Identificación de un modelo ARIMA cuando existen observaciones faltantes

Introducción. I. El proceso de identificación del modelo y estimación de las observaciones faltantes. II. Aplicación un caso simulado. Conclusiones. Referencias

Introducción

Un supuesto común en el análisis de series de tiempo es que las series que van a ser estudiadas disponen de información para cada momento de tiempo del período que se va analizar. Sin embargo, con frecuencia ocurre que faltan datos en la serie, o que algunos de ellos son erróneos. Por ejemplo, en series de tiempo económicas es relativamente común encontrarse con observaciones no registradas debido a huelgas, o a pérdida de la información, o aún a que el dato que existe debe ser eliminado por considerarse equivocado.

En la literatura de Análisis Series de Tiempo, en particular en la de los procesos ARIMA (Box y Jenkins, 1976), se han propuesto diferentes métodos para estimar estas observaciones, pero la mayoría de ellos supone que el **modelo es conocido** o que las observaciones son tales que han permitido identificarlo. Algunos de estos procedimientos se encuentran en Chow y Lin (1976), Jones (1980), Kohn y Ansley (1983), Harvey y Pierse (1984), Maraval y Peña (1988), Nieto (1989), Peña y Maraval (1990), Chen y Liu (1990) y Castañeda (1994). Sin embargo, en la práctica, la distribución de las

observaciones faltantes en la muestra puede ser tal que impida la identificación del modelo ARIMA adecuado.

Este documento presenta una metodología relativamente simple que permite estimar las observaciones faltantes y simultáneamente identificar el modelo ARIMA que generó una serie de tiempo. El procedimiento consta de dos etapas: la primera de ellas produce estimaciones preliminares de los datos faltantes a partir de una aproximación del modelo ARIMA por medio de un modelo autorregresivo de alto orden y el uso del análisis de intervención; en la segunda etapa se identifica el modelo usando las estimaciones preliminares y se reestiman las observaciones usando de nuevo el análisis de intervención sobre el modelo identificado. Experimentos de simulación muestran que el procedimiento parece funcionar adecuadamente.

I. El proceso de identificación del modelo y estimación de las observaciones faltantes

Suponga que Zt es una serie de tiempo que sigue una proceso ARIMA de la forma

$$Zt = \frac{\theta(B)}{\delta(B)\phi(B)} a_t$$

donde B es el operador usual de rezagos, $_{\theta}$(B) es el polinomio de medias móviles con todas sus raíces fuera del círculo unidad, $_{\phi}$(B) es el polinomio autorregresivo con sus raíces fuera del círculo unidad y que no tiene factores comunes con $_{\theta}$(B) y $_{\delta}$(B) es el polinomio de diferencias (que induce estacionaridad) con sus raíces sobre el círculo unidad. En primer lugar consideremos el caso donde hay solamente una observación faltante. Para esto, suponga que la serie se observó durante T períodos y que no se encuentra disponible la observación para el período t=t*. Una caracterización natural de un valor faltante es describirlo como una observación **atípica aditiva.** Esta caracterización ha sido empleada por varios autores entre los que se encuentran Peña y Maraval (1990) y Liu y Chen (1990). La razón es la

siguiente: si suponemos que en el período t=t* ocurre una observación atípica aditiva, podemos representar la serie observada como:

$$NZ_t = Z_t \qquad \text{si } t \neq t^*$$

$$= AI_t^{(t^*)} + Z_t \qquad \text{si } t = t^*$$

donde A indica la cantidad de desviación desde el verdadero valor de Z_{t^*}, y la variable $It^{(t^*)}$ toma el valor de uno cuando t=t* y de cero en otro caso.

En este caso, Chen y Liu (1990) mostraron que el valor ajustado de NZ_{t^*} (es decir, después de remover el efecto atípico sobre NZ_{t^*}) es:

$$NZ^*_{t^*} = \{ \sum_{j=1}^{t-1} [\sum_{k=j}^{n-t^*+j} \pi_k \pi_{k-j}] NZ_{t^*-j} + \sum_{j=1}^{n-t^*} [\sum_{k=j}^{n-t^*} \pi_k \pi_{k-j}] NZ_{t^*+j} \} / \sum_{j=0}^{n-t^*} \pi_j^2 \qquad (1)$$

donde los coeficientes π son obtenidos del polinomio auto-rregresivo $\pi(B)=1-\pi 1B-\pi 2B2-\ldots = \phi(B)/\theta(B)$. Si la serie no es estacionaria entonces la parte autorregresiva debe incluir el operador que induce estacionaridad.

De acuerdo con (1), el valor interpolado $NZ^*_{t^*}$ está basado en las observaciones de la serie anteriores y posteriores de NZ_{t^*}, es decir, en los valores anteriores y posteriores de la serie original Z a la observación faltante Z_{t^*}. Por lo tanto el valor ajustado no tiene nada que ver con la observación atípica NZ_{t^*}.

El resultado anterior sugiere que podemos estimar un valor faltante en una serie de tiempo tratándole como si fuera una observación atípica aditiva.

Por tanto, si **conocemos a** $\theta(B)$, $\phi(B)$ **y** $\delta(B)$, **el** procedimiento para estimar el valor desconocido de Z_{t^*} consiste en asignar un valor atípico cualquiera a la observación faltante y estimar el modelo intervenido:

$$NZ_t = A\, I_t^{(t^*)} + \frac{\theta(B)}{\delta(B)\phi(B)} a_t ,$$

donde NZ_t es igual a la serie Z_t con la observación faltante reemplazada por un valor atípico, A es el coeficiente que indica el impacto de la observación atípica aditiva sobre el nivel medio de la serie en el período t=t*, y donde $It^{(t^*)}$ es una variable indicadora con un uno en el período t=t* y ceros en los demás períodos. La estimación óptima de Z_{t^*} es el valor predicho de NZ_{t^*} menos la estimación de A (Box y Tiao, 1975).

De igual forma se procedería si la serie tuviera m observaciones perdidas no consecutivas en los períodos T1, T2,...,Tm: A cada período donde se desconoce la observación se asigna un valor atípico y en el modelo

$$N Z_t = \sum_{j=1}^{m} A_j I_t^{(T_j)} + \frac{\theta(B)}{\delta(B)\phi(B)} a_t$$

la estimación de Z_{Tj} se obtiene como la predicción de NZ_{Tj} menos la estimación de Aj en el modelo anterior, para j=1,2,...,m..

Sin embargo, el problema en la práctica es más complicado pues, en general, no se dispone del conocimiento de $\theta(B)$, $\phi(B)$ y $\delta(B)$ y debemos tratar de identificarlos a partir de la información incompleta inicial. A continuación se presenta el proceso de identificación propuesto para un modelo donde existen observaciones faltantes:

i) Aproxime a Z_t usando un proceso puro autorregresivo de orden alto. En la elección del orden se deben tener en cuenta la frecuencia del período de observación, y la clase de proceso (estacional o no). En otras palabras Zt puede aproximarse como el proceso autorregresivo puro:

$$Z_t = \frac{1}{\delta(B)\ \phi'(B)} a_t$$

Donde $\phi'(B)$ es el polinomio autorregresivo del orden seleccionado. Para la aproximación de un proceso ARIMA(p,1,q), Said y Dickey (1984) consiguieron el siguiente resultado: Todo proceso ARIMA(p,1,q) puede ser

adecuadamente aproximado por medio de un proceso ARIMA(n,1,0), donde $n \leq T^{1/3}$

ii)A cada una de las observaciones faltantes asigne un valor que sea atípico y ajuste el modelo

$$N Z_t = \sum_{j=1}^{m} A_j I_t^{(T_j)} + \frac{1}{\delta(B)\ \phi'(B)} a_t$$

Estime el modelo anterior y haga un análisis de residuales para verificar la buena aproximación a la estructura de Z_t del modelo estimado. Si los residuales se comportan como ruido blanco el orden elegido aproxima adecuadamente la estructura del proceso.

Una estimación preliminar de las observaciones faltantes está definida por los valores predichos del modelo aproximado estimado y corregido por la estimación del respectivo impacto. Esta estimación tiene en cuenta la estructura de autocorrelación aproximada de la serie.

iii) Reemplace los valores atípicos de la serie por los valores estimados preliminares. Sobre esta serie pueden emplearse las técnicas de identificación de Box-Jenkins para los modelos ARIMA.

iv) Suponga que el modelo identificado es de la forma

$$Zt = \frac{\theta(B)}{\delta(B)\phi(B)} a_t$$

Para obtener una estimación más refinada de las observaciones faltantes ajuste el modelo

$$N Z_t = \sum_{j=1}^{m} A_j I_t^{(T_j)} + \frac{\theta(B)}{\delta(B)\phi(B)} a_t$$

El valor predicho de NZ para el período Tj menos la estimación de Aj es una estimación óptima de Z_{Tj}, j=1,2,...,m.

II. Aplicación: un caso simulado

A continuación presentamos la aplicación del procedimiento de estimación de datos faltantes e identificación del modelo ARIMA a un caso simulado usando el paquete SCA (Scientific Computing Associates). Se presentan los resultados de una simulación donde se aplica en detalle el procedimiento propuesto, y de 1000 simulaciones para tres elecciones del orden del modelo autorregresivo inicial: p=5, p=8 y p=15.

El modelo simulado fué:

$$(1\text{-}B)Z_t = 1 + (1\text{-}.7B+.45B^2)a_t$$

Este es un modelo ARIMA (0,1,2) invertible con las raíces del polinomio de medias móviles fuera del círculo unidad (Las raíces del polinomio son el par de raíces complejas conjugadas 0.777778 ± 1.271725, con un módulo 1.490712). Para el término de error a_t se consideró una distribución n(0,4). Se generaron 150 observaciones de las cuales se eliminaron las observaciones pertenecientes a los períodos T_1=20, T_2=70, T_3=80, T_4=125 y T_1=135. Es decir, suponemos que en la serie faltan cinco observaciones pertenecientes a los períodos mencionados.

A. Aplicación del procedimiento a un caso simulado:

El procedimiento aplicado a la serie simulada es el siguiente:

i) El primer paso es definir un proceso autorregresivo de orden p suficientemente alto que permita captar la estructura de dependencia de las observaciones. Para este caso se tomó p=8 (el resultado de Said y Dickey (1984) sugiere que tomando p=5 podría aproximarse suficientemente bien el modelo; una comparación para diferentes valores de p se encuentra en la siguiente sección).

ii) Asignar valores atípicos a los períodos donde faltan observaciones.

En nuestro caso se asignaron los valores:

$NZ_{20} = 150$
$NZ_{70} = 200$
$NZ_{80} = 200$
$NZ_{125} = 200$
$NZ_{135} = 200$

Ahora bien, no es posible emplear las funciones de autocorrelación y autocorrelación parcial para tratar de identificar el tipo y el orden del proceso debido a los efectos de las observaciones atípicas introducidas (véase por ejemplo Wei, 1990). Para ilustrar lo anterior se encontraron las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de la serie. Los resultados son los siguientes:

AUTOCORRELACIONES

```
          -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
           +----+----+----+----+----+----+----+----+----+----+----+
                                      I
  1   -.52              XXXXXXXXX+XXXI   +
  2    .03                       +    IX   +
  3    .00                       +    I    +
  4   -.01                       +    I    +
  5    .00                       +    I    +
  6    .01                       +    I    +
  7   -.01                       +    I    +
  8    .01                       +    I    +
  9   -.18                       +XXXXI    +
 10    .33                       +    IXXXX+XXX
 11   -.17                       +XXXXI    +
 12   -.01                      +     I     +
 13    .02                      +     I     +
 14   -.01                      +     I     +
 15    .00                      +     I     +
```

AUTOCORRELACIONES PARCIALES

```
                -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                +----+----+----+----+----+----+----+----+----+----+----+
                                          I
  1    -.52                XXXXXXXXX+XXXI   +
  2    -.33                   XXXX+XXXI     +
  3    -.24                     XX+XXXI     +
  4    -.19                      X+XXXI     +
  5    -.17                       XXXXI     +
  6    -.12                      +XXXXI     +
  7    -.11                      +XXXXI     +
  8    -.08                      + XXXI     +
  9    -.38                   XXXXX+XXXI    +
 10    -.02                         +  I    +
 11     .01                         +  I    +
 12    -.02                         +  XI   +
 13    -.01                         +  I    +
 14    -.01                         +  I    +
 15    -.03                         +  XI   +
```

De los resultados anteriores no se desprende claramente la información para identificar el proceso como un ARIMA(0,1,2).

A continuación se estima el modelo autorregresivo con p=8 y con las 5 intervenciones definidas anteriormente.

PARAMETER	VARIABLE	NUM./	FACTOR	ORDER	CONS-	VALUE	STD	T
LABEL	NAME	DENOM.			TRAINT		ERROR	VALUE
1 CONST		CNST	1	0	NONE	.8856	.0968	9.15
2 A20	INT20	NUM.	1	0	NONE	112.9557	1.4718	76.75
3 A70	INT70	NUM.	1	0	NONE	115.9351	1.4878	77.92
4 A80	INT80	NUM.	1	0	NONE	107.7666	1.4845	72.59
5 A125	INT125	NUM.	1	0	NONE	69.7186	1.5310	45.54
6 A135	INT135	NUM.	1	0	NONE	64.1568	1.4892	43.08
7 P1	ZSSIN	D-AR	1	1	NONE	-.7301	.0867	-8.42
8 P2	ZSSIN	D-AR	1	2	NONE	.0172	.1076	.16
9 P3	ZSSIN	D-AR	1	3	NONE	.2986	.1065	2.80
10 P4	ZSSIN	D-AR	1	4	NONE	.1278	.1101	1.16
11 P5	ZSSIN	D-AR	1	5	NONE	-.0186	.1114	-.17
12 P6	ZSSIN	D-AR	1	6	NONE	-.0591	.1115	-.53
13 P7	ZSSIN	D-AR	1	7	NONE	-.1416	.1113	-1.27
14 P8	ZSSIN	D-AR	1	8	NONE	-.1277	.1116	-1.14

Un chequeo de los residuales nos permite determinar si el orden seleccionado es adecuado para aproximar la estructura de la serie. La siguiente función de autocorrelación de los residuales del modelo ajustado parece indicar que el orden p=8 es adecuado.

AUTOCORRELACIONES

```
                -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                  +----+----+----+----+----+----+----+----+----+----+----+
                                       I
  1    -.01                      +     I     +
  2     .02                      +     IX    +
  3     .03                      +     IX    +
  4     .01                      +     I     +
  5    -.03                      +    XI     +
  6    -.01                      +     I     +
  7    -.02                      +     I     +
  8     .03                      +     IX    +
  9     .00                      +     I     +
 10    -.02                      +    XI     +
 11    -.03                      +    XI     +
 12     .02                      +     I     +
 13    -.06                      +    XI     +
 14     .02                      +     IX    +
 15     .11                      +     IXXX  +
 16     .11                      +     IXXX  +
 17    -.07                      +   XXI     +
 18     .02                      +     I     +
 19     .00                      +     I     +
 20     .00                      +     I     +
```

En la tabla anterior los parámetros estimados A_{20}, A_{70}, A_{80}, A_{125} y A_{135} corresponden a las estimaciones de los efectos de las observaciones atípicas aditivas sobre el nivel de la serie. Las estimaciones preliminares se obtienen restando de los valores predichos por el modelo para cada período donde falta la observación el valor de su repectivo afecto Aj. Los resultados son:

Período	Valor Real	Predicción	Aj	Estimación valor faltante
20	38.845	147.962	112.9557	35.298
70	87.237	198.255	115.9351	85.204
80	94.387	199.466	107.7666	92.806
125	130.372	199.325	69.7186	130.923
135	137.432	199.420	64.1568	135.670

iii) A continuación, asignar los valores estimados preliminares a los períodos donde falta la información y emplee la ACF y la PACF para identificar el proceso ARIMA. Los resultados son los siguientes:

AUTOCORRELACIONES

```
                -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                  +----+----+----+----+----+----+----+----+----+----+----+
                                          I
  1   -.64          XXXXXXXXXXXX+XXXI   +
  2    .31                     +    IXXXX+XXX
  3   -.03                    +    XI     +
  4   -.03                    +    XI     +
  5   -.01                    +     I     +
  6    .06                    +     IX    +
  7   -.07                    +   XXI     +
  8   -.03                    +    XI     +
  9    .11                    +     IXXX  +
 10   -.11                    +  XXXI     +
 11    .06                    +     IXX   +
 12    .01                    +     I     +
 13   -.07                    +   XXI     +
 14    .00                    +     I     +
 15    .11                    +     IXXX  +
 16   -.11                    +  XXXI     +
 17    .07                    +     IXX   +
 18   -.03                    +    XI     +
 19   -.01                    +     I     +
 20    .05                    +     IX    +
```

AUTOCORRELACIONES PARCIALES

```
          -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
           +----+----+----+----+----+----+----+----+----+----+----+
                                       I
 1   -.64          XXXXXXXXXXXX+XXXI   +
 2   -.18                     X+XXXI   +
 3    .16                      +   IXXXX
 4    .13                      +   IXXX+
 5   -.05                      +  XI   +
 6   -.01                      +   I   +
 7   -.02                      +   I   +
 8   -.14                      XXXXI   +
 9    .03                      +   IX  +
10    .06                      +   IX  +
11    .01                      +   I   +
12    .02                      +   IX  +
13   -.06                      + XXI   +
14   -.16                      XXXXI   +
15    .07                      +   IXX +
16    .13                      +   IXXX+
17    .07                      +   IXX +
18   -.04                      +  XI   +
19   -.07                      + XXI   +
20    .03                      +   IX  +
```

Del análisis de estos resultados se observa que la ACF y la PACF contienen información consistente con un proceso ARIMA(0,1,2) (Observe el corte sobre la ACF en el rezago 2).

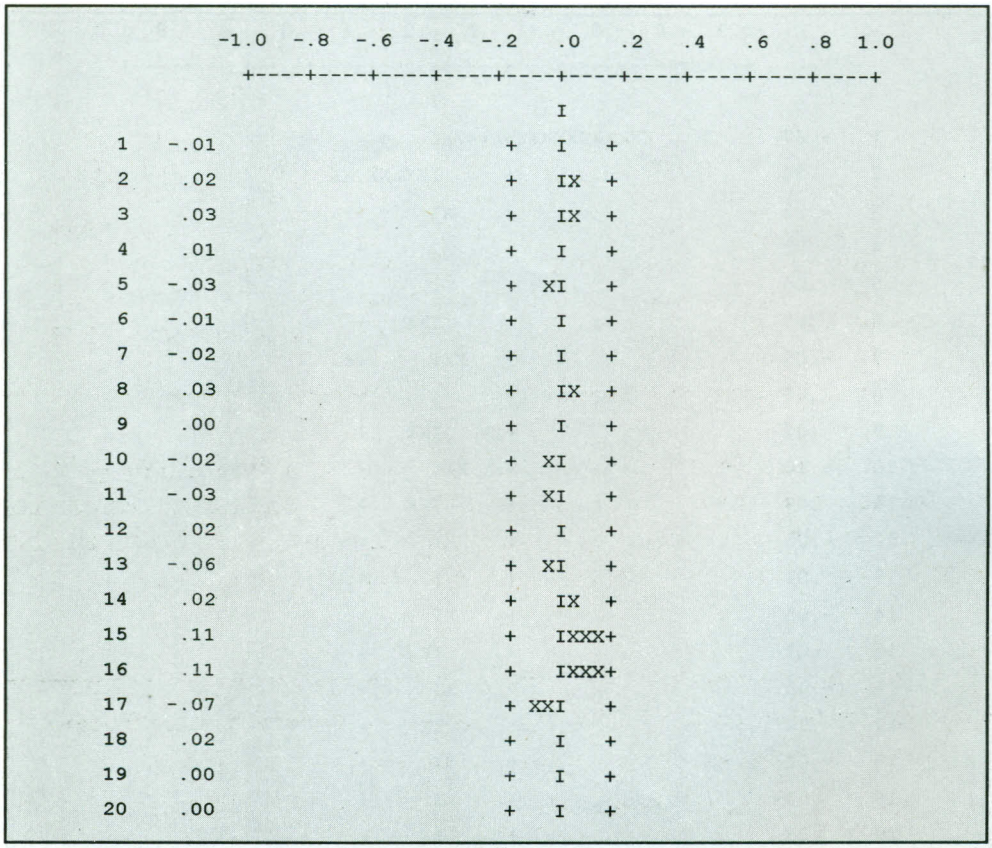
¿Qué tan fiel es el patrón de estas ACF y PACF con respecto a patrón de los datos originales?

Puesto que conocemos los verdaderos datos es interesante observar cuales son la ACF y la PACF de ellos. A continuación se presentan.

AUTOCORRELACIONES

```
                -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                 +----+----+----+----+----+----+----+----+----+----+----+
                                          I
  1    -.67          XXXXXXXXXXXXX+XXXI   +
  2     .34                     +     IXXXXX+XX
  3    -.03                     +    XI     +
  4    -.05                     +    XI     +
  5     .00                     +     I     +
  6     .07                     +     IXX   +
  7    -.09                     +   XXI     +
  8     .00                     +     I     +
  9     .07                     +     IXX   +
 10    -.10                     +   XXI     +
 11     .07                     +     IXX   +
 12    -.05                     +    XI     +
 13     .01                     +     I     +
 14    -.05                     +    XI     +
 15     .09                     +     IXX   +
 16    -.05                     +    XI     +
 17     .04                     +     IX    +
 18    -.02                     +    XI     +
 19     .00                     +     I     +
 20     .03                     +     IX    +
```

AUTOCORRELACIONES PARCIALES



		-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0
		+----+----+----+----+----+----+----+----+----+----+----+
		I
1	-.01	+ I +
2	.02	+ IX +
3	.03	+ IX +
4	.01	+ I +
5	-.03	+ XI +
6	-.01	+ I +
7	-.02	+ I +
8	.03	+ IX +
9	.00	+ I +
10	-.02	+ XI +
11	-.03	+ XI +
12	.02	+ I +
13	-.06	+ XI +
14	.02	+ IX +
15	.11	+ IXXX+
16	.11	+ IXXX+
17	-.07	+ XXI +
18	.02	+ I +
19	.00	+ I +
20	.00	+ I +

Obsérvese el comportamiento similar de las ACF y PACF de la serie con estimadores preliminares a las ACF y PACF de la serie de los datos originales.

iv) Para obtener una estimación más refinada de los datos faltantes, ajustar el modelo identificado usando las mismas intervenciones definidas en ii). Los resultados son los siguientes:

PARAMETER	LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS-TRAINT	VALUE	STD ERROR	T VALUE
1	CONST		CNST	1	0	NONE	.9065	.1259	7.20
2	A20	INT20	NUM.	1	0	NONE	113.7030	1.4923	76.19
3	A70	INT70	NUM.	1	0	NONE	114.8195	1.5010	76.50
4	A80	INT80	NUM.	1	0	NONE	106.4921	1.4918	71.38
5	A125	INT125	NUM.	1	0	NONE	69.0729	1.4836	46.56
6	A135	INT135	NUM.	1	0	NONE	63.5450	1.4951	42.50
7	T1	ZSSIN	MA	1	1	NONE	.7445	.0727	10.25
8	T2	ZSSIN	MA	1	2	NONE	-.5127	.0724	-7.08

En este modelo las A_j contienen estimaciones más refinadas de los efectos de intervención que permiten obtener también estimaciones más refinadas de las observaciones faltantes. A continuación se presentan las estimaciones finales de las observaciones faltantes:

Período	Valor Real	Predicción	Aj	Estimación valor faltante
20	38.845	148.955	113.7030	35.252
70	87.237	198.484	114.8195	83.664
80	94.387	199.531	106.4921	93.037
125	130.372	199.541	69.0729	130.470
135	137.432	199.518	63.5450	135.974

El error cuadrático medio de las estimaciones preliminares con respecto a las observaciones reales es de 4.524, mientras que el de las nuevas estimaciones es de 3.868. Por lo tanto, las nuevas estimaciones son más precisas que las preliminares.

Remplazando finalmente estos valores estimados finales de las observaciones faltantes podemos estimar el modelo identificado. Los resultados son:

PARAMETER LABEL	VARIABLE NAME	NUM/ DENOM.	FACTOR	ORDER	CONS-TRAINT	VALUE	SID ERROR	T VALUE
1 CONST	CNST	1	0	NONE	.9062	.1231	7.36	
2 T1	ZSSIN2	MA	1	1	NONE	.7419	.0720	10.30
3 T2	ZSSIN2	MA	1	2	NONE	-.4913	.0719	-6.83

Obsérvese lo cerca de las estimaciones obtenidas con respecto a los verdaderos parámetros 1.0, 0.70 y -0.45.

Ajustando el modelo a los **datos originales** las estimaciones son las siguientes:

PARAMETER LABEL	VARIABLE NAME	NUM/ DENOM.	FACTOR	ORDER	CONS-TRAINT	VALUE	SID ERROR	T VALUE
1 CONST		CNST	1	0	NONE	.9064	.1275	7.11
2 T1	ZSSIN2	MA	1	1	NONE	.7399	.0720	10.28
3 T2	ZSSIN2	MA	1	2	NONE	-.5078	.0718	-7.08

Comparando las dos tablas anteriores, los resultados del procedimiento se ajustan bastante a lo que se hubiera obtenido si hubiésemos conocidos los datos.

En la siguiente sección se presentan los resultados más importantes obtenidos en la simulación del modelo anterior bajo tres elecciones del orden autorregresivo p.

B. Experimentos Montecarlo

Para confirmar los hallazgos obtenidos para el modelo anterior se realizaron 1000 simulaciones para cada uno de los órdenes autorregresivos p=5, 8 y 15. Un resumen de los resultados se presentan en la siguiente tabla. Las cantidades que aparecen en ella son:

Const: es el promedio de las constantes (const) en las mil simulaciones

θ**1** y θ**2** son los promedios del primer y segundo parámetro de medias móviles del modelo, para las 1000 simulaciones.

Std(x) es el promedio de error estándar para el parámetro x en las 1000 simulaciones.

RESULTADOS PARA 1000 SIMULACIONES DE UN IMA(1,2)

orden p	Const	Std (const)	Theta1 (theta1)	Std	Theta2	std (theta2)
4	0.99448	0.12110	0.69015	0.08227	-0.43971	0.09110
8	0.99465	0.12141	0.69714	0.08241	-0.44594	0.09103
15	1.00320	0.12038	0.69965	0.08077	-0.44493	0.08307

La tabla anterior indica que para este modelo IMA(1, 2) los resultados no parecen estar lejos de los parámetros teóricos para cualquiera de las aproximaciones autorregresivas usadas. Sin embargo, se observa una mejoría leve a medida que se eleva el orden del proceso autorregresivo.

Observaciones:

- El procedimiento puede ser empleado usando software común, pues hace uso de modelos autorregresivos.

- El procedimiento descrito es un caso particular del procedimiento de identificación de un modelo ARIMA contaminado (Castaño, 1995).

Conclusiones

Los métodos propuestos para estimar observaciones faltantes suponen que el modelo se encuentra identificado. En la práctica, la disposición de las observaciones faltantes puede ser tal que impida la identificación del modelo. El procedimiento propuesto permite simultáneamente identificar y estimar las observaciones. Los resultados de la simulación muestran que el procedimiento parece funcionar bien. La aproximación del orden autorregresivo puede basarse en el resultado de Said y Dickey (1984) para el caso de modelos no estacionales.

Referencias

Box, G.E.P. y Jenkins, G.M. (1976) *Time Series Analysis, Forecasting and Control*, Holden-Day, 2ª edición.

Box, G.E.P. y Tiao, G.C. (1975) "*Intervention Analysis with Applications to Economic and Enviromental Problems*", Journal of the American Statistical Associaton, 70, 335-365.

Castañeda, M. (1994) "*Reconstrucción de Series de Tiempo Univariadas Mediante el Enfoque de Pronósticos con Restricciones*", Tesis de Grado, Magister en Estadística, Universidad Nacional de Colombia, Bogotá.

Castaño, Elkin (1995) "*Identificación de un Modelo ARIMA Contaminado*", Lecturas de Economía, Vol. 42, 49-70. Medellín.

Chen, C., y Liu, L-M (1990) "*Joint Estimation of Models Parameters and Outliers Effcets in Time Series*", Working Papers Series, Scientific Computing Associates, P.O.Box 625, DeKalb, Illinois 60115.

Chow, G.C. y Lin, A. (1976) "*Best Linear Unbiased Estimation of Missing Observation in a Economic Time Series*", Journal of the American Statistical Association, 71, 719-721.

Harvey, A.C. y Pierse, R.G. (1984) "*Estimating Missing Observations in Economic Time Series*", Journal of the American Statistical Associaton, 79, 385, 15-131.

Jones, R.H. (1980) "*Maximum Likelihood Fitting of ARMA Models to Time Series with Missing Observations*", Technometrics, 22, 389,395.

Kohn, R. y Ansley, F. (1983) "*Fixed Interval Estimation in State Space Models when Some of the Data are Missing or Aggregated*", Biometrika, 70, 683-688.

Maravall, A. y Peña, D. (1988), "*Missing Observations of Time Series and the Dual Autocorrelation Function*", Documento de trabajo No. 8830, Servicio de Estudios, Banco de España, Madrid.

Nieto, F. (1989) "*Reconstrucción de una serie de Tiempo Censurada usando Filtros de Kalman*", Tesis de Grado, Postgrado de Estadística, Universidad Nacional de Colombia, Bogotá.

Peña, D. y Maraval, A. (1990) "*Interpolation, Outliers and the Inverse Autocorrelations*", Aprobado para ser publicado en Communications in Statistics.

Said, S., y Dickey, D. (1984) "*Testing Unit Roots in Autoregresive-Moving Average Models with Unknown Order*", Biometrika 71, 599-607.

Wei, W.W.S. (1990) *Time Series Analysis : Univariate and Multivariate Methods.* Redwood City CA: Addison-Wesley.