fuera del círculo unitario, y $A(Z^{-1})$ y $B(Z^{-1})$ dentro.

Usando las expresiones en (8) y (10) para $G_x(Z)$, se llega a la siguiente relación, entre las autocovarianzas del proceso y los parámetros del mismo, en (3)

$$\sum_{j=-\infty}^{\infty} \gamma_j Z^j = \frac{\sigma_e^2 B(Z)B(Z^{-1})}{A(Z)A(Z^{-1})} \quad (11)$$

A partir de la igualdad en (11), se puede obtener una expresión para la autocovarianza de orden j , en términos de los parámetros del proceso, igualando los coeficientes de Z^j . Obviamente la relación entre γ_j y los parámetros del modelo es una no lineal.

El lector interesado en un desarrollo más detallado, de las definiciones presentadas en este apéndice, puede consultar Nerlove, Grether y Carvalho (1979), Cap. 3.

APENDICE B


```

CALL SIMULA(1,1),R(1),VARE,TM,NM,MSIM)
CALL SUBROUTINA(TM,MSIM,MACVE,MPAIE)
CALL OPTIM(MACVE,MPAIE)
DO 30 I=1,100
WRITE(12,34)I,(MPAIE(I,J),J=0,NPAE1)
FORMAT(1X,I3,7(1X,F8.4))
CONTINUE
DO 25 I=1,100
DO 24 J=0,NPAE1
ECM(J)=ECM(J)+(MPAIE(I,J)-RPA(J))*(MPAIE(I,J)-RPA(J))
CONTINUE
CONTINUE
DO 26 I=0,NPAE1
ECM(I)=SQRT(ECM(I)/300)
CONTINUE
25 FORMAT(1X,I3,7(3X,F8.4),/)
31 WRITE(12,35)ECM(J),J=0,NPAE1)
35 FORMAT(1X,I3,7(3X,F8.4))
18 DO 10 TO 19
CLOSE(11,DISP=CRUNCH)
CLOSE(12,DISP=CRUNCH)
CLOSE(13,DISP=CRUNCH)
STOP
END

```

```

C*
C* SUBROUTINA QUE GENERA 100 MUESTRAS DE TAMAÑO TM
C* EN CADA LLAMADA.
C* LE LLEGAN LOS SIGUIENTES PARAMETROS:
C* A,B,VARE,P,Q,TM,NM
C* A.- VECTOR DE LONGITUDES P. CONTIENE LOS PARAMETROS
C* B.- VECTOR DE LONGITUDES Q. CONTIENE LOS PARAMETROS
C* DE PROMEDIOS MOVILES DEL MODELO.
C* VARE.- VARIANZA RESIDUAL.
C* TM.- TAMAÑO DE LAS MUESTRAS.
C* NM.- TAMAÑO VALORES 1,2,3 INDICA EL NUMERO DE BLOQUES
C* DE 100 MUESTRAS DE TAMAÑO TM.
C* P.- ORDEN DE LA PARTE AUTORREGRESIVA DEL MODELO.
C* Q.- ORDEN DE LA PARTE DE PROMEDIOS MOVILES DEL MODELO.
C* REGRESA:
C* LA MATRIZ MSIM.
C* MSIM.- MATRIZ DE LONGITUD (1:100,1:TM). ESTA MATRIZ CONTIENE
C* LAS 100 MUESTRAS GENERADAS DE TAMAÑO TM.
C* CADA BLOQUE DE MSIM CONTIENE UNA MUESTRA DE TAMAÑO TM.
C* PARAMETROS INTERNOS:
C* XSIMI.- VECTOR DE LONGITUD P. CONTIENE LOS VALORES
C* INICIALES PARA GENERAR LAS MUESTRAS.
C* CUANDO NM=1 XSIMI CONTIENE CEROS.
C* CUANDO NM>1 XSIMI CONTIENE LOS ÚLTIMOS P VALORES
C* DE LA ÚLTIMA MUESTRA GENERADA.
C* ESTA SUBROUTINA LLAMA A LA SUBROUTINA FTGEN DEL PAQUETE IMSL,
C* LA CUAL, GENERA UNA SERIE DE TIEMPO DADOS LOS ORDENES P Y Q
C* DEL PROCESO GENERADOR.
C*
C* SUBROUTINE SIMULA(ASIM,BSIM,VARE,TM,NM,MSIM)
C* COMMON P,Q,P1,Q1
C*
C* VARIABLES ELEMENTOS DE UN COMMON.

```

```

C*      INTEGER P,Q,P1,J1
C*      VARIABLES FORMALES
C*      INTEGER TM,NM
C*      REAL ASIM(P),BSIM(Q),VARE,MSIM(1:100,1:TM)
C*      VARIABLES LOCALES
C*      INTEGER TM2,I,J,I1
C*      REAL PMAC,WA(403),XSIM1(400)
3SET OWN REAL XSIMI(3)
        DOUBLE PRECISION DSEED
$RESET OWN
        PMAC=0.0
        GO TO (0),N*
        TM2=50+TM
        GO TO 51
        TM2=100+TM
50      DSEED=123457.000
        IF(P.EQ.0)GO TO 51
        DO 49 I=1,P
        XSIMI(I)=0.0
49      CONTINUE
C*      LLAMADA A FTGEN DEL PAQUETE IMSL.
C*
51      CALL FTGEN(ASIM,BSIM,PMAC,XSIMI,VARE,DSEED,P,Q,TM2,XSIM1,WA)
        DO 52 J=1,TM
        XSIM1(J)=XSIM1(TM2-TM+J)
52      CONTINUE
        IF(P.EQ.0)GO TO 53
        DO 53 I=1,P
        XSIMI(I)=XSIM1(TM2-P+I)
53      CONTINUE
        DO 54 I=1,TM2
        XSIM1(I)=0.0
54      CONTINUE
        TM2=50+TM
        DO 60 I1=2,100
        PMAC=0.0
C*      LLAMADA A FTGEN DEL PAQUETE IMSL.
C*
        CALL FTGEN(ASIM,BSIM,PMAC,XSIMI,VARE,DSEED,P,Q,TM2,XSIM1,WA)
        DO 55 J=1,TM
        XSIM1(J)=XSIM1(TM2-TM+J)
55      CONTINUE
        IF(P.EQ.0)GO TO 56
        DO 56 I=1,P
        XSIMI(I)=XSIM1(TM2-P+I)
56      CONTINUE
        DO 57 J=1,TM2
        XSIM1(J)=0.0
57      CONTINUE
60      CONTINUE
        RETURN
        END

```



```

67 CONTINUE
*****
SE SUPONE QUE LA MEDIA DEL PROCESO ES CERO, POR LO TANTO SE
LE DA EL VALOR DE CERO A XBAR Y SE ESTIMA LA VARIANZA DE LA
SERIE BAJO ESTE SUPUESTO.
XBAR Y ACV(O) SON PARAMETROS DE ENTRADA A FTAUTO.
XBAR=0.0
SUM=0.0
DO 60 J=1, TM
SUM=SUM+REVSIM(J)*PEVSIM(J)
69 CONTINUE
ACV(O)=SUM/TM
***
LLAMADA A FTAUTO DEL PAQUETE IMSL.
CALL FTAUTO(REVSIM,TM,LAG,L,ISW,XBAR,ACV(O),ACV(1),AC,PACV,
WKAREB)
DO 61 J=0, LAG
MACV(I1,J)=ACV(J)
61 CONTINUE
CC
CALCULA ESTIMADORES INICIALES DE LOS PARAMETROS ARMA.
AIE(O)=1.0
IF(P.NE.O) GO TO 71
AIE(1)=0.0
GO TO 53
***
LLAMADA A FTARPS DEL PAQUETE IMSL
71 CALL FTARPS(ACV,XBAR,P,0,AIE(1),PMAC,WAA,IER1)
WRITE(13,/)I1," IER1 = ",IER1
CC
SE VERIFICA CONDICION DE ESTACIONARIEDAD PARA LOS PARAMETROS AR.
DO 59 J=0,P
AT1(J+1)=AIE(J)
59 CONTINUE
CALL REST(P1,AT1,AT2,P)
DO 58 J=0,P
AIE(J)=AT2(J+1)
58 CONTINUE
DO 62 J=0,P
PAIE(I1,J)=AIE(J)
62 CONTINUE
SIG(O)=1.0
IF(S.NE.O) GO TO 72
SIE(1)=0.0
CC
SI Q=0, SE ASIGNA VALOR INICIAL AL ESTIMADOR DE LA VARIANZA
RESIDUAL.
VAREIE=.5
GO TO 43
***
LLAMADA A FTMPMS DEL PAQUETE IMSL.
72 CALL FTMPMS(ACV,AIE(1),P,0,SIE(1),VAREIE,WA,IER2)
WRITE(13,/)I1," IER2 = ",IER2

```



```

PA(J)=(VD(J)-SUM)/AE(O)
PA(-J)=RA(J)
IF(L.E3,LAG)GO TO 11
16 CONTINUE
DO 18 J=O+1,LAG
SUM=O.O
DO 18 K=1,P
SUM=SUM+AE(K)+RA(J-K)
17 CONTINUE
J1=ABS(J-I*)
RA(J)=- (R(J1)+SUM)/AE(O)
PA(-J)=RA(J)
13 CONTINUE
11 DO 12 JJ=O,LAG
12 FJAC(JJ,II)=PA(JJ)
12 CONTINUE
32 CONTINUE

C*
C**
C***
CALCULO DE LAS DERIVADAS CON RESPECTO A B(J),J=1,0
BAN=3
C*
C**
C***
EL PRIMER PARAMETRO MA SE SUPONE CONSTANTE E IGUAL A 1, POR LO
TANTO LAS DERIVADAS CON RESPECTO A ESTE PARAMETRO SON IGUALES
A CERO Y NO ENTAN AL JACOBIANO.
C*
C**
C***
DO 22 II=1,6
CALL CALDB(O,AE,OB,II)
CALL CALS(P,AE,O,OB,SB)
CALL CALR(O,P,AE,SB,OB,LAG,BAN,VC)
DO 23 JJ=O,LAG
FJAC(JJ,II)=RB(JJ)
23 CONTINUE
22 CONTINUE

C*
C**
C***
CALCULO DE LAS DERIVADAS CON RESPECTO A SIGMA (AL CUADRADO)
BAN=4
DO 26 I=O,3
DSY(I)=D(I)/VAREE
25 CONTINUE
CALL CALS(P,AE,O,DSIG,SSIG)
CALL CALR(O,P,AE,SSIG,PSIG,LAG,BAN,VC)
DO 27 II=O,LAG
FJAC(II,NPAE2)=PSIG(I)
27 CONTINUE
GO TO 333

C*
C**
C***
ERROR
900 IFLAG=-1
999 RETUR
END

C*
C**
C***
ESTA SUBROUTINA OBTIENE LOS COEFICIENTES "D'S",
NECESARIOS PARA EL CALCULO DE LA FUNCION DE
AUTOCOVARIANZAS.

SUBROUTINE CALD(O1,O2,O3)
COMMON/JND/VAREE

```



```

C*
C**
C**
REAL DB2(G:DB1),DB3(-3:DB1)
INTEGER DB1,DB4
VARIABLES LOCALES.
INTEGER I,JJM,JIME
DO 201 I=D,DR1
JIM=DB4+I
JIME=DB4-I
IF(O.LE.JJM.AND.JJM.LE.DB1.AND>O.LE.JIME.AND.JIME.LE.DB1)GO TO 202
IF(JIME.LT>O.AND.JJM.LE.DB1)GO TO 203
IF(JIM.GE>O.AND.JJM.GT.DB1)GO TO 204
DB2(I)=J
GO TO 201
202 DB3(I)=VAREE*(DB2(JJM)+DB2(JIME))
GO TO 201
203 DB3(I)=VAREE+DB2(JJM)
GO TO 201
204 DB3(I)=VAREE+DB2(JJM)
201 CONTINUE
RETURN
END

```

```

C*
C**
C**
ESTA SUBROUTINA VERIFICA QUE LOS PARAMETROS DEL
MODELO SATISFAGAN LAS CONDICIONES DE ESTACIONARIEDAD
EN EL CASO DE LOS PARAMETROS AR, Y DE INVERTIBILIDAD
EN EL CASO DE LOS PARAMETROS MA.
C*
C**
ESTA SUBROUTINA LLAMA A ZRPOLY DEL PAQUETE IMSL, LA
CUAL OBTIENE LAS RAICES DE UN POLINOMIO CON
COEFICIENTES REALES.

```

```

C*
C**
SUBROUTINE REST(N1,C1,CN,N)
VARIABLES FORMALES.
INTEGER N,N1
REAL C1(N1),CN(N1)
VARIABLES LOCALES.
INTEGER IER
REAL RZ(4),RZ,C2(4)
COMPLEX Z(4),SUMR,SUMR1(4),SUMR2(4)
EQUIVALENCE (RZ,Z)
C2(N1)=C1(1)
DO 10 I=D,N-1
C2(I+1)=C1(N1-I)
10 CONTINUE

```

```

C*
C**
LLAMADA A ZRPOLY DEL IMSL.
CALL ZRPOLY(C2,N,RZ,IER)
WRITE(13,/) IERRES = ,IER
I=1
11 IF(I.GT.N)GO TO 15

```

```

C*
C**
SE VERIFICA QUE LA NORMA DE LAS RAICES DEL POLINOMIO CARAC-
TERISTICO (1/Z(I)) SEA MAYOR QUE 1, ES DECIR QUE !Z(I)!<1.

```


APENDICE C

 D E C O M P

DECOMP SUBROUTINA FORTRAN PARA CALCULAR LA DESCOMPOSICION TRIANGULAR DE UNA MATRIZ CUADRADA REAL

 1. DESCRIPCION

1.1 PROPOSITO

DADA UNA MATRIZ CUADRADA REAL A DE ORDEN N. LA SUBROUTINA DECOMP CALCULA LA DESCOMPOSICION TRIANGULAR DE A DE LA FORMA

$$PA=LU,$$

DONDE L ES UNA MATRIZ TRIANGULAR SUPERIOR CON UNOS EN SU DIAGONAL, U ES UNA MATRIZ TRIANGULAR SUPERIOR, Y P ES UNA MATRIZ DE PERMUTACION.

CALCULA ADEMAS UNA ESTIMACION DEL NUMERO DE CONDICION DE A COMO

$$K(A) = \|A\| \|INV(A)\|$$

DONDE LA NORMA UTILIZADA ES LA NORMA 1, PARA DETERMINAR QUE LA MATRIZ ES SINGULAR.

1.2 METODO

SE APLICAN $K(K=1, \dots, N-1)$ ETAPAS DE ELIMINACION GAUSSIANA CON PIVOTEO PARCIAL POR RENGLONES PARA ELIMINAR LOS $N-K$ ELEMENTOS DE LA K -ESIMA COLUMNA DE LA MATRIZ A PARA REDUCIRLA A UNA MATRIZ TRIANGULAR SUPERIOR. EN CADA ETAPA SE CALCULAN Y SE SALVAN LOS MULTIPLICADORES USADOS PARA ELIMINAR LOS $N-K$ ELEMENTOS DE LA K -ESIMA COLUMNA Y EL INDICE DE CADA FILA PIVOTE. AL TERMINO DE LA REDUCCION SE HAN OBTENIDO TRES FACTORES: UNA MATRIZ P ASOCIADA A LA ESTRATEGIA DEL PIVOTE, UNA MATRIZ TRIANGULAR INFERIOR L CUYOS ELEMENTOS SON LOS MULTIPLICADORES USADOS CON 1'S EN SU DIAGONAL (NO SE ALMACENAN), Y UNA MATRIZ TRIANGULAR SUPERIOR U CUYOS ELEMENTOS SON LOS DE LA MATRIZ A. L ES ALMACENADA EN LA PARTE INFERIOR DE A, U EN LA PARTE SUPERIOR Y, P ES ALMACENADA EN UN VECTOR QUE CONTIENE UNICAMENTE LOS INDICES DE LAS FILAS PIVOTES. PARA CALCULAR EL NUMERO DE CONDICION DE LA MATRIZ A, LA SUBROUTINA UTILIZA LA SIGUIENTE RELACION

$$K(A) = C * \|A(J)\| * \|Z\| / \|Y\|$$

DONDE $A(J)$ ES LA J -ESIMA COLUMNA DE A; Y, Y Y Z SON DOS VECTORES CALCULADOS POR LA SUBROUTINA TAL QUE $\|Z\| / \|Y\|$ ES APROXIMADAMENTE LA NORMA DE LA MATRIZ INVERSA DE A Y C ES UNA CONSTANTE CON VALOR 1 O -1.

PARA UNA INFORMACION MAS DETALLADA DEL ALGORITMO UTILIZADO CONSULTESE LA REFERENCIA 1, PAGES. 30-30.

 2. USO DE LA SUBROUTINA

2.1 USO GENERAL:

EN EL PROGRAMA PRINCIPAL SE DEBE DIMENSIONAR LOS ARREGLOS POR

USAR, ESPECIFICAR EL ORDEN DE LA MATRIZ, Y LA MATRIZ QUE VA A SER FACTORIZADA.
SI EL ALGORITMO DETERMINA QUE LA MATRIZ TIENE UN NUMERO DE CONDICION MUY GRANDE, ENTONCES LA SUBROUTINA LO REGRESARA EN EL PARAMETRO " COND " .

- 2.2 DESCRIPCION DE LOS PARAMETROS:
EL USO DE LA SUBROUTINA SE REALIZA POR EL SIGUIENTE ENUNCIADO:

CALL DECOMP(NDIM,N,A,COND,INPUT,WORK)

TENIENDO LOS PARAMETROS EL SIGUIENTE SIGNIFICADO:

PARAMETROS DE ENTRADA

NDIM ES EL NUMERO MAXIMO DE LOS RENGLONES DE LA MATRIZ A, DECLARADO EN EL PROGRAMA PRINCIPAL.

N ES EL ORDEN DE LA MATRIZ A, Y DE LOS VECTORES IPVT Y WORK. N IGUAL O MENOR QUE NDIM.

A MATRIZ ORIGINAL QUE VA A FACTORIZARSE. ES DE DIMENSION NDIM*N.

PARAMETROS DE SALIDA

A CONTIENE EN SU PARTE TRIANGULAR INFERIOR UNA VERSION PERMUTADA DE UNA MATRIZ TRIANGULAR INFERIOR L, Y EN SU PARTE SUPERIOR UNA MATRIZ TRIANGULAR U, TAL QUE PA=LU, SIENDO P UNA MATRIZ DE PERMUTACION.

COND UNA APROXIMACION DEL NUMERO K(A). SI COND+1=COND, SE DICE ENTONCES QUE A ES SINGULAR. SI COND=1.06+32 SE TIENE UNA SINGULARIDAD EXACTA.

IPVT CONTIENE LOS INDICES DE LAS FILAS PIVOTES ASOCIADOS A P. IPVT(K)=INDICE DE LA K-ESIMA FILA PERMUTA (K-ESIMA COLUMNA DE P). IPVT(N)=(-1)**(NUMERO DE INTERCAMBIOS).

WORK ARREGLO AUXILIAR DE DIMENSION N, SU CONTENIDO PUEDE SER IGNORADO.

- 2.3 CONDICIONES DE ERROR Y RETORNOS:
NINGUNO.

- 2.4 REQUERIMIENTOS DE MEMORIA:
EN EL PROGRAMA PRINCIPAL HAY QUE DIMENSIONAR LOS SIGUIENTES ARREGLOS, A(NDIM,N), IPVT(N) Y WORK(N).

- 2.5 SUBPROGRAMAS REQUERIDOS:
SOLVE, CALCULA LA SOLUCION DE LA ECUACION AX=B A PARTIR DE LA DESCOMPOSICION TRIANGULAR DE A.

3. APLICABILIDAD Y RESTRICCIONES

1. LA INVERSA DE UNA MATRIZ CUADRADA A PUEDE CALCULARSE A PARTIR SU DESCOMPOSICION TRIANGULAR, COMO

$$\text{INV}(A) = \text{INV}(U) \text{INV}(L) \text{INV}(P)$$

2. LA SOLUCION DE UN SISTEMA DE N ECUACIONES LINEALES $AX = B$, PUEDE CALCULARSE A PARTIR DE LA DESCOMPOSICION TRIANGULAR LU DE LA MATRIZ A, RESOLVIENDO DOS SISTEMAS LINEALES

$$\begin{aligned} LY &= \text{INV}(P) B \\ UX &= Y \end{aligned}$$

3. EL DETERMINANTE DE UNA MATRIZ CUADRADA A PUEDE CALCULARSE A PARTIR DE SU DESCOMPOSICION TRIANGULAR LU, COMO

$$\text{DET}(A) = [(-1)^{**K}] U(1,1) * \dots * U(N,N)$$

DONDE K ES EL NUMERO DE INTERCAMBIOS DE LOS RENGLONES DE A Y $U(I,I), I=1,N$, SON LOS ELEMENTOS DIAGONALES DE U.

4. REFERENCIAS

1. FORSYTHE, G.E., MALCOLM, M.A. AND MOLER, C.B., COMPUTER METHODS FOR MATHEMATICAL COMPUTATIONS. ENGLEWOOD CLIFFS, N.J., PRENTICE HALL, 1977.
2. STEWART, G.W., INTRODUCTION TO MATRIX COMPUTATIONS. ACADEMIC PRESS, 1973.
3. FORSYTHE, G.E. AND MOLER, C.B., COMPUTER SOLUTION OF LINEAR ALGEBRAIC SYSTEMS. ENGLEWOOD CLIFFS, N.J., PRENTICE-HALL, 1967.
4. CONTE, S.D. AND DE BOOR, C., ELEMENTARY NUMERICAL ANALYSIS, AN ALGORITHMIC APPROACH, 2ND. ED. NEW YORK, MCGRAW HILL.

5. OBSERVACIONES Y SUGERENCIAS

LA ESTIMACION DEL NUMERO DE CONDICION DE UNA MATRIZ ES UNA MEDIDA MUCHO MAS CONFIABLE Y UTIL DE LA SINGULARIDAD DE UNA MATRIZ QUE UN DETERMINANTE O UN PIVOTE MUY CHICO. AL RESOLVER UN SISTEMA DE ECUACIONES LINEALES ES RECOMENDABLE VER ESTE NUMERO PARA DETERMINAR QUE TAN EXACTA VA A SER LA SOLUCION ESPERADA. ESTO ES, SI EL NUMERO DE CONDICION ES MUY GRANDE ENTONCES EL NUMERO DE CIFRAS DECIMALES EXACTAS SERAN MUY POCAS, EN CAMBIO SI ES MUY PEQUEÑO ENTONCES EL NUMERO DE CIFRAS AUMENTARA EN EXACTITUD.

6. EJEMPLO DE PRUEBA

EL DETERMINANTE Y EL NUMERO DE CONDICION DE LA MATRIZ

$$A = \begin{pmatrix} 33 & 16 & 72 \\ -24 & -10 & -57 \\ -8 & -4 & -17 \end{pmatrix}$$

FUERON CALCULADOS EN UNA COMPUTADORA BURROUGHS 6800 CON ARITMETICA EN SIMPLE PRECISION.

PARAMETROS: NDIM=5, N=3.

RESPUESTAS DEL ALGORITMO : $\text{DET}(A) = 6.0$
 $\text{COND}(A) = 7672.44$

 APENDICE A. LISTADO DEL PROGRAMA PRUEBA.

```

SET AUTOBIND
BIND = FROM (ISAK)OBJECT/BIBDAN/=
RESET FREE
*****
C  PROBLEMA: CALCULAR EL DETERMINANTE Y EL NUMERO DE CONDICION DE *
C              LA SIGUIENTE MATRIZ A. *
C              SE EFECTUA PRIMERO LA DESCOMPOSICION TRIANGULAR DE *
C              LA MATRIZ PARA CALCULAR EL DETERMINANTE. *
*****
C  PROGRAMA PRUEBA PARA LA SUBROUTINA DECOMP
C
      REAL A(5,3), WORK(3), COND, CONDP1
      INTEGER IPVT(3), I, J, N, NDIM
      NDIM = 5
      N = 3
      A(1,1) = 33.0
      A(2,1) = -24.0
      A(3,1) = -8.0
      A(1,2) = 16.0
      A(2,2) = -10.0
      A(3,2) = -4.0
      A(1,3) = 72.0
      A(2,3) = -57.0
      A(3,3) = -17.0
C
      DO 10 I = 1, N
        WRITE(6,21) (A(I,J), J=1,N)
C 10 CONTINUE
C
      CALL DECOMP(NDIM, N, A, COND, IPVT, WORK)
C
      WRITE(6,22)
      PI=1.0
      DO 12 I=1,N
C 12 PI= PI*A(I,I)
      DET= IPVT(N)*PI
      DO 14 I=1,N
C 14 WRITE(6,23) ( A(I,J),J=1,N )
      WRITE(6,24) DET,COND
C
C 21 FORMAT( 5X,4F8.0 )
C 22 FORMAT( //,15X,"MATRIZ FACTORIZADA",/ )
C 23 FORMAT( 5X,3F14.7 )
C 24 FORMAT( //,6X,"EL DETERMINANTE DE A ES ",F5.1,///,6X,
*           "EL NUMERO DE CONDICION DE A ES ",E14.7 )
      STOP
      END
  
```

33.	16.	72
-24.	-10.	-57.
-8.	-4.	-17.

MATRIZ FACTORIZADA

33.0000000	16.0000000	72.0000000
------------	------------	------------

0.7272727	1.6363636	-4.6363636
0.2424242	0.0740741	0.1111111

EL DETERMINANTE DE A ES 6.0

EL NUMERO DE CONDICION DE A ES .7672440E+04

 SOLVE

SOLVE SUBROUTINA FORTRAN PARA OBTENER LA SOLUCION DE UN SISTEMA DE ECUACIONES LINEALES A PARTIR DE LA DESCOMPOSICION TRIANGULAR DE LA MATRIZ COEFICIENTE.

 1. DESCRIPCION

1.1 PROPOSITO:
 DADA LA FACTORIZACION TRIANGULAR LU DE UNA MATRIZ A DE ORDEN N, LOS INDICES DE LOS RENGLONES PERMUTADOS P Y UN VECTOR DERECHO B ESTA SUBROUTINA EFECTUA LA SUSTITUCION HACIA ATRAS PARA OBTENER LA SOLUCION DEL SISTEMA DE ECUACIONES LINEALES $AX=B$.

1.2 METODO:
 EL METODO CONSISTE DE DOS PARTES:

1) ELIMINACION HACIA ADELANTE. EN ESTA PARTE SE RESUELVE EL SISTEMA

$$LY = INV(P) B$$

ESTO ES, SE APLICAN LOS MULTIPLICADORES USADOS DURANTE LA DESCOMPOSICION TRIANGULAR AL VECTOR DERECHO B.

2) SUSTITUCION HACIA ATRAS. EN ESTA PARTE SE RESUELVE EL SISTEMA

$$UX = Y.$$

EL VECTOR X ES ALMACENADO EN EL VECTOR B.

PARA UNA INFORMACION MAS DETALLADA DEL ALGORITMO CONSULTESE LA REFERENCIA 2, PAGES. 127-129.

 2. USO DE LA SUBROUTINA

2.1 USO GENERAL:
 EN EL PROGRAMA PRINCIPAL HAY QUE DIMENSIONAR LOS ARREGLOS QUE USA ESTA SUBROUTINA. ADEMÁS HAY QUE OBTENER PRIMERO LA FACTORIZACION TRIANGULAR DE A USANDO LA SUBROUTINA DECOMP. SI DECOMP DETERMINO QUE A ES SINGULAR ENTONCES SOLVE NO SE USARA.

2.2 DESCRIPCION DE LOS PARAMETROS:
 EL USO DE LA SUBROUTINA SE EFECTUA POR EL SIGUIENTE ENUNCIADO

CALL SOLVE(NDIM,N,A,B,IPVT)

TENIENDO LOS PARAMETROS EL SIGUIENTE SIGNIFICADO:

PARAMETROS DE ENTRADA

NDIM DIMENSION MAXIMA DE LOS RENGLONES DE A, DECLARADO EN EN PROGRAMA PRINCIPAL.
 N ORDEN DE LA MATRIZ A) (N MENOR O IGUAL QUE NDIM).
 A ARREGLO DE DIMENSION NDIM*N QUE CONTIENE LA DESCOMPOSICION TRIANGULAR LU.
 B VECTOR DE DIMENSION N QUE CONTIENE EL VECTOR DERECHO B.
 IPVT VECTOR DE DIMENSION N QUE CONTIENE LOS INDICES DE LAS

FILAS PIVOTES.

PARAMETROS DE SALIDA

B CONTIENE LA SOLUCION CALCULADA X DEL SISTEMA.

2.3 CONDICIONES DE ERROR Y RETORNOS:
NINGUNO

2.4 REQUERIMIENTOS DE MEMORIA:
HAY QUE DIMENSIONAR, ADEMAS DE LOS ARREGLOS A, IPUT, Y W, EL
ARREGLO B(N) EN EL PROGRAMA PRINCIPAL.

2.5 SUBPROGRAMAS REQUERIDOS:
NINGUNO.

3. APLICABILIDAD Y RESTRICCIONES

LA INVERSA $INV(A)$ DE UNA MATRIZ A PUEDE CALCULARSE TAMBIEN APROVE-
CHANDO LA DESCOMPOSICION TRIANGULAR DE A, RESOLVIENDO N ECUACIO-
NES LINEALES,

$$AX=E(I) \quad , \quad I=1,2,\dots,N$$

DONDE E(I) ES EL I-ESIMO VECTOR UNITARIO. LA SOLUCION PARA EL
I-ESIMO LADO DERECHO E(I) CORRESPONDE A LA I-ESIMA COLUMNA DE $INV(A)$

ESTA SUBROUTINA DEBERA USARSE JUNTO CON LA SUBROUTINA DECOMP.

4. REFERENCIAS

1. FORSYTHE, G.E., MALCOLM, M.A. AND MOLER, C.B., COMPUTER METHODS FOR MATHEMATICAL COMPUTATIONS. ENGLEWOOD CLIFFS, PRENTICE-HALL, 1977.
2. CONTE, S.D, AND DE BOOR, C., ELEMENTARY NUMERICAL ANALYSIS: AN ALGORITHMIC APPROACH, 2ND. ED., NEW YORK:MC GRAW-HILL, 1972.
3. STEWART, G.W., INTRODUCTION TO MATRIX COMPUTATIONS. ACADEMIC PRESS, 1973.
4. DAHLQUIST, G, AND BJORCK, A., NUMERICAL METHODS. ENGLEWOOD CLIFFS, N.J.:PRENTICE-HALL, 1974.

5. OBSERVACIONES Y SUGERENCIAS

EL METODO DE SUSTITUCION HACIA ATRAS ES LA ETAPA COMPLEMENTARIA PARA RESOLVER LA ECUACION $AX=B$. ESTA SUBROUTINA NO DEBERA USARSE SI LA SUBROUTINA DECOMP DETERMINA QUE A ES SINGULAR. ES DECIR, SI $COND1=COND+1$, ENTONCES A ES SINGULAR.

6. EJEMPLO DE PRUEBA

LA SOLUCION DEL SISTEMA DE ECUACIONES LINEALES

$$\begin{aligned} 33X1 + 16X2 + 72X3^0 &= -359 \\ -24X1 - 10X2 - 57X3 &= 281 \\ -8X1 - 4X2 - 17X3 &= 85 \end{aligned}$$

FUE CALCULADA, EN UNA COMPUTADORA BURROUGHS 6800 CON ARITMETICA EN SIMPLE PRECISION.

PARAMETROS: NDIH=5, N=3

RESPUESTA DEL ALGORITMO :

X1 = 1.0
X2 = -2.0
X3 = -5.0

APENDICE A. LISTADO DEL PROGRAMA PRUEBA.

```
4 SET AUTOBIND
  BIND = FROM (ISAK)OBJECT/BIBDAN/=
  RESET FREE
C *****
C PROBLEMA: RESOLVER EL SIGUIENTE SISTEMA DE ECUACIONES LINEALES*
C           A X = B, DADOS A Y B. *
C           SE USA LA SUBROUTINA DECOMP PARA EFECTUAR LA. *
C           DESCOMPOSICION TRIANGULAR DE LA MATRIZ A. *
C *****
C C
C C PROGRAM A PRUEBA PARA LA SUBROUTINA SOLVE
C
  REAL A(5,3),B(3),WORK(3),COND,CONDP1
  INTEGER IPVT(3), I, J, N, NDIH
  NDIH = 5
  N = 3
  A(1,1) = 33.0
  A(2,1) = -24.0
  A(3,1) = -8.0
  A(1,2) = 16.0
  A(2,2) = -10.0
  A(3,2) = -4.0
  A(1,3) = 72.0
  A(2,3) = -57.0
  A(3,3) = -17.0
  B(1) = -359.0
  B(2) = 281.0
  B(3) = 85.0
C
  DO 10 I = 1, N
    WRITE(6,21) (A(I,J), J=1,N),B(I)
  10 CONTINUE
C
  CALL DECOMP(NDIH, N, A, COND, IPVT, WORK)
  WRITE(6,22) COND
C
  CONDP1 = COND + 1
  IF (CONDP1 .EQ. COND) GO TO 20
C
  CALL SOLVE(NDIH, N, A, B, IPVT)
C
  WRITE(6,23) ( I,B(I),I=1,N )
  STOP
  20 WRITE(6,24) CONDP1
  21 FORMAT( 5X,8F4.0 )
  22 FORMAT( /,5X,"COND=",E14.7,/)
C
```

```
23 FORMAT( 5X,'X(',I2,')=',E14.7 )
24 FORMAT( 5X,'LA MATRIZ ES SINGULAR. CONDP1=',E14.7 )
STOP
END
```

33.	16.	72.	-359.
-24.	-10.	-57.	281.
-8.	-4.	-17.	85.

COND= .7672440E+04

X(1)= .1000000E+01
X(2)= -.2000000E+01
X(3)= -.5000000E+01

A

L Y D E R 1

SUBROUTINA FORTRAN PARA MINIMIZAR LA SUMA DE LOS CUADRADOS DE FUNCIONES NO LINEALES

1. PROPOSITO

LA SUBROUTINA LYDER1 UTILIZA UNA MODIFICACION DEL ALGORITMO DE LEVENBERG-MARQUARDT PARA MINIMIZAR LA SUMA DE LOS CUADRADOS DE M FUNCIONES NO LINEALES EN N VARIABLES, M.G.E. N, DE LA FORMA

$$F(X) = \sum R(X) = R(X) R(X).$$

EL USUARIO DEBERA DAR UNA SUBROUTINA PARA EVALUAR LAS FUNCIONES Y SU JACOBIANO.

2. METODO

EL METODO NUMERICO USADO EN ESTA PASADA ES UNA VERSION MODIFICADA DEL ALGORITMO DE LEVENBERG-MARQUARDT EN DONDE, A PARTIR DE UNA DADA, LA SUBROUTINA GENERA UNA SUCCESION DE PUNTOS QUE SE APROXIMAN A UN MINIMO LOCAL DE LA FUNCION F(X). LAS MODIFICACIONES MAS IMPORTANTES DEL ALGORITMO DE LEVENBERG-MARQUARDT SON EL USO DE VARIABLES ESCALADAS IMPLICITAMENTE Y UNA PERMISIVIDAD OPTIMA DE LA LONGITUD DE LA CORRECCION EN CUALQUIER DIRECCION DONDE LA FUNCION CAMBIANDO RAPIDAMENTE. LA ELECCION OPTIMA DE LA CORRECCION GARANTIZA LA CONVERGENCIA AUN CUANDO LA APROXIMACION INICIAL SE ENCONTRA LEJOS DE LA SOLUCION Y LA CONVERGENCIA ES MUY RAPIDA CUANDO SE TIENEN SOLUCIONES CON RESIDUO PEQUEÑO. EL PARAMETRO DE LEVENBERG-MARQUARDT SE CALCULA SIGUIENDO UN ESQUEMA DEBIDO A HEDDEN. LA CONVERGENCIA DEL ALGORITMO SE OBTIENE CUANDO SE CUMPLE CUALQUIERA DE LOS TRES TIPOS DE COMPACTACIONES QUE SE EFECTUAN ENTRE LA APROXIMACION X Y UNA MEJOR SOLUCION XSOL.

PARA UNA INFORMACION MAS DETALLADA DEL ALGORITMO CONSULTESE LAS REFERENCIAS 1 Y 2.

3. COMO USAR LA SUBROUTINA

3.1 USO GENERAL

EL USUARIO EN SU PROGRAMA PRINCIPAL DEBERA DIMENSIONAR LOS ARREGLOS POR USAR, ESPECIFICAR EL NUMERO DE FUNCIONES A MINIMIZAR, EL NUMERO DE VARIABLES, UNA TOLERANCIA DEL ERROR PARA INDICAR QUE TAN EXACTA DEBEVA A CALCULAR LA SOLUCION, Y UNA SUBROUTINA PARA EVALUAR

LA FUNCION $R(X)$ Y SU JACOBIANO. LA SUBROUTINA REQUIERE DE UNA APROXIMACION INICIAL PARA EMPEZAR EL PROCESO DE MINIMIZACION. SI LA SUBROUTINA TERMINA EXITOSAMENTE EL MINIMO ESTARA EN X EN CASO CONTRARIO LA SUBROUTINA DARA LA INFORMACION NECESARIA PARA SABER QUE TIPO DE ERROR OCURRIO.

3.2 DESCRIPCION DE LOS PARAMETROS

EL USO DE LA SUBROUTINA SE EFECTUA CON EL SIGUIENTE ENUNCIADO

CALL LMDER1 (FCN, M, N, X, F, JAC, MDIM, TOL, INFO, IPRINT,
PVT, W, LW)

TENIENDO LOS PARAMETROS EL SIGUIENTE SIGNIFICADO

PARAMETROS DE ENTRADA

FCN ES EL NOMBRE DE LA SUBROUTINA DADA POR EL USUARIO PARA EVALUAR LA FUNCION Y SU JACOBIANO. SE DEBERA DECLARAR EXTERNAL EN EL PROGRAMA PRINCIPAL.

M ES UNA VARIABLE ENTERA QUE ESPECIFICA EL NUMERO DE FUNCIONES A MINIMIZAR.

N ES UNA VARIABLE ENTERA QUE REPRESENTA EL NUMERO DE VARIABLES. N.LE. M

X ES UN ARREGLO REAL DE DIMENSION N QUE CONTIENE LA APROXIMACION INICIAL DEL MINIMO.

MDIM ES UNA VARIABLE ENTERA QUE ESPECIFICA EL NUMERO MAXIMO DE RENGLONES DEL ARREGLO JAC TAL Y COMO FUE ESPECIFICADO EN EL PROGRAMA PRINCIPAL. MDIM.GE. M.

TOL ES UNA VARIABLE REAL NO NEGATIVA QUE REPRESENTA LA TOLERANCIA DEL ERROR REPERADO EN LA SOLUCION. SE OBTIENE LA CONVERGENCIA DEL ALGORITMO CUANDO EL ERROR RELATIVO EN LA SUMA DE LOS CUADRADOS SEA MENOR O IGUAL A TOL, O CUANDO EL ERROR RELATIVO ENTRE X Y LA SOLUCION XSOL SEA MENOR O IGUAL A TOL. EL ALGORITMO CONVERGE TAMBIEN SI $F(X)$ ES ORTOGONAL A LAS COLUMNAS DEL GRADIENTE DE $F(X)$, CON RESPECTO A LA PRECISION DE LA MAQUINA. SI TOL ES MAS PEQUEÑO QUE LA PRECISION DE LA MAQUINA LMDER1 INTENTARA SATISFACER SOLAMENTE LA PRUEBA DEFINIDA POR LA PRECISION DE LA MAQUINA. POR LO TANTO, UN VALOR RECOMENDABLE PARA TOL ES LA RAIZ CUADRA DE LA PRECISION DE LA MAQUINA.

IPRINT NUMERO ENTERO POSITIVO QUE CONTROLA LA ESCRITURA DE LOS RESULTADOS INTERMEDIOS.
 = 0 SI NO SE DESEA NINGUNA IMPRESION DE RESULTADOS.
 = K > 0 SI SE DESEA IMPRIMIR CADA K-ITERACIONES. ESTA ES UNA OPCION UTIL ESPECIALMENTE SI EL USUARIO REQUIERE MUCHA INFORMACION DE LA ITERACION. EN CASO CONTRARIO SE RECOMIENDA EVITAR ESTA OPCION.

W ES UN ARREGLO REAL DE LONGITUD LW USADO COMO UN ARREGLO DE TRABAJO

LW ES UNA VARIABLE ENTERA QUE ESPECIFICA LA LONGITUD DEL ARREGLO W. LW.GE. 5*N + M.

PARAMETROS DE SALIDA

- X CONTIENE EL MINIMO CALCULADO.
- F ES UN VECTOR REAL DE DIMENSION M QUE CONTIENE LOS VALORES DE LAS FUNCIONES EN EL VECTOR X.
- JAC ES UN APREGLO REAL DE DIMENSION MDIM*N QUE CONTIENE EN UNA SUBMATRIZ DE DIMENSION N*N UNA MATRIZ TRIANGULAR SUPERIOR R CON ELEMENTOS DIAGONALES NO CRECIENTES TAL QUE
- $$P (J J) P = R R$$
- DONDE P ES UNA MATRIZ DE PERMUTACION Y J ES EL JACOBIANO FINAL CALCULADO. LA J-ESIMA COLUMNA DE P ES LA IPVT(J)-ESIMA COLUMNA DE LA MATRIZ IDENTIDAD.
- INFO ES UNA VARIABLE ENTERA QUE ESPECIFICA LA CONDICION QUE CAUSO EL REGRESO AL PROGRAMA PRINCIPAL.
- = 0 SI LOS PARAMETROS DE ENTRADA FUERON INVALIDOS.
 - = 1, 2, 3, O 4 PARA UN REGRESO NORMAL.
 - > 4 SI LA SUBROUTINA FALLO.
- IPVT ES UN APREGLO ENTERO DE DIMENSION N QUE DEFINE A LA MATRIZ DE PERMUTACION P TAL QUE JP=OR, DONDE J ES EL JACOBIANO FINAL, Q ES UNA MATRIZ ORTOGONAL (NO ALMACENADA) Y R ES UNA MATRIZ TRIANGULAR SUPERIOR.

3.3 DETECCION DE ERRORES Y RETORNOS

SI LA SUBROUTINA CALCULO EL MINIMO SIN NINGUNA COMPLICACION ENTONCES EL VALOR DE INFO EN LA SALIDA DE LMDER1 SERA 1, 2, 3, O 4.

- INFO = 1, SI $|| F(X) || \leq (1 + \text{TOL}) * || F(XSOL) ||$. EN ESTE CASO SI $\text{TOL} = 10^{-(K)}$, LA NORMA FINAL DE $F(X)$ TIENE K DIGITOS DECIMALES SIGNIFICATIVOS.
- INFO = 2, SI $|| D(X-XSOL) || \leq \text{TOL} * || D * XSOL ||$. EN ESTE CASO SI $\text{TOL} = 10^{-(K)}$ ENTONCES LA COMPONENTE MAS GRANDE DE $D * X$ TIENE K DIGITOS SIGNIFICATIVOS. HAY UN PELIGRO CUANDO LAS COMPONENTES MAS PEQUEÑAS DE $D * X$ PUEDAN TENER ERRORES RELATIVOS GRANDES, PERC LA ELECCION DE D ES TAL QUE LA EXACTITUD DE LAS COMPONENTES DE X ESTEN RELACIONADAS A SU SENSIBILIDAD.
- INFO = 3, SI SE CUMPLE TANTO PARA INFO = 1 COMO INFO = 2.
- INFO = 4, SI $F(X)$ ES ORTOGONAL A LAS COLUMNAS DEL JACOBIANO DE $F(X)$. ESTE CASO HABRA QUE EXAMINARLO CUIDADOSAMENTE, YA QUE LA PRUEBA SE PUEDE SATISFACER TAMBIEN EN OTROS PUNTOS CRITICOS, TALES COMO UN PUNTO MAXIMO Y UN PUNTO SILLA.

LOS POSIBLES ERRORES QUE PUEDEN RESULTAR SON AHORA DESCritos, ALGUNOS SON MAS DESASTROSOS QUE OTROS Y EL USUARIO PUEDE DECIDIR, EN CADA CASO, QUE ACCION TOMAR.

SI 200*(4+1) EVALUACIONES DE LA FUNCION NO FUERON SUFICIENTES PARA

ENCONTRAR EL MINIMO ENTONCES INFO REGRESARA CON EL VALOR 5. EN
ESTE CASO ES POSIBLE CONTINUAR CON EL ULTIMO PUNTO CALCULADO,
USANDOLO COMO UNA APROXIMACION INICIAL PARA INTENTAR OTRO NUMERO
IGUAL DE EVALUACIONES.

SI ES IMPOSIBLE REDUCIR EL RESIDUO CON TOL MUY PEQUENA ENTON-
CES INFO REGRESARA CON EL VALOR 5.

SI LA SOLUCION APROXIMADA ES IMPOSIBLE MEJORARLA CON TOL MUY
PEQUENA ENTONCES INFO REGRESARA CON EL VALOR 7.

SI LOS PARAMETROS DE ENTRADA SON INVALIDOS ENTONCES INFO REGRESARA
CON EL VALOR 0.

3.4 SUBPROGRAMAS REQUERIDOS:

FCN ES EL NOMBRE DE LA SUBROUTINA DADA POR EL USUARIO PARA
EVALUAR LA FUNCION Y SU JACOBIANO. DEBERA DECLARARSE
EXTERNAL EN EL PROGRAMA PRINCIPAL Y DEBERA ESCRIBIRSE
DE LA FORMA SIGUIENTE:

```

SUBROUTINE FCN(M,N,X,F,FJAC,MDIM,IFLAG)
INTEGER M,N,MDIM,IFLAG
REAL X(N), F(X), FJAC(MDIM,N)
-----
SI IFLAG = 1  EVALUAR LA FUNCION EN X Y REGRESAR
               EL VALOR EN F NO ALTERAR JAC
SI IFLAG = 2  EVALUAR EL JACOBIANO EN X Y REGRESAR
               ESTE VALOR EN LA MATRIZ FJAC. EL
               I-ESIMO PUNGLON DE FJAC ES EL GRADIENTE
               DE LA I-ESIMA FUNCION CON RESPECTO A X.
               NO ALTERAR F.
-----
RETURN
END

```

LA SUBROUTINA NO DEBERA CAMBIAR EL VALOR DE IFLAG, A MENOS DE QUE
EL USUARIO DEBE TERMINAR LA EJECUCION DE LMDERT. SI ESTE ES EL
CASO, HAY QUE ASIGNARLE UN VALOR NEGATIVO A IFLAG.

APLICABILIDAD Y RESTRICCIONES

1. LA SOLUCION DE UN SISTEMA DE M ECUACIONES NO LINEALES CON N
VARIABLES, $G(X)=0$, CON M MAYOR O IGUAL QUE N, PUEDE CALCULARSE
AL MINIMIZAR LA FUNCION,

$$F(X) = \sum G(X) \quad !! = G(X) \quad G(X)$$

SI X ES EL MINIMO CALCULADO, TAL QUE $F(X) = 0$, ENTONCES X
ES LA SOLUCION DEL SISTEMA DE ECUACIONES NO LINEALES.

2. EL PROBLEMA DE AJUSTAR UN CONJUNTO DE DATOS DADO CON UNA
CURVA EN EL SENTIDO DE MINIMOS CUADRADOS NO LINEAL, PUEDE
RESOLVERSE UTILIZANDO ESTA SUBROUTINA, MINIMIZANDO LA
FUNCION,

$$F(X) = \sum G(X;W) - Y \quad !!$$

DONDE LA NORMA UTILIZADA ES LA NORMA 2. $G(X;W)$ ES LA
 FUNCION NO LINEAL QUE SE DESEA AJUSTAR AL CONJUNTO DE DATOS
 $(W(I), Y(I))$, $I=1, \dots, M$; X ES UN VECTOR DE DIMENSION N .
 ADEMÁS, M ES MAYOR O IGUAL A N .

3. EL MAXIMO DE LA SUMA DE LOS CUADRADOS DE M FUNCIONES NO LINEALES
 $G(X)$, PUEDE SER CALCULADO USANDO L'YDEP, AL MINIMIZAR LA
 FUNCION $F(X) = -G(X)$, RECORDANDO QUE

$$\max G(X) = -\min F(X).$$

5. REFERENCIAS

- MORE, J.J., THE LEVENBERG-MARQUARDT ALGORITHM: IMPLEMENTATION
 AND THEORY. IN NUMERICAL ANALYSIS, G.A. WATSON (EDITOR), LECTURES
 NOTES IN MATHEMATICS, NO. 430, SPRINGER VERLAG, NEW YORK, 1977,
 105-114.
- MORE, J.J., IMPLEMENTATION AND TESTING OF OPTIMIZATION SOFTWARE.
 EN PERFORMANCE EVALUATION OF NUMERICAL SOFTWARE, L.D. FOSDICK,
 EDITOR, NORTH-HOLLAND, 1979, 253-266.

6. OBSERVACIONES Y SUGERENCIAS

LA SUBROUTINA NO INCORPORA NINGUNA ESTRATEGIA PARA OBTENER LA
 CONVERGENCIA GLOBAL. LA CONVERGENCIA ESTA GARANTIZADA SI LA APROXI-
 MACION INICIAL ESTA MUY CERCA DEL MINIMO Y OCURRE EN UN MINIMO ---
 LOCAL.

PARA RESOLVER UN SISTEMA DE ECUACIONES NO LINEALES ES MAS VENTA-
 JOSO UTILIZAR UNA SUBROUTINA ADECUADA PARA ESTOS PROBLEMAS, PUESTO QUE
 L'YDEP REQUIERE LAS DERIVADAS PARCIALES DE LAS ECUACIONES Y EN LA
 MAYORIA DE LOS CASOS ES IMPOSIBLE TENERLAS EXPLICITAMENTE, Y AUN
 CUANDO SE TENGAN SON MUY DIFICILES DE EVALUAR. SE RECOMIENDA USAR
 BIENTI, HIBRID O L'YDEP, SEGUN EL TIPO DE PROBLEMA.

7. EJEMPLO DE PRUEBA

EL PROBLEMA ES DETERMINAR LOS VALORES DE $X(1)$, $X(2)$, Y $X(3)$
 QUE DAN EL MEJOR AJUSTE, EN EL SENTIDO DE MINIMOS CUADRADOS, DE LA
 FUNCION

$$G(X;W) = X(1) + X(2) + \exp(W(I)) + X(3), \quad I=1, \dots, 6$$

A LOS DATOS

$$Y = (127, 151, 379, 421, 460, 426)$$

DONDE $W = (-5, -3, -1, 1, 3, 5)$. LA I -ESIMA COMPONENTE DE $F(X)$
 ESTA DEFINIDA POR

Las subrutinas del paquete IMSL que se utilizaron, no se incluyeron en este apéndice, por considerar que dicho paquete es bastante usual. El lector interesado puede consultar IMSL (1980).

Las subrutinas usadas en este trabajo fueron:

- FTAUTO
- _ FTARPS
- FTMPS
- _ FTGEN
- _ ZRPOLY

REFERENCIAS

AIGNER, D. J. (1971). A Compendium on Estimation of the Autoregressive - Moving Average Model from Time Series Data., International Economic Review, XII.

ANDERSON, T. W. (1971). The Statistical Analysis of Time Series., John Wiley and Sons: New York.

BARD, Y. (1974). Nonlinear Parameter Estimation., Academic Press: New York.

BOX and JENKINS (1970). Time Series Analysis, Forecasting and Control., Holden Day: San Francisco.

BUSTOS, A. y AHSANULLAH, M. (1984). On the Estimates of the Parameters of the First Order Autoregressive Process., Communications in Statistics, Theory and Methods Simulation and Computation., 13(4), 489-505.

CALDERON, A. y NOCEDAL, J. (1981). Guía para el uso de

la Biblioteca Básica de Programas de Análisis Numérico.,
Comunicaciones Técnicas, Serie Azul, No. 50, IIMAS, UNAM.

CHATFIELD, C. (1974). The Analysis of Time Series,
Theory and Practice., Chapman and Hall: London.

DRAPER, N. R. y SMITH, H. (1981). Applied Regression
Analysis, Segunda Edición., John Wiley and Sons: New York.

FARAH, J. L. y ESTANDIA, A. (1980). Identificación de
Modelos ARMA de Procesos Estacionarios Escalares., Notas de
Clase, IIMAS, UNAM.

FULLER, W.A. (1976). Introduction to Statistical Time
Series., John Wiley and Sons: New York.

GRANGER, C.W. y NEWBOLD, P. (1977). Forecasting
Economic Time Series., Academic Press: New York.

GRENANDER, U. y ROSENBLATT, H.M. (1957). Statistical
Analysis of Stationary Time Series., John Wiley and Sons:
New York. Chichester: England.

HANNAN, E. J. (1970). Multiple Time Series., Wiley: New
York.

HARVEY, A. C. (1981). Time Series Models., Philip
Allan: Oxford.

IMSL, Library Reference Manual (1980): U. S. A.

CALDERON, A. y NOCEDAL, J. (1981). Guía para el uso de la Biblioteca Básica de Programas de Análisis Numérico., Comunicaciones Técnicas, Serie Azul, No.50, IIMAS, UNAM.

KENDALL, M.G. y STUART, A. (1966). The Advanced Theory of Statistics, Vol. 3., Griffin: London.

NERLOVE, M., GREY, D. M., CARVALHO, J. L. (1979). Analysis of Economic Time Series a Synthesis., Academic Press: New York.

NEWBOLD, P. (1974). The Exact Likelihood Function for a Mixed Autoregressive - Moving Average Process., Biometrika 61, 3, p. 423.

PACK, D. J. (1974). computer Programs for the Analysis of Univariate Time Series Models and Single Input Transfer Function Models using the Methods of Boz and Jenkins., The Ohio State University.

PAGAN, A. (1970). The Estimation of Models with Moving Average Disturbance Terms., Australian National University.

PHILLIPS, A. W. (1966). Estimation of System of Difference Equations with Moving Average Disturbance.,

Australian National University.

SHAMAN, P. (1973). On the Inverse of the Covariance Matrix for an Autoregressive - Moving Average Process., *Biometrika* 60, 1, p. 193.

TIAO, G. C. y ALI, M. M. (1971). Analysis of Correlated Random Effects: Linear Model with two Random Componets., *Biometrika* 58, 37-51.

WILSON, G. T. (1969). Factorization of the Generating Funtions of a Pure Moving Average Process., *Siam J. Num. Anal.* 6, 1-7.

WINCEK, M. A. y REINSEL, G. C. (1984). An Exact Maximum Likelihood Estimation Procedure for Regression - ARMA Time Series Models with possibly non consecutive data., Technical Report No. 749, The University of Wisconsin.