UNIVERSIDAD DE CANTABRIA Facultad de Ciencias Económicas y Empresariales Departamento de Economía



TESIS DOCTORAL

Modelos econométricos dinámicos no invertibles

Doctorando Cristina Mazas Pérez-Oleaga

Director Prof. Dr. José Luis Gallego Gómez

Santander, 2016

Dedicada a mis padres.

Agradecimientos

En primer lugar quiero que este apartado sea de agradecimiento a mi director de tesis, José Luis. Hace tiempo que perdí a mis padres y algunos aspectos José Luis se ha convertido en algo parecido. Tengo que agradecerle el valor del trabajo, y sobre todo del trabajo riguroso. Sin su ayuda hubiese sido imposible acabar esta tesis. Gracias José Luis por tu infinita paciencia y comprensión.

A Carlos, con quién he pasado algunos de los momentos más significativos de esta dura experiencia. Ha sido un compañero generoso y es un buen amigo.

A todos aquellos los que han colaborado para que esta tesis llegue a término. A Sofía y a Juan, que me abrieron las puestas de la Universidad de Cantabria. A todos aquellos con quienes he compartido despacho. A Marijose, Ingrid y Marta con quienes he compartido muy buenos momentos.

Al Departamento de Economía de la Universidad de Cantabria.

También a mis compañeros de trabajo que entendieron lo importante que para mí era terminar esta tesis y han sabido suplir y disculpar algunas ausencias.

Y por supuesto también a mi familia, y en especial a Mario, Mencía, Víctor y Purita.

Muchísimas gracias a todos.

Índice general

1.	. Introducción						
2.	El n	l modelo clásico generalizado: contrastes óptimos					
	2.1.	Introducción	5				
	2.2.	El modelo clásico generalizado	6				
	2.3.	Contrastes óptimos sobre la matriz de covarianzas	10				
		2.3.1. Principio de invarianza	11				
		2.3.2. Distribución condicional	19				
		2.3.3. Verosimilitud concentrada	19				
	2.4.	Distribución muestral de una ratio de formas cuadráticas	20				
		2.4.1. Momentos	21				
		2.4.2. Tabulación	23				
3.	3. Contrastes óptimos de no invertibilidad regular						
	3.1. Introducción		27				
	3.2. El contraste de Nyblom y Mäkeläinen						
		3.2.1. Formulación matricial del modelo de nivel local	30				
		3.2.2. Estadístico de contraste	31				
		3.2.3. Distribución exacta	32				
		3.2.4. Distribución límite	35				

		3.2.5.	El contraste invariante óptimo puntual	37
		3.2.6.	Valores críticos y potencia del contraste	39
		3.2.7.	Inclusión de una deriva	41
	3.3.	Los co	ontrastes de Tanaka y Saikkonen-Luukkonen	46
		3.3.1.	La aproximación de Tanaka	46
		3.3.2.	La aproximación de Saikkonen-Luukkonen	50
		3.3.3.	La equivalencia entre los tres estadísticos	51
		3.3.4.	Distribución exacta	54
		3.3.5.	Distribución límite	56
		3.3.6.	Inclusión de una deriva	58
	3.4.	El con	traste de Nyblom y Mäkeläinen corregido	58
		3.4.1.	La corrección no paramétrica de KPSS	59
		3.4.2.	La corrección no paramétrica de Tanaka	60
		3.4.3.	La corrección paramétrica de Leybourne y McCabe	63
		3.4.4.	La corrección paramétrica de Saikkonen y Luukkonen	65
		3.4.5.	La corrección paramétrica de Tam y Reinsel	68
3.5. Contraste de no invertibilidad en un proceso ARIMA				71
		3.5.1.	El modelo de nivel local generalizado	73
		3.5.2.	Estadísticos de contraste	75
		3.5.3.	Distribución exacta	76
		3.5.4.	Estadísticos alternativos	78
		3.5.5.	Múltiples raíces MA unitarias	80
		3.5.6.	Evaluación de los contrastes	80
4.	Con	trastes	óptimos de no invertibilidad estacional	89
	4.1.	Introd	- ucción	89

	4.2. Contrastes de no invertibilidad en un modelo $MA(1)_s$					
		4.2.1.	El contraste de Tam y Reinsel	93		
		4.2.2.	La aproximación de Saikkonen-Luukkonen	95		
		4.2.3.	La aproximación de Nyblom y Mäkeläinen	97		
		4.2.4.	Valores críticos y potencia del contraste	100		
	4.3. Contraste de raíz MA unitaria estacional en un proceso ARIMA g					
		4.3.1.	El contraste para el nivel local estacional generalizado	103		
		4.3.2.	Estadísticos de contraste	105		
		4.3.3.	Distribución exacta	106		
		4.3.4.	Estadísticos aproximados	107		
	4.4. Constraste de no invertibilidad para un MA(2)					
		4.4.1.	la extensión del contraste de Tam y Reinsel	109		
		4.4.2.	La aproximación de Saikkonen y Luukkonen	110		
		4.4.3.	Contraste para de estabilidad paramétrica para una frecuencia .	113		
		4.4.4.	El contraste para las frecuencias estacionales generalizado	114		
		4.4.5.	Estadísticos de contraste	116		
		4.4.6.	Distribución exacta	118		
		4.4.7.	Evaluación de los contrastes	119		
	4.5.	Contra	aste para múltiples raíces unitarias en un modelo general	122		
5.	Con	clusion	nes	125		
A.	Algo	oritmos	s para el estadístico de Tam y Reinsel (1997)	129		
B.	Desa	arrollos	s del modelo MA(2) en la aproximación de Saikkonen y Luukko	-		
	nen			132		
		B.0.1.	Estadístico de contraste	132		
		B.0.2.	Distribución del estadístico	133		

Capítulo 1

Introducción

Una decisión clave en el análisis de series temporales siguiendo la metodología de Box y Jenkins (véase, por ejemplo, Box et al. 2015) es la de identificar el orden de integración o número de diferencias necesarias para transformar una serie no estacionaria en estacionaria. La inspección visual de determinados gráficos, como el gráfico temporal o el correlograma, es tremendamente útil para tomar esta decisión, pero a veces se critica por ser subjetiva o no tener asociado un determinado nivel de significación. Esto motivó el desarrollo de contrastes más formales, los contrastes de raíces unitarias, siguiendo dos aproximaciones diferentes: infra o sobrediferenciación de datos, siendo claramente la primera la más popular.

El contraste de referencia en la primera aproximación es el de Dickey y Fuller (1979). Estos autores consideraron un proceso simple AR(1)

$$Y_t = \phi Y_{t-1} + u_t,$$

y formularon las hipótesis

$$H_0: \phi = 1;$$

 $H_1: \phi < 1;$

donde la hipótesis nula de no estacionariedad se contrasta frente a la alternativa de estacionariedad. El interés que ha generado este problema de contraste se debe no tanto a su relevancia práctica como al desafío teórico que plantea, dado que no se aplica la teoría convencional de inferencia estadística. Dickey y Fuller derivaron las propiedades estadísticas del estimador de mínimos cuadrados de ϕ bajo H_0 y derivaron la distribución límite de la ratio-*t*, proporcionando así un método para realizar el contraste. El contraste de Dickey y Fuller se ha extendido en tres principales direcciones: correlación serial en los errores, las correcciones paramétrica de Said y Dickey (1985) y no paramétrica de Phillips y Perron (1988); raíces estacionales, Dickey et al. (1984) y Hylleberg et al. (1990) y cambio estructural, Perron (1990). Estas son solo unas referencias básicas dentro de una extensísima producción bibliográfica, que puede consultarse en diversos textos especializados como Patterson (2011) y Choi (2015).

La segunda aproximación para identificar el orden de integración se basa en la sobrediferenciación de datos. El modelo de referencia es el IMA(1,1)

$$(1-B)Y_t = (1-\theta B)a_t$$

donde se formulan las hipótesis

$$H_0: \theta = 1;$$

 $H_1: \theta < 1.$

Bajo H_0 , el modelo contiene un factor común en los dos lados de la ecuación revelando que la transformación diferencia no es necesaria, es decir, el proceso está sobrediferenciado. Se dice también que el modelo no es invertible. Aquí, el problema a contrastar es el dual al considerado por Dickey-Fuller: la hipótesis nula de estacionariedad se contrasta frente a la alternativa de no estacionariedad. Este contraste estadístico también ha atraído el interés de muchos académicos por ser irregular y no poder tratarse con la teoría convencional. Además, se presenta una dificultad adicional porque en muestras pequeñas existe una probabilidad significativa de que el estimador de máxima verosimilitud del parámetro verdadero sea exactamente uno cuando H_1 es verdadera. Este efecto amontonamiento, encontrado en algunos estudios de simulación como el de Ansley y Newbold (1983), fue probado por Cryer y Ledolter (1981) y Sargan y Bhargava (1983). La consecuencia de nuevo es que los contrastes clásicos de Wald o de razón de verosimilitudes no deberían usarse en este contexto, lo que motivó el desarrollo de una gran diversidad de procedimientos para detectar no invertibilidad. De entre todos estos contrastes, en esta tesis se toma como punto de partida el de Nyblom y Mäkeläinen (1983) por sus propiedades estadísticas: localmente invariante, óptimo e insesgado (LBIU). Este contraste guarda cierto paralelismo con el de Dickey-Fuller: es un contraste de raíz (MA) unitaria básico, que ha sido extendido en tres principales direcciones: correlación serial en los errores, las correcciones paramétricas de Saikkonen y Luukkonen (1993) y no paramétricas de Tanaka (1990), Kwiatkowski et al. (1992); raíces estacionales, Tam y Reinsel (1997), Canova y Hansen (1995), Harvey y Busetti (2003) y Taylor (2003); cambio estructural, Busetti y Harvey (2001) y Busetti y Taylor (2003). Entre los textos especializados cabe destacar Tanaka (Tanaka (1996)).

Una vez que se han definido las dos aproximaciones para el contraste de raíces unitarias es conveniente señalar que el interés de esta tésis se centra en el desarrollo de contrastes de no invertibilidad para la hipótesis nula de una o varias raíces unitarias en el contexto de un modelo ARIMA generalizado.

El punto de partida es el marco teórico común que comparten los contraste de no invertibilidad y que se desarrolla en el capítulo 2. En este capítulo se revisan los contrastes óptimos sobre la matriz de covarianzas en un modelo lineal general, donde uno de los problemas de contraste habituales es decidir si la matriz de covarianzas es escalar. Utilizando la teoría sobre invarianza en Lehman (1986) y los contrastes óptimos referidos en Ferguson (1967), King (1980) y King y Hillier (1985) se enuncian varios teoremas que permiten la obtención de contrastes óptimos para distintas hipótesis alternativas. Los contrastes que se obtienen se pueden expresar como una ratio de formas cuadráticas de variables normales que se pueden evaluar mediante procedimientos numéricos (Imhof 1961; Davies 1973) que también se recogen en la última sección de este mismo capítulo. Una vez establecido el marco general para el desarrollo de contrastes óptimos, en el Capítulo 3 se pasa al marco concreto de los contrastes de no invertibilidad regular en un modelo ARIMA. En este capítulo se realiza una revisión de los contrastes más relevantes para la hipótesis nula, $H_0: \theta = 1$ frente a $H_0: \theta \neq 1$ sobre la base de los teoremas básicos para contrastes óptimos del capítulo 2. Los trabajos más citados para este contraste son los anteriormente mencionados, Nyblom y Mäkeläinen (1983), Saikkonen y Luukkonen (1993) o Tanaka (1990) que se desarrollaron de forma independiente, como si no fuesen equivalentes. Con objeto evidenciar sus similitudes, en este capítulo se proponen nuevas relaciones matriciales útiles que además facilitan la obtención de sus distribuciones límite. También en este capítulo se abordan las correcciones para recoger correlación en el término de error, de carácter paramétrico y no paramétrico. En esta sección se propone un estadístico basado en los residuos exactos, y que constituye una de las principales aportaciones de esta tesis. Este estadístico es LBIU y combina las representaciones estructural y ARIMA en un modelo generalizado. Esto permite que, a diferencia de otros contastes de no invertibilidad, puede aplicarse en presencia de una raíz MA extra. Su funcionamiento se evalúa a través de un estudio Monte Carlo que ofrece unos resultados muy explicativos de cómo se comporta en presencia de un raíz unitaria adicional, y se compara con otros estadísticos existentes como Tanaka (1990), Saikkonen y Luukkonen (1993), Leybourne y McCabe (1994) y Tam y Reinsel (1997).

A continuación, en el Capítulo 4 se abordan las extensiones al marco estacional de los estadísticos de no invertibilidad regular. Uno de los artículos referentes es el de Tam y Reinsel (1997) que propone un estadístico para contrastar una raíz unitaria estacional en un modelo MA(1)_s. Sin embargo, este contraste ha recibido críticas puesto que no se puede aplicar de forma individualizada sobre los componentes que se obtienen de la factorización del polinomio $(1 - \theta B^s)$. Por otra parte, en los modelos estructurales la literatura al respecto es prolija; algunas aportaciones destacadas son las de Canova y Hansen (1995), Caner (1998), Harvey y Busetti (2003) o Taylor (2003) en las que se solventa la limitación que tiene su forma reducida en el contraste de Tam y Reinsel (1997), y se proponen contrastes en las frecuencias estacionales. Esta cuestión se debe a que la aditividad de los componentes en los modelos estructurales hace que el desarrollo de los contrastes sea más sencillo que en su forma reducida, aunque ambas representaciones son equivalentes. Por esto, en el Capítulo 4 se aportan nuevos estadísticos de raíces unitarias en el marco de los modelos ARIMA y estructural para contrastar una raíz unitaria estacional, así como en los factores de un MA(1)_s, tomando como referencia Gallego y Treadway (1995). También se extiende el estadístico basado en los residuos exactos a los modelos estacionales, y por último se apunta un procedimiento general para contrastar hipótesis distintas.

Finalmente, en el apartado de las conclusiones se informa brevemente de los principales resultados de esta tesis y las futuras líneas de investigacion.

CAPÍTULO 2

El modelo clásico generalizado: contrastes óptimos sobre la matriz de covarianzas

2.1. Introducción

En el marco del modelo clásico generalizado se plantea con frecuencia el problema de decidir si la matriz de covarianzas de las perturbaciones se reduce a una matriz escalar. Este problema de decisión tiene interés por dos razones principales. En primer lugar, por su múltiples aplicaciones. Bastantes contrastes de hipótesis pueden parametrizarse de esta forma. Algunos ejemplos son los contrastes de autocorrelación, heterocedasticidad, estabilidad paramétrica, anomalías, errores compuestos y, los que son objeto de estudio en esta tesis, los contrastes de no invertibilidad en modelos ARIMA. Y, en segundo lugar, por tratarse de un contraste irregular. La hipótesis nula especifica que un determinado parámetro se encuentra en la frontera del espacio paramétrico, por lo que no resulta de aplicación la teoría clásica de contrastación de hipótesis.

Usando la teoría de invarianza descrita en Lehman (1959, cap. 6) y los contrastes localmente óptimos de Ferguson (1967), King y Hillier (1985) propusieron diferentes resultados para contrastar la hipótesis nula de matriz de covarianzas escalar frente a varias formas de dependencia. Otros trabajos relacionados son los de Kadiyala (1970), Kariya (1977), Kariya (1980), King (1980), King y Wu (1997) y Muller (1998). La metodología de contraste propuesta por King y Hillier (1985) ha sido ampliamente aplicada en el desarrollo de contrastes de estabilidad y de no invertibilidad, los cuales serán objeto de revisión en los capítulos siguientes. Sin embargo, y a pesar de su relativa relevancia en el análisis de series temporales no estacionarias y en el tratamiento de otros problemas econométricos, apenas se discute en los libros de texto de estas materias. Es por esto que se ha creído conveniente dedicar este capítulo a ofrecer una exposición detallada de la misma, probando todos los resultados en que se basa.

El capítulo se organiza del siguiente modo. En primer lugar, se revisa el modelo clásico generalizado con el doble propósito de establecer una notación general y definir el problema de decisión. Después, se revisa el trabajo de King y Hillier (1985) proporcionando una derivación detallada de sus dos estadísticos de contraste: el contraste invariante localmente óptimo (LBI, del inglés Locally Best Invariant Unbiased) y el contraste invariante e insesgado localmente óptimo (LBIU, del inglés Locally Best Invariant Unbiased). Ambos estadísticos se definen como una ratio de formas cuadráticas en variables normales cuya evaluación requiere el uso de procedimientos numéricos, que son objeto de revisión en la última sección.

2.2. El modelo clásico generalizado

El modelo lineal general con perturbaciones aleatorias no esféricas, también denominado modelo de regresión lineal generalizado o modelo clásico generalizado, se define como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(0, \sigma^2 \mathbf{\Omega}_e(\theta)),$$
 (2.1)

en donde **y** es un vector columna de *T* observaciones, **X** es una matriz no estocástica de orden $T \times k$ y rango k < T, β es un vector columna de k coeficientes de regresión y **e** es un vector columna de *T* perturbaciones aleatorias. La notación **e** $\sim N(0, \sigma^2 \Omega_e(\theta))$ indica que **e** sigue una distribución normal multivariante con vector de medias $E(\mathbf{e}) =$ **0** y matriz de varianzas y covarianzas $V(\mathbf{e}) = \sigma^2 \Omega_e(\theta)$, dependiente de dos parámetros θ y σ^2 . Se supone, además, que $\Omega_e(\theta)$ es una matriz definida positiva y diferenciable en el conjunto de definición de $\theta \in \Theta \subset \mathbb{R}$.

Bajo estos supuestos, es claro que $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{\Omega}_e(\theta))$. Por tanto, la función de densi-

dad conjunta de y viene dada por

$$p(\mathbf{y}|\boldsymbol{\beta},\sigma^{2},\theta) = (2\pi\sigma^{2})^{-T/2} |\boldsymbol{\Omega}_{e}(\theta)|^{-1/2} \exp\left(-\frac{1}{2\sigma^{2}}S(\boldsymbol{\beta}|\mathbf{y},\mathbf{X},\theta)\right), \qquad (2.2)$$

que permite expresar el logaritmo neperiano de la función de verosimilitud de una realización particular de **y** como

$$\ell(\boldsymbol{\beta}, \sigma^2 | \mathbf{y}, \mathbf{X}, \theta) = -\frac{T}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \ln|\boldsymbol{\Omega}_e(\theta)| - \frac{1}{2\sigma^2} S(\boldsymbol{\beta} | \mathbf{y}, \mathbf{X}, \theta),$$
(2.3)

en donde $S(\beta|\mathbf{y}, \mathbf{X}, \theta) = (\mathbf{y} - \mathbf{X}\beta)' \Omega_e(\theta)^{-1} (\mathbf{y} - \mathbf{X}\beta)$ y se supone que el parámetro θ es conocido.

Los estimadores de máxima verosimilitud (MV) de β y σ^2 se obtienen resolviendo el sistema de ecuaciones lineales formado por las condiciones de primer orden para maximizar (2.3)

$$\frac{\partial \ell(\boldsymbol{\beta}, \sigma^2 | \mathbf{y}, \mathbf{X}, \theta)}{\partial \boldsymbol{\beta}} = -\frac{1}{\sigma^2} \left(-\mathbf{X}' \mathbf{\Omega}_e(\theta)^{-1} \mathbf{y} + \mathbf{X}' \mathbf{\Omega}_e(\theta)^{-1} \mathbf{X} \right) = \mathbf{0} \quad \mathbf{y}$$
$$\frac{\partial \ell(\boldsymbol{\beta}, \sigma^2 | \mathbf{y}, \mathbf{X}, \theta)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma_e^4} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{\Omega}_e(\theta)^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{0},$$

cuya solución viene dada por

$$\tilde{\boldsymbol{\beta}}_{MV} = (\mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{X})^{-1} \mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{y} \quad \mathbf{y} \quad \tilde{\sigma}_{MV}^2 = -\frac{S(\tilde{\boldsymbol{\beta}}_{MV} | \mathbf{y}, \mathbf{X}, \theta)}{T}.$$

El estimador $\tilde{\beta}_{MV}$ es equivalente al estimador de mínimos cuadrados generalizados (MCG), $\tilde{\beta}_{MCG}$, que se obtiene minimizando la suma de cuadrados de los residuos en el modelo de regresión transformado

$$\mathbf{y}_* = \mathbf{X}_* \boldsymbol{\beta} + \mathbf{e}_*, \quad \mathbf{e}_* \sim N(0, \sigma^2 \mathbf{I}_T), \tag{2.4}$$

en donde $\mathbf{y}_* = \mathbf{\Omega}_e(\theta)^{-1/2} \mathbf{y}, \mathbf{X}_* = \mathbf{\Omega}_e(\theta)^{-1/2} \mathbf{X}, \mathbf{e}_* = \mathbf{\Omega}_e(\theta)^{-1/2} \mathbf{e} \mathbf{y} \mathbf{\Omega}_e(\theta)^{1/2}$ es el factor de Cholesky de $\mathbf{\Omega}_e(\theta) = \mathbf{\Omega}_e(\theta)^{1/2} \mathbf{\Omega}_e(\theta)^{1/2}$. Así, definiendo el residuo $\mathbf{\tilde{e}}_*(\mathbf{\tilde{\beta}}) = \mathbf{y}_* - \mathbf{X}_*\mathbf{\tilde{\beta}}$, la función SCR $\mathbf{\tilde{e}}_*(\mathbf{\tilde{\beta}})'\mathbf{\tilde{e}}_*(\mathbf{\tilde{\beta}}) = (\mathbf{y}_* - \mathbf{X}_*\mathbf{\tilde{\beta}})'(\mathbf{y}_* - \mathbf{X}_*\mathbf{\tilde{\beta}}) = (\mathbf{y} - \mathbf{X}\mathbf{\beta})'\mathbf{\Omega}_e(\theta)^{-1}(\mathbf{y} - \mathbf{y})$ $\mathbf{X}oldsymbol{eta})=S(oldsymbol{ ildeeta}|\mathbf{y},\mathbf{X}, heta)$ alcanza un mínimo en

$$\tilde{\boldsymbol{\beta}}_{MCG} = (\mathbf{X}'_*\mathbf{X}_*)^{-1}\mathbf{X}_*\mathbf{y}_* = (\mathbf{X}'\boldsymbol{\Omega}_e(\theta)^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}_e(\theta)^{-1}\mathbf{y} = \tilde{\boldsymbol{\beta}}_{MV}.$$

Por el teorema de Gauss-Markov, $\tilde{\beta}_{MCG}$ (en adelante $\tilde{\beta}$) es el estimador más eficiente en la clase de estimadores lineales e insesgados del parámetro β . Además, por el supuesto de normalidad, su distribución muestral viene dada por

$$\tilde{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{X})^{-1}),$$

de manera que su función de densidad conjunta es

$$p(\tilde{\boldsymbol{\beta}}) = (2\pi\sigma^2)^{-k/2} |\mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{X}|^{1/2} \exp[-\frac{1}{2\sigma^2} (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' (\mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{X})^{-1} (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})].$$

Respecto a la estimación de σ^2 , el estimador MCG $\tilde{\sigma}^2_{MCG}$ difiere de $\tilde{\sigma}^2_{MV}$ en muestras finitas porque la SCR se corrige por los grados de libertad del modelo para conseguir insesgadez

$$\tilde{\sigma}_{MCG}^2 = \frac{(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{\Omega}_e(\theta)^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})}{m},$$

en donde m = T - k. Conviene notar que al reemplazar β por $\tilde{\beta}$ en el modelo de interés (2.1) y en el transformado (2.4) se obtienen dos tipos de residuos, $\tilde{\mathbf{e}} = \mathbf{y} - \mathbf{X}\tilde{\beta}$ y $\tilde{\mathbf{e}}_* = \mathbf{y}_* - \mathbf{X}_*\tilde{\beta}$, respectivamente, cuya relación viene dada por

$$\tilde{\mathbf{e}}_* = \mathbf{y}_* - \mathbf{X}_* \tilde{\boldsymbol{\beta}} = \boldsymbol{\Omega}_e(\theta)^{-1/2} (\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}) = \boldsymbol{\Omega}_e(\theta)^{-1/2} \tilde{\mathbf{e}}.$$
(2.5)

Cada uno de estos residuos puede contemplarse como una transformación lineal del vector de observaciones **y**

$$\mathbf{\tilde{e}}_* = \mathbf{M}_*(\theta)\mathbf{y}_* \quad \mathbf{y} \quad \mathbf{\tilde{e}} = \mathbf{M}(\theta)\mathbf{y},$$

en donde

$$\mathbf{M}_*(\theta) = \mathbf{I}_T - \mathbf{X}_*(\mathbf{X}'_*\mathbf{X}_*)^{-1}\mathbf{X}'_* = \mathbf{I}_T - \mathbf{\Omega}_e(\theta)^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{\Omega}_e(\theta)^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}_e(\theta)^{-1/2},$$

у

$$\mathbf{M}(\theta) = \mathbf{I}_T - \mathbf{X} (\mathbf{X}' \mathbf{\Omega}_e(\theta)^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{\Omega}_e(\theta)^{-1}$$

Mientras que $\mathbf{M}_*(\theta)$ es una matriz simétrica, $\mathbf{M}(\theta)$ no lo es. Pero ambas son idempotentes, anulan a la matriz \mathbf{X} , $\mathbf{M}_*(\theta)\mathbf{X} = \mathbf{M}(\theta)\mathbf{X}$, y tienen rango *m*.

Dependiendo del tipo de residuo usado, la función SCR puede expresarse bien como

$$S(\tilde{\boldsymbol{\beta}}|\mathbf{y},\mathbf{X},\theta) = \tilde{\mathbf{e}}'_{*}\tilde{\mathbf{e}}_{*} = \mathbf{y}'_{*}\mathbf{M}_{*}(\theta)\mathbf{y}_{*} = \mathbf{e}'_{*}\mathbf{M}_{*}(\theta)\mathbf{e}_{*}$$

o bien como

$$S(\tilde{\boldsymbol{\beta}}|\mathbf{y},\mathbf{X},\theta) = \tilde{\mathbf{e}}'\boldsymbol{\Omega}_{e}(\theta)^{-1}\tilde{\mathbf{e}} = \mathbf{y}'\mathbf{M}(\theta)'\boldsymbol{\Omega}_{e}(\theta)^{-1}\mathbf{M}(\theta)\mathbf{y} = \mathbf{e}'\mathbf{M}(\theta)'\boldsymbol{\Omega}_{e}(\theta)^{-1}\mathbf{M}(\theta)\mathbf{e}.$$

Y aplicando los resultados sobre distribuciones de formas cuadrativas, se tiene que $S(\tilde{\boldsymbol{\beta}}|\mathbf{y}, \mathbf{X}, \theta) / \sigma^2 \sim \chi_m^2$, en donde χ_m^2 es la distribución Chi-cuadrado con *m* grados de libertad. De aquí, la función de densidad de $\tilde{\sigma}_{MCG}^2$ (en adelante $\tilde{\sigma}^2$) es

$$p(\tilde{\sigma}^2) = \frac{m^{m/2} (\tilde{\sigma}^2)^{m/2-1} \exp(-m\tilde{\sigma}^2/2\sigma^2)}{\Gamma(m/2)(2\sigma^2)^{m/2}}.$$

El supuesto de normalidad junto con el resultado $\mathbf{M}(\theta)\mathbf{X} = \mathbf{0}$ garantizan la independencia estadística de $\tilde{\boldsymbol{\beta}}$ y $\tilde{\mathbf{e}}$, por lo que la función de densidad conjunta

$$p(\tilde{\boldsymbol{\beta}}, \tilde{\sigma}^2) = \frac{|\mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{X}|^{1/2} m^{m/2} (\tilde{\sigma}^2)^{m/2-1}}{\pi^{k/2} \Gamma(m/2) (2\sigma^2)^{T/2}} \exp\left(-\frac{1}{2\sigma^2} S(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X}, \theta)\right), \quad (2.6)$$

dado que

$$m\tilde{\sigma}^2 + (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\mathbf{X}'\boldsymbol{\Omega}_e(\theta)^{-1}\mathbf{X})^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = S(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X}, \theta).$$

Todos los resultados anteriores pueden particularizarse trivialmente para el caso especial más relevante de (2.1) dado por el modelo de regresión lineal clásico, que aparece cuando se supone que la matriz de varianzas y covarianzas de **e** es escalar, $\Omega_e(\theta) = \mathbf{I}_T$. En este marco, el mejor estimador en el sentido de Gasuss-Markov es el de mínimos cuadrados ordinarios

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y},$$

que da lugar al vector de residuos ordinarios

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{M}\mathbf{y},$$

en donde $\mathbf{M} = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ es una matriz idempotente y simétrica con descomposición espectral $\mathbf{M} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{M}^2 = \mathbf{P}\mathbf{\Lambda}^2\mathbf{P}'$, siendo $\mathbf{\Lambda}$ una matriz diagonal con autovalores unitarios o nulos y \mathbf{P} la correspondondiente matriz de autovectores. Además, como la traza de \mathbf{M} es T - k, $\mathbf{\Lambda}$ contiene T - k unos y k zeros en la diagonal principal. Asimismo, \mathbf{P} admite la forma particionada $\mathbf{P} = [\mathbf{P}_1 \ \mathbf{P}_2]$, siendo \mathbf{P}_1 la submatriz de orden $T \times (T - k)$ que contiene los autovectores asociados a los autovalores unitarios y \mathbf{P}_2 la submatriz de orden $k \times (T - k)$ que contiene los autovectores asociados a los autovalores nulos. Usando esta partición de \mathbf{P} y la partición compatible de $\mathbf{\Lambda}$,

$$\mathbf{\Lambda} = egin{pmatrix} \mathbf{I}_{T-k} & \mathbf{0} \ \mathbf{0} & \mathbf{0} \end{pmatrix}$$
 ,

la descomposición espectral de M se reduce a

$$\mathbf{M} = \mathbf{P}_1 \mathbf{P}_{1\prime}^{\prime} \tag{2.7}$$

cumpliéndose que $\mathbf{P}'_1\mathbf{P}_1 = \mathbf{I}_{T-k}$ y $\mathbf{P}'_1\mathbf{X} = \mathbf{0}$. Esta descomposición juega un papel fundamental en la derivación de los contrastes invariantes que se estudian en la siguiente sección.

2.3. Contrastes óptimos sobre la matriz de covarianzas

En el modelo (2.1), el interés se centra frecuentemente en decidir si el parámetro θ es igual a un determinado valor θ_0 . Este problema de decisión puede formularse bien como un contraste de hipótesis nula simple frente a alternativa unilateral

$$H_0: \theta = \theta_0$$
 versus $H_1: \theta > \theta_0$ (o $H_1: \theta < \theta_0$), (2.8)

o bien como un contraste de hipótesis nula simple frente a alternativa bilateral

$$H_0: \theta = \theta_0$$
 versus $H_1: \theta \neq \theta_0$, (2.9)

dependiendo de si $\Omega_e(\theta)$ está definida sólo para $\theta \ge \theta_0$ ($\theta \le \theta_0$) o para cualquier $\theta \in \mathbb{R}$. En ambos casos, tanto H_0 como H_1 son invariantes a cambios de origen y escala de los datos porque la distribución multivariante de la transformación lineal $g(\mathbf{y}) =$ $a\mathbf{y} + \mathbf{X}\mathbf{b}, g(\mathbf{y}) \sim N(\mathbf{X}(a\boldsymbol{\beta} + \mathbf{b}), a^2\sigma^2\Omega(\theta))$, sigue dependiendo de $\Omega_e(\theta)$. De Lehman y Romano (1986 p. 214), el principio de invarianza sugiere centrar la atención a la clase de contrastes invariantes, definida por todos los estadísticos $T(\mathbf{y})$ que pueden expresarse en términos de un invariante maximal de $\mathbf{y}, \mathbf{v}(\mathbf{y})$.

2.3.1. Principio de invarianza

En general, se dice que una función de las observaciones, $\nu(\mathbf{y})$, es un invariante maximal con respecto a una transformación *g* cuando cumple las propiedades de invarianza y maximalidad. El estadístico $\nu(\mathbf{y})$ es invariante a *g* si se cumple que $\nu(g(\mathbf{y})) = \nu(\mathbf{y})$ para cualquier realización de **y**; y un maximal, si para dos realizaciones distintas de **y**, \mathbf{y}_1 e \mathbf{y}_2 , $\nu(\mathbf{y}_1) = \nu(\mathbf{y}_2)$ implica que $\mathbf{y}_1 = g(\mathbf{y}_2)$.

En el caso particular del modelo clásico generalizado, el estadístico

$$\boldsymbol{\nu}(\mathbf{y}) = \mathbf{P}_1' \mathbf{y} / \sqrt{\mathbf{y}' \mathbf{P}_1 \mathbf{P}_1' \mathbf{y}}$$
(2.10)

es un invariante maximal con respecto a la transformación lineal $g(\mathbf{y}) = a\mathbf{y} + \mathbf{X}\mathbf{b}$, en donde \mathbf{P}_1 se ha definido en (2.7). La demostración de que $v(g(\mathbf{y})) = v(\mathbf{y})$ se prueba fácilmente notando que $\mathbf{P}'_1\mathbf{X} = \mathbf{0}$ y $\mathbf{P}'_1g(\mathbf{y}) = a\mathbf{P}'_1\mathbf{y}$. De aquí, es claro que la premuntiplicación de \mathbf{y} por \mathbf{P}'_1 anula el cambio de origen, mientras que la división de este producto por $(\mathbf{y}\mathbf{P}'_1\mathbf{P}'_1\mathbf{y})^{1/2}$ cancela el cambio de escala. De ahora en adelante, para simplificar la notación, se omite el argumento de $v(\mathbf{y})$ y se considera el vector $v = (v_1 \ v_2 \ \dots \ v_m)'$ con m = T - k.

La función de densidad conjunta de v puede obtenerse a partir de la distribución de $\mathbf{z} = \mathbf{P}_1 \mathbf{y} \sim N(\mathbf{0}, \sigma^2 \mathbf{P}'_1 \mathbf{\Omega}_e(\theta) \mathbf{P}_1)$. Definiendo $r = \mathbf{z}' \mathbf{z}$, la trasformación de $\mathbf{z} = (z_1 \ z_2 \ \dots \ z_m)'$ a un sistema de coordenadas polares viene dada por

$$z_j = r\sin(w_1)\dots\sin(\omega_{j-1})\cos(\omega_j) \quad (j = 1,\dots,m)$$

en donde r > 0, $0 \le \omega_i \le \pi$ para $i = 1, \dots, m-2$, $0 \le \omega_{m-1} \le 2\pi$ y $\omega_m = 0$. De aquí,

la función de densidad conjunta de r y Ω_e es

$$f(r, \mathbf{\Omega}_e) = f(\mathbf{z}) ||\mathcal{J}||,$$

en donde $\Omega_e = (\omega_1 \ \omega_2 \ \dots \ \omega_{m-1})' \ y \ ||\mathcal{J}||$ es el valor absoluto del determinante de la matriz Jacobiana de la transformación de \mathbf{z} a $(r \ \Omega'_e)$

$$|\mathcal{J}| = \begin{vmatrix} \frac{\partial z_1}{\partial r} & \frac{\partial z_1}{\partial \omega_1} & \cdots & \frac{\partial z_1}{\partial \omega_{m-1}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial z_m}{\partial r} & \frac{\partial z_m}{\partial \omega_1} & \cdots & \frac{\partial z_m}{\partial \omega_{m-1}} \end{vmatrix} = r^{m-1} \prod_{j=1}^{m-1} \sin(\omega_j)^{m-1-j}.$$

Puesto que $\mathbf{z} = r \boldsymbol{v}$, resulta que

$$f(r, \mathbf{\Omega}_e) \propto |\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1|^{-1/2} \exp\left(-\frac{1}{2\sigma^2} r^2 \nu' (\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \nu\right) r^{m-1} \prod_{j=1}^{m-1} \sin(\omega_j)^{m-1-j},$$

siendo $(2\pi\sigma^2)^{-(m)/2}$ la constante de proporcionalidad. Notando que

$$f(\mathbf{\Omega}_e) = f(\mathbf{\nu}) \prod_{j=1}^{m-1} \sin(\omega_j)^{m-1-j},$$

la expresión de f(v) se obtiene resolviendo la integral

$$f(\mathbf{\Omega}_e) = \int_0^\infty f(r, \mathbf{\Omega}_e) dr.$$

Para ello se realiza el cambio de variable

$$\alpha = \frac{1}{2\sigma^2} r^2 (\boldsymbol{\nu}' (\mathbf{P}'_1 \boldsymbol{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \boldsymbol{\nu}),$$

que conduce a

$$r = (2\sigma^2)^{1/2} (\nu' (\mathbf{P}'_1 \mathbf{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \nu)^{-1/2} \alpha^{1/2}$$

у

$$dr = \frac{1}{2} (2\sigma^2)^{1/2} (\nu' (\mathbf{P}'_1 \mathbf{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \nu)^{-1/2} \alpha^{-1/2} d\alpha.$$

Teniendo en cuenta que

$$\int_0^\infty e^{-\alpha} \alpha^{(m)/2-1} = \Gamma\left(\frac{m}{2}\right),$$

se obtiene finalmente que

$$f(\boldsymbol{\nu}) = \frac{1}{2} \Gamma\left(\frac{m}{2}\right) \pi^{-(m)/2} |\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1|^{-1/2} (\boldsymbol{\nu}' (\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \boldsymbol{\nu})^{-(m)/2}.$$
 (2.11)

Dentro de la clase de contrastes invariantes, el mejor contraste de tamaño α para el problema de decisión (2.8) o (2.9), en el sentido de ser uniformemente más potente (contraste UMPI), se obtiene aplicando el lema de Neyman y Pearson y viene dado por

$$\mathcal{T}(\boldsymbol{\nu}) = \begin{cases} 1 & \text{si} \quad f(\boldsymbol{\nu}|\boldsymbol{\theta}) > \kappa f(\boldsymbol{\nu}|\boldsymbol{\theta}_0), \\ 0 & \text{si} \quad f(\boldsymbol{\nu}|\boldsymbol{\theta}) < \kappa f(\boldsymbol{\nu}|\boldsymbol{\theta}_0), \end{cases}$$
(2.12)

que, para una realización de ν dada, sugiere rechazar H_0 cuando $\mathcal{T}(\nu) = 1$ o aceptar H_0 cuando $\mathcal{T}(\nu) = 0$. Si $\mathcal{V} \subset \mathbb{R}^m$ denota el espacio muestral de ν , entonces la función potencia de $\mathcal{T}(\nu)$, o probabilidad de rechazar H_0 cuando $\theta \in \Theta$ es el valor paramétrico verdadero, se define por

$$\beta_{\mathcal{T}}(\theta) = E_{\theta}[\mathcal{T}(v)] = \int_{\mathcal{V}} \mathcal{T}(v) f(v|\theta) dv,$$

que proporciona la probabilidad del error de tipo I cuando $\theta = \theta_0$ y uno menos la probabilidad del error de tipo II en otro caso. Así, la constante κ se elige de forma que $\beta(\theta_0) = \alpha$. De (2.11) y (2.12), y usando las relaciones $\mathbf{P}'_1 \mathbf{P}_1 = \mathbf{I}_m$ y $\nu'\nu = 1$, se tiene que la región crítica del contraste viene dada por

$$\frac{f(\boldsymbol{\nu}|\boldsymbol{\theta})}{f(\boldsymbol{\nu}|\boldsymbol{\theta}_0)} = |\mathbf{P}_1'\mathbf{\Omega}_e(\boldsymbol{\theta})\mathbf{P}_1|^{-1/2} (\boldsymbol{\nu}'(\mathbf{P}_1'\mathbf{\Omega}_e(\boldsymbol{\theta})\mathbf{P}_1)^{-1}\boldsymbol{\nu})^{-(m)/2} > \kappa_e$$

que es equivalente a

$$\boldsymbol{\nu}'(\mathbf{P}_1'\boldsymbol{\Omega}_e(\theta)\mathbf{P}_1)^{-1}\boldsymbol{\nu} = \frac{\mathbf{y}'\mathbf{P}_1(\mathbf{P}_1'\boldsymbol{\Omega}_e(\theta)\mathbf{P}_1)^{-1}\mathbf{P}_1'\mathbf{y}}{\mathbf{y}'\mathbf{M}\mathbf{y}} < \kappa_1.$$
(2.13)

Notando que $\mathbf{M}_*(\theta) = \mathbf{\Omega}_e(\theta)^{1/2} \mathbf{P}_1(\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \mathbf{P}_1' \mathbf{\Omega}_e(\theta)^{1/2}$ es una matriz simétrica

e idempotente que anula a la matriz $X_* = \Omega_e(\theta)^{-1/2}X$, el numerador de (2.13) puede escribirse como

$$\mathbf{y}' \mathbf{\Omega}_{e}(\theta)^{-1/2} \mathbf{\Omega}_{e}(\theta)^{1/2} \mathbf{P}_{1}(\mathbf{P}_{1}' \mathbf{\Omega}_{e}(\theta) \mathbf{P}_{1})^{-1} \mathbf{P}_{1}' \mathbf{\Omega}_{e}(\theta)^{-1/2} \mathbf{\Omega}_{e}(\theta)^{1/2} \mathbf{y} = \mathbf{y}_{*}' \mathbf{M}_{*}(\theta)' \mathbf{y}_{*}.$$

Teorema 1. Para el problema de decisión (2.8) o (2.9), el contraste UMPI rechaza H_0 cuando el estadístico

$$\frac{\tilde{\mathbf{e}}' \mathbf{\Omega}_e(\theta)^{-1} \tilde{\mathbf{e}}}{\hat{\mathbf{e}}' \hat{\mathbf{e}}} < \kappa, \tag{2.14}$$

en donde **ẽ** y **ê** son los residuos de la estimación MCG y MCO, respectivamente.

Cuando se contrasta H_0 : $\theta = \theta_0$ frente a una alternativa específica H_0 : $\theta = \theta_1$, el contraste basado en (2.14) se denomina óptimo puntual. En general, para alternativas unilateres o bilaterales, como la región crítica depende del valor de θ fijado bajo H_1 , no existe un contraste UMPI para todas las alternativas. En esta situación, se consideran contrastes invariantes que sean localmente óptimos en un entorno de H_0 .

Siguiendo a Ferguson (1967, p. 224-225), dentro de la clase de contrastes invariantes, el contraste localmente más potente (Locally Best Invariant, en inglés) de tamaño α para el problema de decisión (2.8) viene dado por

$$\mathcal{T}(\boldsymbol{\nu}) = \begin{cases} 1 & \text{si} \quad f'(\boldsymbol{\nu}|\theta_0) > \kappa f(\boldsymbol{\nu}|\theta_0), \\ 0 & \text{si} \quad f'(\boldsymbol{\nu}|\theta_0) < \kappa f(\boldsymbol{\nu}|\theta_0), \end{cases}$$
(2.15)

en donde $f'(\nu|\theta_0)$ es la función derivada de $f(\nu|\theta)$ con respecto de θ evaluada bajo H_0 . La prueba de esta proposición se basa en un razonamiento similar al usado en el lema de Neyman y Pearson.

Sea $\mathcal{T}_*(\nu)$ un contraste invariante alternativo con tamaño menor o igual que α , $\beta_{\mathcal{T}_*}(\theta_0) \leq \alpha$. Se cumple siempre que

$$[\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_*(\boldsymbol{\nu})][f'(\boldsymbol{\nu}|\theta_0) - \kappa f(\boldsymbol{\nu}|\theta_0)] \ge 0.$$

Lógicamente, cuando $\mathcal{T}(\boldsymbol{\nu})$ rechaza H_0 , $[f'(\boldsymbol{\nu}|\theta_0) - \kappa f(\boldsymbol{\nu}|\theta_0)] > 0$ y $[\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_*(\boldsymbol{\nu})] \ge 0$ dependiendo de si $\mathcal{T}_*(\boldsymbol{\nu})$ acepta ($\mathcal{T}_*(\boldsymbol{\nu}) = 0$) o rechaza ($\mathcal{T}_*(\boldsymbol{\nu}) = 1$) H_0 . Análogamente, cuando $\mathcal{T}(\boldsymbol{\nu})$ acepta H_0 , $[f'(\boldsymbol{\nu}|\theta_0) - \kappa f(\boldsymbol{\nu}|\theta_0)] < 0$ y $[\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_*(\boldsymbol{\nu})] \leq 0$ dependiendo de si $\mathcal{T}_*(\boldsymbol{\nu})$ rechaza ($\mathcal{T}_*(\boldsymbol{\nu}) = 1$) o acepta ($\mathcal{T}_*(\boldsymbol{\nu}) = 0$) H_0 . Por tanto,

$$\begin{split} 0 &\leq \int_{\mathcal{V}} [\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_{*}(\boldsymbol{\nu})] [f'(\boldsymbol{\nu}|\theta_{0}) - \kappa f(\boldsymbol{\nu}|\theta_{0})] d\boldsymbol{\nu}, \\ 0 &\leq \int_{\mathcal{V}} [\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_{*}(\boldsymbol{\nu})] f'(\boldsymbol{\nu}|\theta_{0}) \boldsymbol{\nu} - \kappa \int_{\mathcal{V}} [\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_{*}(\boldsymbol{\nu})] f(\boldsymbol{\nu}|\theta_{0})] d\boldsymbol{\nu}, \\ 0 &\leq \left(\beta_{\mathcal{T}}'(\theta_{0}) - \beta_{\mathcal{T}_{*}}'(\theta_{0})\right) - \kappa \left(\beta_{\mathcal{T}}(\theta_{0}) - \beta_{\mathcal{T}_{*}}(\theta_{0})\right), \end{split}$$

en donde $\beta'_{\mathcal{T}}(\theta_0)$ y $\beta'_{\mathcal{T}_*}(\theta_0)$ son las pendientes de las funciones potencias de $T(\nu)$ y $T_*(\nu)$ en θ_0 . Como $\beta_{\mathcal{T}}(\theta_0) = \alpha$ y $\beta_{\mathcal{T}_*}(\theta_0) \leq \alpha$, resulta que $\beta'_{\mathcal{T}}(\theta_0) \geq \beta'_{\mathcal{T}_*}(\theta_0)$, es decir, $\mathcal{T}(\nu)$ es el contraste invariante con función potencia más inclinada en θ_0 . Usando ahora la aproximación lineal de $\beta_{\mathcal{T}}(\theta)$ en un entorno θ_0

$$\beta_{\mathcal{T}}(\theta) \simeq \beta_{\mathcal{T}}(\theta_0) + (\theta - \theta_0)\beta'_{\mathcal{T}}(\theta_0)$$

se observa que $\beta_{\mathcal{T}}(\theta) > \beta_{\mathcal{T}_*}(\theta)$ porque $\beta'_{\mathcal{T}}(\theta_0) \ge \beta'_{\mathcal{T}_*}(\theta_0)$.

Notando que la región crítica del contraste (2.15) puede expresarse en términos de la derivada del logaritmo de $f(\nu|\theta)$ respecto de θ

$$\left.\frac{d\ln f(\boldsymbol{\nu}|\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} > \kappa,$$

se tiene para (2.11) que

$$-\frac{1}{2} \left. \frac{d\ln |\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1|}{d\theta} \right|_{\theta=\theta_0} - \frac{m}{2} \left. \frac{d\ln (\nu' (\mathbf{P}_1' \mathbf{\Omega}_e(\theta) \mathbf{P}_1)^{-1} \nu)}{d\theta} \right|_{\theta=\theta_0} > \kappa.$$

Aplicando las fórmulas de la derivada de un determinante

$$\frac{d \left| \mathbf{P}_{1}^{\prime} \mathbf{\Omega}_{e}(\theta) \mathbf{P}_{1} \right|}{d\theta} = tr\left((\mathbf{P}_{1}^{\prime} \mathbf{\Omega}_{e}(\theta) \mathbf{P}_{1})^{-1} \frac{d(\mathbf{P}_{1}^{\prime} \mathbf{\Omega}_{e}(\theta) \mathbf{P}_{1})}{d\theta} \right)$$

y de la inversa de una matriz

$$\frac{d(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)^{-1}}{d\theta} = -(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)^{-1}\frac{d(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)}{d\theta}(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)^{-1},$$

se obtiene que

$$\begin{split} \frac{d\ln f(\boldsymbol{\nu}|\boldsymbol{\theta})}{d\boldsymbol{\theta}} &= -\frac{1}{2\left|\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1}\right|} tr\left((\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1})^{-1}\frac{d(\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1})}{d\boldsymbol{\theta}}\right) \\ &+ \frac{m}{2}\frac{\boldsymbol{\nu}^{\prime}(\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1})^{-1}\frac{d(\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1})}{d\boldsymbol{\theta}}(\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1})^{-1}\boldsymbol{\nu}}{\boldsymbol{\nu}^{\prime}(\mathbf{P}_{1}^{\prime}\boldsymbol{\Omega}_{e}(\boldsymbol{\theta})\mathbf{P}_{1})^{-1}\boldsymbol{\nu}}, \end{split}$$

cuya evaluación en $\theta = \theta_0$ conduce a la región crítica

$$\frac{d\ln f(\boldsymbol{\nu}|\theta)}{d\theta}\Big|_{\theta=\theta_0} = -\frac{1}{2}tr\left(\left.\frac{d\boldsymbol{\Omega}_e(\theta)}{d\theta}\right|_{\theta=\theta_0}\right) + \frac{m}{2}\boldsymbol{\nu}'\mathbf{P}_1'\left(\left.\frac{d\boldsymbol{\Omega}_e(\theta)}{d\theta}\right|_{\theta=\theta_0}\right)\mathbf{P}_1\boldsymbol{\nu} > \kappa,$$

que es equivalente a

$$\frac{\mathbf{y}'\mathbf{P}_1(\mathbf{P}_1'\mathbf{\Omega}_e(\theta_0)\mathbf{P}_1)^{-1}\mathbf{P}_1'\left(\left.\frac{d\mathbf{\Omega}_e(\theta)}{d\theta}\right|_{\theta=\theta_0}\right)\mathbf{P}_1(\mathbf{P}_1'\mathbf{\Omega}_e(\theta_0)\mathbf{P}_1)^{-1}\mathbf{P}_1'\mathbf{y}}{\mathbf{y}'\mathbf{P}_1(\mathbf{P}_1'\mathbf{\Omega}_e(\theta_0)\mathbf{P}_1)^{-1}\mathbf{P}_1'\mathbf{y}} > \kappa_1.$$

Teorema 2. Para el problema de decisión (2.8), el contraste LBI rechaza H₀ cuando

$$\frac{\tilde{\mathbf{e}}' \mathbf{\Omega}_{e}(\theta_{0})^{-1} \left(\left. \frac{d \mathbf{\Omega}_{e}(\theta)}{d \theta} \right|_{\theta=\theta_{0}} \right) \mathbf{\Omega}_{e}(\theta_{0})^{-1} \tilde{\mathbf{e}}}{\tilde{\mathbf{e}}' \mathbf{\Omega}_{e}(\theta_{0})^{-1} \tilde{\mathbf{e}}} > \kappa,$$
(2.16)

en donde $\tilde{\mathbf{e}} = \mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_{MCG} = \mathbf{M}(\theta_0)\mathbf{y}$ son los residuos en la estimación MCG. En el caso especial $\boldsymbol{\Omega}_e(\theta_0) = \mathbf{I}_T$, la región crítica se reduce a

$$\frac{\hat{\mathbf{e}}'\left(\left.\frac{d\mathbf{\Omega}_{e}(\theta)}{d\theta}\right|_{\theta=\theta_{0}}\right)\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}}>\kappa,$$

en donde $\hat{\mathbf{e}} = \mathbf{M}\mathbf{y}$ son los residuos de la estimación MCO del modelo.

Del Teorema 2 se deriva el siguiente corolorario: cuando la media de es un vector nulo, X = 0, el contraste LBI rechaza H_0 cuando

$$\frac{\mathbf{y}' \mathbf{\Omega}_{e}(\theta_{0})^{-1} \left(\left. \frac{d \mathbf{\Omega}_{e}(\theta)}{d \theta} \right|_{\theta=\theta_{0}} \right) \mathbf{\Omega}_{e}(\theta_{0})^{-1} \mathbf{y}}{\mathbf{y}' \mathbf{\Omega}_{e}(\theta_{0})^{-1} \mathbf{y}} > \kappa.$$
(2.17)

Análogamente, dentro de la clase de contrastes invariantes insesgados, el contraste lo-

calmente óptimo de tamaño α para el problema de decisión (2.9) viene dado por

$$\mathcal{T}(\boldsymbol{\nu}) = \begin{cases} 1 & \text{si} \quad f''(\boldsymbol{\nu}|\theta_0) > \kappa_1 f(\boldsymbol{\nu}|\theta_0) + \kappa_2 f'(\boldsymbol{\nu}|\theta_0), \\ 0 & \text{si} \quad f''(\boldsymbol{\nu}|\theta_0) < \kappa_1 f(\boldsymbol{\nu}|\theta_0) + \kappa_2 f'(\boldsymbol{\nu}|\theta_0), \end{cases}$$
(2.18)

en donde $f''(\nu|\theta_0)$ es la segunda derivada de $f(\nu|\theta)$ con respecto de θ evaluada bajo H_0 .

Para probar la proposición, en primer lugar, hay que tener en cuenta que el contraste $\mathcal{T}(\boldsymbol{\nu})$ es insesgado cuando su función potencia es tal que $\beta_{\mathcal{T}}(\theta) \ge \beta_{\mathcal{T}}(\theta_0)$. Si la hipótesis alternativa es bilateral, entonces la función potencia tendrá un mínimo en θ_0 y se cumplirá $\beta'_{\mathcal{T}}(\theta_0) = 0$ y $\beta''_{\mathcal{T}}(\theta_0) > 0$. En segundo lugar, hay que notar que para cualquier otro contraste invariante insesgado $\mathcal{T}_*(\boldsymbol{\nu})$, se cumplirá que

$$[\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_*(\boldsymbol{\nu})][f''(\boldsymbol{\nu}|\theta_0) - \kappa_1 f(\boldsymbol{\nu}|\theta_0) - \kappa_2 f'(\boldsymbol{\nu}|\theta_0)] \ge 0$$

y, por tanto,

$$\int_{\mathcal{V}} [\mathcal{T}(\boldsymbol{\nu}) - \mathcal{T}_*(\boldsymbol{\nu})] [f''(\boldsymbol{\nu}|\theta_0) - \kappa_1 f(\boldsymbol{\nu}|\theta_0) - \kappa_2 f'(\boldsymbol{\nu}|\theta_0)] d\mathcal{V} \ge 0.$$

De aquí, se obtiene que

$$\beta_{\mathcal{T}}''(\theta_0) \ge \beta_{\mathcal{T}_*}''(\theta_0).$$

Usando ahora la aproximación cuadrática de $\beta_{\mathcal{T}}(\theta)$ en θ_0

$$eta_{\mathcal{T}}(heta) \simeq eta_{\mathcal{T}}(heta_0) + (heta - heta_0)eta_{\mathcal{T}}'(heta_0) + rac{1}{2}(heta - heta_0)^2eta_{\mathcal{T}}''(heta_0)$$

se observa que $\beta_{\mathcal{T}}(\theta) \ge \beta_{\mathcal{T}_*}(\theta)$ porque $\beta'_{\mathcal{T}}(\theta_0) = \beta'_{\mathcal{T}_*}(\theta_0) = 0$ y $\beta''_{\mathcal{T}}(\theta_0) \ge \beta''_{\mathcal{T}_*}(\theta_0)$.

La región crítica del contraste (2.18) puede expresarse como

$$\frac{d^2 \ln f(\nu|\theta)}{d\theta^2}\Big|_{\theta=\theta_0} + \left(\frac{d \ln f(\nu|\theta)}{\theta}\Big|_{\theta=\theta_0}\right)^2 > \kappa_1 + \kappa_2 \left.\frac{d \ln f(\nu|\theta)}{d\theta}\right|_{\theta=\theta_0}.$$

Dada la forma de $f(\nu|\theta)$ en (2.11), puede comprobarse que $\beta'_{\mathcal{T}}(\theta_0)$ será igual a cero cuando $f'(\nu|\theta) = 0$, es decir, cuando $f(\nu|\theta)$ alcance un máximo o un mínimo en θ_0 .

En este caso, $f'(\nu|\theta) = 0$ cuando $d\Omega_e(\theta)/d\theta|_{\theta=\theta_0}$ sea una matriz escalar. Por tanto, la región crítica del contraste (2.18) para (2.11) se reduce a

$$\frac{d^2 \ln f(\boldsymbol{\nu}|\boldsymbol{\theta})}{d\boldsymbol{\theta}^2}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} > \kappa,$$

que, calculando la segunda derivada del determinate y de la inversa, conduce a

$$\frac{\mathbf{y}'\mathbf{P}_1(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)^{-1}\mathbf{P}_1'\frac{d^2\mathbf{\Omega}_e(\theta)}{d\theta^2}\mathbf{P}_1(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)^{-1}\mathbf{P}_1'\mathbf{y}}{\mathbf{y}'\mathbf{P}_1(\mathbf{P}_1'\mathbf{\Omega}_e(\theta)\mathbf{P}_1)^{-1}\mathbf{P}_1'\mathbf{y}}\Bigg|_{\theta=\theta_0} > k.$$

Teorema 3. *Para el problema de decisión (2.9) con hipótesis alternativa bilateral, el contraste LBIU rechaza H*₀ *cuando*

$$\frac{\tilde{\mathbf{e}}' \mathbf{\Omega}_{e}(\theta_{0})^{-1} \left(\left. \frac{d^{2} \mathbf{\Omega}_{e}(\theta)}{d\theta^{2}} \right|_{\theta=\theta_{0}} \right) \mathbf{\Omega}_{e}(\theta_{0})^{-1} \tilde{\mathbf{e}}}{\tilde{\mathbf{e}}' \mathbf{\Omega}_{e}(\theta_{0})^{-1} \tilde{\mathbf{e}}} > \kappa,$$
(2.19)

en donde $\tilde{\mathbf{e}} = \mathbf{M}(\theta_0)\mathbf{y}$ son los residuos de la estimación MCG del modelo. En el caso especial $\mathbf{\Omega}_e(\theta_0) = \mathbf{I}_T$, la región crítica se reduce a

$$\frac{\hat{\mathbf{e}}'\left(\left.\frac{d^2\mathbf{\Omega}_e(\theta)}{d\theta^2}\right|_{\theta=\theta_0}\right)\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}} > \kappa,$$

en donde $\hat{\mathbf{e}} = \mathbf{M}\mathbf{y}$ son los residuos de la estimación MCO del modelo.

En resumen, cuando $\Omega_e(\theta_0) = \mathbf{I}_T$, los dos estadísticos LBI y LBIU derivados en esta sección pueden escribirse en una forma común como

$$L_T = \frac{\mathbf{\hat{e}}' \mathbf{A}(\theta_0) \mathbf{\hat{e}}}{\mathbf{\hat{e}}' \mathbf{\hat{e}}},$$

en donde $\mathbf{A}(\theta) = d\mathbf{\Omega}_e(\theta)/d\theta$ para (2.16) y $\mathbf{A}(\theta) = d^2\mathbf{\Omega}_e(\theta)/d\theta^2$ para (2.19).

2.3.2. Distribución condicional

Honda (1989) siguió una aproximación alternativa basada en la distribución condicionada de y dados $\tilde{\beta}$ y $\tilde{\sigma}^2$, cuya función de densidad viene dada por

$$p(\mathbf{y}|\boldsymbol{\tilde{\beta}}, \tilde{\sigma}^2, \theta) = \frac{p(\mathbf{y}|\boldsymbol{\beta}, \sigma^2, \theta)}{p(\boldsymbol{\tilde{\beta}}, \tilde{\sigma}^2)}$$

De (2.2) y (2.6) se obtiene que

$$p(\mathbf{y}|\boldsymbol{\tilde{\beta}}, \tilde{\sigma}^2, \theta) = \frac{\Gamma(m/2)}{(m\pi)^{m/2} |\mathbf{X}' \boldsymbol{\Omega}_e(\theta)^{-1} \mathbf{X}|^{1/2} |\boldsymbol{\Omega}_e(\theta)|^{1/2} (\tilde{\sigma}^2)^{m/2-1}}.$$
(2.20)

Aplicando ahora el lema de Neyman y Pearson a la función de densidad condicional (2.20), se obtiene la región crítica

$$\frac{p(\mathbf{y}|\boldsymbol{\tilde{\beta}}, \tilde{\sigma}^2, \theta)}{p(\mathbf{y}|\boldsymbol{\hat{\beta}}, \hat{\sigma}^2, \theta_0)} > \kappa.$$

2.3.3. Verosimilitud concentrada

Se propone ahora una tercera aproximación a la derivación de los contrastes óptimos basada en la función de verosimilitud concentrada. Reemplazando β y σ^2 en (2.3) por sus estimadores de MV, se obtiene la función de verosimilitud concentrada

$$L(\theta|\mathbf{y}, \mathbf{X}, \tilde{\boldsymbol{\beta}}, \tilde{\sigma}^2) = \left(2\pi\tilde{\sigma}^2\right)^{-T/2} |\mathbf{\Omega}_e(\theta)|^{-1/2} \exp\left(-\frac{T}{2}\right).$$
(2.21)

Aplicando el lema de Neyman y Pearson a (2.21), se obtiene que el contraste uniformemente más potente tiene región crítica dada

$$\frac{L(\theta|\mathbf{y}, \mathbf{X}, \tilde{\boldsymbol{\beta}}, \tilde{\sigma}^2)}{L(\theta_0|\mathbf{y}, \mathbf{X}, \hat{\boldsymbol{\beta}}, \hat{\sigma}^2)} = \frac{S(\theta|\mathbf{y}, \mathbf{X}, \tilde{\boldsymbol{\beta}})}{S(\theta_0|\mathbf{y}, \mathbf{X}, \tilde{\boldsymbol{\beta}})} > \kappa.$$

Análogamente, para obtener el contraste LBI, es necesario derivar el logaritmo neperiano de (2.21)

$$\ell(\theta|\mathbf{y},\mathbf{X},\tilde{\boldsymbol{\beta}},\tilde{\sigma}^2) = -\frac{T}{2}\ln(\frac{2\pi}{T}+1) - \frac{1}{2}\ln|\boldsymbol{\Omega}_e(\theta)| - \frac{T}{2}\ln S(\theta|\mathbf{y},\mathbf{X},\tilde{\boldsymbol{\beta}}),$$

en donde $S(\theta|\mathbf{y}, \mathbf{X}, \tilde{\boldsymbol{\beta}}) = \mathbf{y}' \mathbf{M}(\theta)' \mathbf{\Omega}_{e}(\theta)^{-1} \mathbf{M}(\theta) \mathbf{y}$. Notando que

$$\mathbf{M}(\theta_0)'\mathbf{\Omega}_e(\theta_0)^{-1}\left(\left.\frac{d\mathbf{M}(\theta)}{d\theta}\right|_{\theta=\theta_0}\right) = 0 \quad \mathbf{y} \quad \left.\frac{d\mathbf{\Omega}_e(\theta)^{-1}}{d\theta}\right|_{\theta=\theta_0} = -\left.\frac{d\mathbf{\Omega}_e(\theta)}{d\theta}\right|_{\theta=\theta_0},$$

se demuestra que la región crítica

$$\left. \frac{d\,\ell(\theta|\mathbf{y},\mathbf{X},\tilde{\boldsymbol{\beta}},\tilde{\sigma}^2)}{d\theta} \right|_{\theta=\theta_0} > \kappa$$

es equivalente a (2.16).

2.4. Distribución muestral de una ratio de formas cuadráticas

El estadístico L_T puede expresarse en términos del vector de perturbaciones aleatorias **e** como

$$L_T = \frac{\mathbf{e}' \mathbf{M} \mathbf{A}(\theta_0) \mathbf{M} \mathbf{e}}{\mathbf{e}' \mathbf{M} \mathbf{e}},$$

en donde se observa que L_T es la ratio de dos formas cuadráticas en variables aleatorias normales. De (2.1) y (2.7), $\mathbf{P}'_1 \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{P}'_1 \mathbf{\Omega}_e(\rho) \mathbf{P}_1)$. Por tanto, definiendo $\boldsymbol{\epsilon} = (\mathbf{P}'_1 \mathbf{\Omega}_e(\rho) \mathbf{P}_1)^{-1/2} \mathbf{P}'_1 \mathbf{e} / \sigma \sim N(\mathbf{0}, \mathbf{I}_m)$, se tiene que

$$L_T = \frac{\boldsymbol{\epsilon}'(\mathbf{P}_1' \mathbf{\Omega}_e(\rho) \mathbf{P}_1)^{1/2} (\mathbf{P}_1' \mathbf{A}(\theta_0) \mathbf{P}_1) (\mathbf{P}_1' \mathbf{\Omega}_e(\rho) \mathbf{P}_1)^{-1/2} \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}'(\mathbf{P}_1' \mathbf{\Omega}_e(\rho) \mathbf{P}_1) \boldsymbol{\epsilon}}$$

es la ratio de dos formas cuadráticas en variables normales tipiticadas. Bajo H_0 : $\Omega_e(\rho_0) = \mathbf{I}_T$, $\mathbf{P}'_1 \Omega_e(\rho_0) \mathbf{P}_1$ y la ratio L_T ser reduce a

$$L_T = \frac{\boldsymbol{\epsilon}' \mathbf{P}_1' \mathbf{A}(\theta_0) \mathbf{P}_1 \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}' \boldsymbol{\epsilon}},$$

en donde $\mathbf{P}'_1 \mathbf{A}(\theta_0) \mathbf{P}_1$ una matriz cuadrada de orden *m* con descomposición espectral

$$\mathbf{P}_1'\mathbf{A}(\theta_0)\mathbf{P}_1 = \sum_{t=1}^m \lambda_t \mathbf{p}_t \mathbf{p}_t',$$

en donde λ_t el es *t*-ésimo autovalor de una matriz simétrica cuyo autovector asociado \mathbf{p}_t cumple las propiedades

$$\mathbf{p}_t'\mathbf{p}_t = 1$$
, $\mathbf{p}_t'\mathbf{p}_s = 0 \ (t \neq s)$ y $\sum_{t=1}^m \mathbf{p}_t\mathbf{p}_t' = \mathbf{I}_m$.

De aquí,

у

$$L_T = \frac{\boldsymbol{\epsilon}' \left(\sum_{t=1}^m \lambda_t \mathbf{p}_t \mathbf{p}'_t \right) \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}' \left(\sum_{t=1}^m \mathbf{p}_t \mathbf{p}'_t \right) \boldsymbol{\epsilon}},$$

en donde $\mathbf{p}'_t \boldsymbol{\epsilon}_t \sim \operatorname{iid} N(0, 1)$. Definiendo ahora $\xi_t = \boldsymbol{\epsilon}' \mathbf{p}_t \mathbf{p}'_t \boldsymbol{\epsilon} \sim \operatorname{iid} \xi_1^2$, se tiene que

$$L_T = \frac{\sum_{t=1}^m \lambda_t \xi_t}{\sum_{t=1}^m \xi_t}.$$
(2.22)

2.4.1. Momentos

Los momentos de L_T son difíciles de calcular directamente porque la distribución exacta de este tipo de ratios de formas cuadráticas en variables normales es desconocida. Coniffe y Spencer (2001) atribuyen a Geary (1933) un resultado que permite calcular los momentos de una ratio como la ratio de momentos. Esta propiedad, que ha pasado frecuentemente inadvertida, se presenta cuando la ratio es independiente de su denominador, de forma que

$$E\left[\left(\frac{\sum_{t=1}^{m}\lambda_{t}\epsilon_{t}^{2}}{\sum_{t=1}^{m}\epsilon_{t}^{2}}\right)\left(\sum_{t=1}^{m}\epsilon_{t}^{2}\right)\right]^{r} = E\left(\frac{\sum_{t=1}^{m}\lambda_{t}\epsilon_{t}^{2}}{\sum_{t=1}^{m}\epsilon_{t}^{2}}\right)^{r} E\left(\sum_{t=1}^{m}\epsilon_{t}^{2}\right)^{r}$$
$$E\left(\frac{\sum_{t=1}^{m}\lambda_{t}\epsilon_{t}^{2}}{\sum_{t=1}^{m}\epsilon_{t}^{2}}\right)^{r} = E\left(\sum_{t=1}^{m}\lambda_{t}\epsilon_{t}^{2}\right)^{r} / E\left(\sum_{t=1}^{m}\epsilon_{t}^{2}\right)^{r}$$

La independencia de la ratio de su propio denominador puede probarse transformando ϵ_t a coordenadas polares

$$\epsilon_t = r \sin(w_1) \dots \sin(\omega_{t-1}) \cos(\omega_t) \quad (t = 1, \dots, m),$$

en donde $r > 0, 0 \le \omega_t \le \pi$ para $t = 1, ..., m - 2, 0 \le \omega_{m-1} \le 2\pi$ y $\omega_m = 0$. Notando que el denominador $\sum_{t=1}^m \epsilon_t^2 = r^2$, es claro que ϵ_t^2/r^2 y, por tanto, la ratio dependen sólo de $\omega_1, \omega_2, ..., \omega_m$, pero no de *r*. Como la función de densidad conjunta de *r*, $\omega_1, ..., \omega_{m-1}$ es

$$f(r,\omega_1,\ldots,\omega_{m-1})=f(\epsilon_1,\ldots,\epsilon_m)r^{m-1}\prod_{j=1}^{m-1}\sin(\omega_j)^{m-1-j},$$

se tiene que

$$f(r^{2},\omega_{1},\ldots,\omega_{m-1})=\frac{1}{2}(2\pi)^{-(m)/2}\exp(-r^{2}/2)r^{(m)/2-1}\prod_{j=1}^{m-1}\sin(\omega_{j})^{m-1-j},$$

en donde se observa que

$$f(r^2,\omega_1,\ldots,\omega_{m-1})=f(r^2)f(\omega_1)\ldots f(\omega_{m-1})$$

siendo

$$\begin{aligned} f(\omega_{m-1}) &= 2\pi, \quad 0 \le \omega_{m-1} \le 2\pi, \\ f(\omega_{m-j}) &= c_j \sin^{m-j-1}(\omega_{m-j-1}), \quad 0 \le \omega_{m-j-1} \le \pi, \\ f(r^2) &= \frac{1}{2^{(m)/2} \Gamma((m)/2)} (r^2)^{(m)/2-1} e^{-(r^2)/2}, \quad 0 \le r^2 \end{aligned}$$

la función de densidad de una distribución chi-cuadrado con *m* grados de libertad.

El momento no centrado de orden *s*, *a*_s, de una forma cuadrática del tipo $\sum_{t=1}^{m} \lambda_t \epsilon_t$ puede calcularse en términos de los cumulantes

$$a_s = \kappa_s - \sum_{j=1}^{s-1} \binom{r-1}{j-1} \kappa_j a_{s-j},$$

en donde el cumulante de orden j está dado por

$$\kappa_j = 2^{j-1}(j-1)! \sum_{t=1}^m \lambda_t^j.$$

Usando estas fórmulas se comprueba fácilmente que los dos primeros momentos de la ratio L_t son

$$E(L_T) = \bar{\lambda}_1 \quad \text{y} \quad \text{var}(L_T) = \frac{2}{m+2} [\bar{\lambda}_2 - \bar{\lambda}_1^2],$$
 (2.23)

en donde $\lambda_1 = \sum_{t=1}^m \lambda_t / m$ y $\lambda_2 = \sum_{t=1}^m \lambda_t^2 / m$. Estos dos primeros momentos pueden usarse para aproximar la distribución desconocida de L_T por una distribución paramétrica.

2.4.2. Tabulación

La función de distribución del estadístico L_T puede expresarse como

$$F(\kappa) \equiv \Pr(L_T < \kappa) = \Pr\left[\frac{\sum_{t=1}^m \lambda_t \xi_t}{\sum_{t=1}^m \xi_t} < \kappa\right] = \Pr\left[\sum_{t=1}^m (\lambda_t - \kappa) \xi_t\right] < 0$$

cuya tabulación puede realizarse usando el procedimiento propuesto por Imhof (1961) para evaluar la función de distribución de una suma ponderada de variables χ^2 independientes

$$Q=\sum_{r=1}^m \omega_k \chi^2_{h_r;\delta_r},$$

en donde ω_k son los pesos (que pueden ser negativos) y $\chi^2_{h_r;\delta_r}$ es una variable χ^2 con h_r grados de libertad y parámetro de no centralidad δ_r . La evaluación de la función de distribución $F(\kappa) = \Pr(Q < \kappa)$ se basa en el teorema de inversión de Gil-Pelaez (1951)

$$F(\kappa) = \frac{1}{2} - \int_0^\infty \operatorname{Im}\left[\frac{\phi(t)e^{-it\kappa}}{2\pi t}\right] dt,$$

en donde $\phi(t)$ es la función característica de Q dada por

$$\phi(t) = \prod_{r=1}^{m} (1 - 2i\omega_r t)^{-h_r/2} \exp\left(i\sum_{r=1}^{m} \frac{\delta_r^2 \omega_r t}{1 - 2i\omega_r t}\right)$$

 $i = \sqrt{-1}$, Im(z) es la parte imaginaria del número completo z, en particular, Im $[\phi(t)e^{-it\kappa}] = [\phi(t)e^{-it\kappa} - \phi(-t)e^{it\kappa}]/i$. Usando las relaciones

$$\arg[(1-ibt]^{-g} = g \tan^{-1}(bt), \quad |(1-ibt)^{-g}| = (1+b^2t^2)^{-g/2},$$

 $\arg[\exp\{iat/(1-bit)\}] = at/(1+b^2t^2), \quad |\exp\{iat/(1-bit)\}| = \exp\{-abt^2(1+b^2t^2)\},$

se tiene que

$$F(\kappa) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\sin[\theta(t)]}{t\rho(t)} dt,$$

en donde¹

$$\theta(t) = \sum_{r=1}^{m} [(h_r/2) \tan^{-1}(2t\omega_r) + \delta_r^2 t\omega_r (1 + 4t^2\omega_r^2)^{-1}] - \kappa t \quad \mathbf{y}$$

$$\rho(t) = \prod_{r=1}^{m} (1 + 4t^2\omega_r^2)^{h_r/4} \exp\left\{\frac{1}{2}\sum_{r=1}^{m} (2\delta_r t\omega_r)^2 / (1 + 4t^2\omega_r^2)\right\}.$$

Como el denominador del integrando, $t\rho(t)$, es una función monótona creciente y el numerador está acotado en [-1, 1], la integración puede realizarse en el intervalo finito $[0, \tau]$, de manera que la aproximación

$$F(\kappa) \simeq \frac{1}{2} - \frac{1}{\pi} \int_0^\tau \frac{\sin[\theta(t)]}{t\rho(t)} dt,$$
(2.24)

tiene asociado un error de truncamiento

$$|e_{\tau}| = \frac{1}{\pi} \int_{\tau}^{\infty} \frac{\sin[\theta(t)]}{t\rho(t)} dt,$$

que tiene una cota superior E_{τ} viene dada por

$$E_{\tau} = (\pi k)^{-1} (2\tau)^{-k} \prod_{r=1}^{m} |\omega_{r}|^{-h_{r}/2} N(\tau),$$
$$N(\tau) = \exp\left\{-2\sum_{r=1}^{m} \left[\delta_{r}^{2} \omega_{r}^{2} \tau^{2} / (1 + 4\omega_{r}^{2} \tau^{2})\right]\right\},$$

en donde $k = (1/2) \sum_{r=1}^{m} h_r$. Si e_a es el máximo error admisible en la aproximación, entonces el punto de truncamiento τ se elige por tanteo hasta conseguir que $E_{\tau} < e_a$. La integral definida puede calcularse numericamente usando la regla del trapecio o el método de Simpson, que tienen asociados un error de integración e_i . Koerts y Abrahamse (1971) proporcionaron una derivación detallada del procedimiento de Imhof (1961) y un programa FORTRAM para su aplicación, en el que fijaron los dos errores e_a y e_i en 10^{-5} . Esta rutina fue traducida a Pascal por Farebrother (1990), cuya principal mejora consiste en ignorar los pesos $|\omega_r|$ pequeños en el cálculo de E_{τ} porque pueden dar lugar a un valor de τ muy grande.

¹Imhof (1961) realiza el cambio de variable u = 2t, que aquí no se aplica para facilitar la comparación con el procedimiento de Davies (1973) que se revisa más adelante.

Davies (1973) extendió la forma cuadrática *Q* incluyendo una variable aleatoria *Z* con distribución normal estándar e independiente del resto de variables

$$Q_1 = \sum_{r=1}^m \omega_k \chi^2_{h_r;\delta_r} + \sigma Z,$$

cuya función característica es

$$\phi_1(t) = \phi(t) \exp(-i\sigma^2 t^2/2) = \prod_{r=1}^m (1 - 2i\omega_r t)^{-h_r/2} \exp\left(i\sum_{r=1}^m \frac{\delta_r^2 \omega_r t}{1 - 2i\omega_r t} - \frac{i\sigma^2 t^2}{2}\right)$$

Davies (1973) siguió un método de Fourier similar a los descritos en Bohman (1972), para el desarrollo Usando las relaciones dadas por Imhof (1961) y siguiendo los mismos pasos, se obtiene que la función de distribución de Q_1 está dada por

$$F_1(\kappa) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\sin[\theta(t)]}{t \rho_1(t)} dt,$$

en donde $\rho_1(t) = \rho(t) \exp(t^2 \sigma^2/2)$. Evaluando la integral en un intevalo finito [0, τ], Davies (1980) propuso tres cotas para el error de truncamiento

$$\begin{split} E_{1,\tau} &= [2/(\pi s)] N(\tau) \exp(-\tau^2 \sigma^2/2) \prod_{(i)} (1 + 4\tau^2 \omega_j^2)^{-n_j/4} \prod_{(ii)} (4\tau^2 \omega_j^2)^{-n_j/4} \\ E_{2,\tau} &= [1/(\pi \tau^2 \sigma^2)] N(\tau) \exp(-\tau^2 \sigma^2/2) \prod_{j=1}^r (1 + 4\tau^2 \omega_j^2)^{-n_j/4} \quad y \\ E_{3,\tau} &= (2.5/\pi) N(\tau) \exp(-\tau^2 \sigma^2/2) \prod_{j=1}^r (1 + 4\tau^2 \omega_j^2)^{-n_j/4}, \end{split}$$

en donde los subíndices (*i*) y (*ii*) en los productos indican que $\omega_r < 1$ y $\omega_r > 1$, respectivamente, y $s = \sum_{(ii)} h_r$. El parámetro τ se elige mediante el método de tanteo, partiendo de un $\tau_0 = 4/\sigma_{Q'}$, en donde $\sigma_{Q'}^2$ es la varianza de Q' dada por

$$\sigma_{Q'}^2 = \sigma^2 + \sum_{r=1}^m \omega_r^2 (2h_r + 4\delta_j).$$

Si $E_{\tau} = \max\{E_{1,\tau}, E_{2,\tau}, E_{2,\tau}, \}$ y e_a es el máximo error admisible en la aproximación, entonces para un τ dado se tiene que $E_{\tau} > e_a/2$ o $E_{\tau} \le e_a/2$. En el primer caso, se aplica sucesivamente un factor un factor c > 1 a τ hasta conseguir que $e_{\tau} \le e_a/2$ (por ejemplo, c = 4); en el segundo, se aplica sucesivamente un factor un factor c < 1 a τ hasta conseguir que $e_{\tau} > e_a/2$ (por ejemplo, c = 1/4).

La evaluación de $F_1(\kappa)$ se realiza usando el método del trapecio

$$F_1(\kappa) \simeq rac{1}{2} - rac{\Delta}{\pi} \sum_{k=0}^K rac{\sin\{\theta[(k+1/2)\Delta]\}}{[(k+1/2)\Delta]
ho_1[(k+1/2)\Delta]}$$
 ,

en donde $\tau = (K + 1/2)\Delta$, siendo Δ el intervalo de integración que se elige de manera que

máx {
$$F_1(\kappa - 2\pi/\Delta), 1 - F_1(\kappa + 2\pi/\Delta)$$
} < $e_a/2,$

que puede acotarse del siguiente modo. Sean $\Psi(t) = E[e^{tQ_1}]$ la función generatriz de momentos de Q_1 y $\psi(t) = \ln[\Psi(t)]$, entonces

$$E[\mathcal{I}(Q_1 > \kappa) - e^{t(Q_1 - \kappa)}] = 1 - F_1(\kappa) - \Psi(t)e^{-t\kappa} \le 0$$

у

$$1 - F_1(\kappa) \le e^{-t\kappa + \psi(t)}.$$

Así, eligiendo $\kappa = \psi'(t)$, se tiene que

$$1 - F_1(\psi'(t)) \le e^{-t\psi'(t) + \psi(t)},$$

que permite encontrar un *t* tal que $1 - F_1(\psi'(t)) < e_a$.

Lu y King (2002) proponen fórmulas alternativas para calcular el punto de truncamiento τ y compararon su eficiencia con los arriba descritos. Concluyen que el procedimiento de Davies (1973) no es necesariamente más eficiente que el de Imhof (1961), y sugieren calcular todas las cotas y elegir la menor de todas ellas. Su estudio, sin embargo, no es exhaustivo y no consideran la mejora propuesta por Farebrother (1990), que en esta tesis se ha encontrado fundamental para muestras grandes. Este procedimiento y el Davies (1980) se han traducido aquí a C++ y están disponibles en el programa Empiricus.

Capítulo 3

Contrastes óptimos de no invertibilidad regular

3.1. Introducción

Tanaka (1990) y Saikkonen y Luukkonen (1993) derivaron el contraste LBUI para el problema de decidir si un modelo IMA(1,1) puro es estrictamente no-invertible o, dicho de otro modo, tiene una raíz MA unitaria. Además, extendieron el contraste de noinvertiblidad al modelo IMA(1,1) impuro o modelo IMA(1,1) con errores serialmente correlacionados, que puede contemplarse como un proceso ARIMA(p, 1, q + 1) general, en donde los órdenes (p, 1, q + 1) aseguran que el modelo contiene una estructura IMA(1,1). Mientras que Tanaka (1990) propuso modificar el estadístico de contraste LBIU aplicando un factor de corrección no paramétrico, Saikkonen y Luukkonen (1993) sugieron una corrección paramétrica basada en la estructura ARMA(p, q) de los errores. Tam y Reinsel (1997) extendieron estos dos contrastes para detectar no invertibilidad en el modelo estacional IMA(1, 1) $_s$ y propusieron una corrección paramétrica alternativa para tratar la correlación serial. En un segundo artículo, Tam y Reinsel (1998) ampliaron el modelo IMA(1, 1) $_s$ impuro para permitir la presencia de un término constante.

El contraste LBIU de no-invertibilidad en un modelo IMA(1,1) es equivalente al contraste LBI de Nyblom y Mäkeläinen (1983) para detectar la estabilidad de la media en el modelo estructural de nivel local. La equivalencia entre estos contrastes tiene su origen tanto en la analogía de los modelos estocásticos usados, como en la de los problemas de decisión considerados. De hecho, en la motivación de su contraste, Nyblom y Mäkeläinen (1983) señalaron que el modelo de nivel local pertenece a la familia de modelos IMA(1,1) y que la hipótesis de estabilidad paramétrica se corresponde con la de no invertibilidad. Sin embargo, se decantaron por la especificación estructural porque la representación IMA no proporciona ninguna simplificación del contraste. Así mismo, la extensión de este contraste propuesta por Nyblom (1986) para incluir una deriva determinista es equivalente a la de Tam y Reinsel (1998) para detectar no invertibilidad en un modelo IMA(1,1) con término constante. No obstante, la estrecha relación entre estos contrastes no ha sido siempre advertida, como lo pone de manifiesto el hecho de que los trabajos de Nyblom y Mäkeläinen (1983) y Nyblom (1986) no sean citados por Tanaka (1990) ni Tam y Reinsel (1997, 1998).

La estrecha relación entre los contrastes LBIU de no invertibilidad y los contrastes LBI de estabilidad paramétrica también se encuentra en los modificaciones propuestas para permitir correlación serial en los errores, aunque no siempre es resaltada suficientemente. Así, el popular contraste de Kwiatkowski et al. (1992) puede contemplarse como una extensión no paramétrica de los contrastes de Nyblom y Mäkeläinen (1983) y Nyblom (1986) y es comparable al contraste no paramétrico de Tanaka (1990). De la misma manera, el contraste de Leybourne y McCabe (1994), que se presenta como la versión paramétrica del contraste de Kwiatkowski et al. (1992), es similar a los contraste de Saikkonen y Luukkonen (1993) y Tam y Reinsel (1997, 1998). Curiosamente Leybourne y McCabe (1994) formularon el problema de decisión en un modelo estructural, pero basaron el procedimiento de contraste en un modelo ARIMA(p, 1, 1).

En este capítulo se revisan todos estos contrastes siguiendo una aproximación común basada en nuevas identidades matriciales que permiten no sólo apreciar claramente sus similitudes y diferencias, sino también simplificar la derivación de bastantes resultados. Esta revisión comienza con una completa derivación del contraste de Nyblom y Mäkeläinen (1983), que se considera aquí como el trabajo pionero en la teoría de contrastes óptimos de raíces MA unitarias y que permite ilustrar la aplicación de esta metodología en el marco de un modelo muy simple. Después, se revisan las extensiones de este contraste básico agrupándolas en dos categorías según propongan un tratamiento paramétrico o no paramétrico de la correlación serial. Como principal aportación se propone una nueva corrección paramétrica del contraste LBI de Nyblom y Mäkeläinen (1983) para detectar estabilidad en la media de un modelo de nivel local con correlación serial, que se corresponde con el contraste LBIU de Saikkonen y Luukkonen (1993) para detectar no invertibilidad en un modelo ARIMA(p, 1, q + 1). A diferencia de estos últimos, aquí se deriva la distribución límite del estadístico corregido tanto bajo la hipótesis nula como bajo una secuencia de alternativas locales. Se demuestra que los resultados distribucionales asintóticos establecidos por Nyblom y Mäkeläinen (1983) en el modelo simple se aplican también en el generalizado. Además, se propone un procedimiento eficiente para calcular el estadístico de contraste basado en los residuos exactos del modelo estructural o ARIMA. Una ventaja comparativa de la nueva corrección paramétrica es que permite la presencia de otras raíces MA unitarias, aunque en este caso deben usarse valores críticos distintos por el desplazamiento que experimenta la distribución muestral del estadístico, que pueden calcularse por integración numérica siguiendo los procedimientos de Imhof (1961) o Davies (1973). En particular, el nuevo estadístico corregido permite obtener como casos especiales el contraste de Nyblom (1986) para detectar una tendencia lineal determinista y otros contrastes de estacionalidad determinista que se revisan en el siguiente capítulo. Finalmente, se realize un estudio comparativo de las potencias de los diferentes contrastes basado, primero, en la evaluación exacta de las diferentes distribuciones cuando los parámetros del modelo son conocidos y, después, en un experimento Monte Carlo cuando los parámetros son desconocidos.

El capítulo se organiza del siguiente modo. En la primera sección se proporciona una revisión completa del contraste LBI de Nyblom y Mäkeläinen (1983) y se proponen unas identidades matriciales que simplifican la derivación de su distribución muestral. La notación establecida en esta sección se utiliza después para obtener los contrastes LBIU de no invertibilidad siguiendo las aproximaciones de Tanaka (1990) y Saikkonen y Luukkonen (1993), y para probar la equivalencia entre estos tres estadísticos. En la tercera sección se revisan tanto las correcciones no paramétricas propuestas por Tanaka (1990) y Kwiatkowski et al. (1992), como las correcciones paramétricas de Saikkonen y Luukkonen (1993), Leybourne y McCabe (1994) y Tam y Reinsel (1997); y se señalan al-
gunas limitaciones que justifican la propuesta, en la sección cuarta, de nueva corrección paramétrica que se aplica también al caso de múltiples raíces MA unitarias. El capítulo finaliza con un estudio comparativo de las potencias de los diferentes contrastes en presencia de autocorrelación.

3.2. El contraste de Nyblom y Mäkeläinen

La derivación de un contraste de hipótesis siguiendo la metodología de King y Hillier (1985) descrita en el capítulo anterior requiere formular el modelo de interés en la notación matricial usada para el modelo clásico generalizado (2.1). Una vez completado este paso, la aplicación de los teoremas 2 o 3, es decir, del contraste LBI o LBUI dependerá de si la hipótesis nula se contrasta frente a una hipótesis alternativa unilateral o bilateral, respectivamente. El estadístico de contraste resultante será una ratio de formas cuadráticas en variables normales, cuya distribución bajo H_0 y H_1 puede evaluarse usando el procedimiento de Imhof (1961), cuyo coste computacional se reduce considerablemente si se obtienen expresiones analíticas para los autovalores de las matrices que definen las formas cuadráticas. Estos autovalores se usan también para derivar la distribución límite.

3.2.1. Formulación matricial del modelo de nivel local

El modelo de nivel local (véase, p. ej., Harvey 1989) viene definido por la ecuación

$$y_t = \mu_t + u_t, \quad t = 1, \dots, n,$$

 $\mu_t = \mu_{t-1} + v_t,$
(3.1)

en donde la serie temporal observada y_t se expresa como la suma de un paseo aleatorio μ_t y un proceso de ruido blanco u_t . Se supone, además, que el valor premuestral μ_0 es no estocástico y que los errores u_t y v_t cumplen las siguientes propiedades: $u_t \sim$ iid $N(0, \sigma_u^2), v_t \sim \text{iid}N(0, \sigma_v^2)$ y $E(u_t v_{t-k}) = 0 \forall k$.

Ordenando las observaciones temporales en el vector columna $\mathbf{y} = (y_1 \ y_2 \ \dots \ y_n)'$, (3.1) puede escribirse en notación matricial como

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{u},$$

$$\mathbf{D}\boldsymbol{\mu} = \mathbf{i}_{(1)}\boldsymbol{\mu}_0 + \mathbf{v},$$
(3.2)

en donde $\mu = (\mu_1 \ \mu_2 \ \dots \ \mu_n)'$, $\mathbf{u} = (u_1 \ u_2 \ \dots \ u_n)'$ y $\mathbf{v} = (v_1 \ v_2 \ \dots \ v_n)'$ son vectores columna de orden n, $\mathbf{i}_{(1)} = (1 \ 0 \ \dots \ 0)'$ es un vector $n \times 1$ de zeros con un 1 en la primera posición, y

$$\mathbf{D} = \begin{pmatrix} 1 & & \\ -1 & 1 & \mathbf{0} \\ & \ddots & \ddots \\ \mathbf{0} & & -1 & 1 \end{pmatrix}$$
(3.3)

es matriz cuadrada bidiagonal de orden n con todos los elementos de la diagonal principal iguales a 1 y todos los elementos de la primera subdiagonal iguales a -1. Notando que la inversa de **D** es la matriz acumulativa o matriz triangular inferior de unos

$$\mathbf{C} = \mathbf{D}^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \dots & 1 \end{pmatrix}$$
(3.4)

y que $Ci_{(1)} = i$ es un vector $n \times 1$ de unos, (3.2) puede escribirse como

$$\mathbf{y} = \mathbf{i}\mu_0 + \mathbf{C}\mathbf{v} + \mathbf{u},\tag{3.5}$$

que es un caso especial del modelo clásico generalizado (2.1) con X = i, $\beta = \mu_0 y$ e = u + Cv. Consecuentemente, la distribución muestral de y es

$$\mathbf{y} \sim N\left(\mathbf{i}\mu_0, \sigma_u^2 \mathbf{\Omega}_e(\rho)\right),\tag{3.6}$$

en donde $\Omega_e(\rho) = \mathbf{I}_n + \rho \mathbf{C}\mathbf{C}', \rho = \sigma_v^2 / \sigma_u^2$ es la denominada ratio señal-ruido y $\mathbf{C}\mathbf{C}' = [\min(i, j)].$

3.2.2. Estadístico de contraste

Nyblom y Mäkeläinen (1983) plantearon el contraste de hipótesis

$$H_0: \rho = 0$$
 versus $H_1: \rho > 0.$ (3.7)

Bajo H_0 , el modelo de nivel local (3.1) se reduce a la regresión sobre un término constan-

te y se dice que la serie y_t es estacionaria o integrada de orden cero, $y_t \sim I(0)$; mientras que bajo H_1 , la serie y_t es no-estacionaria o integrada de orden uno, $y_t \sim I(1)$. Así, el contraste de ratio señal-ruido nula es de hecho un contraste de estacionariedad frente a no-estacionariedad.

El problema estadístico (3.7) es invariante a transformaciones lineales de la serie y_t y puede considerarse como un caso especial del contraste de matriz de varianzas y covarianzas escalar descrito en el capítulo anterior. Como la hipótesis alternativa es unilateral, resulta de aplicación el Teorema 2, según el cual el contraste LBI rechaza H_0 cuando

$$\mathcal{NM}_n = \frac{\hat{\mathbf{e}}'\mathbf{C}\mathbf{C}'\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}} > \kappa, \tag{3.8}$$

en donde $d\Omega_e(\rho)/d\rho = \mathbf{C}\mathbf{C}'$ y $\hat{\mathbf{e}}$ son los residuos minimocuadráticos en la regresión de y sobre un vector de unos i, esto es, $\hat{\mathbf{e}} = \mathbf{M}\mathbf{y} = [y_t - \bar{y}]$ con $\mathbf{M} = \mathbf{I}_n - \mathbf{i}(\mathbf{i}\mathbf{i})^{-1}\mathbf{i}$. Así, el vector columna $\mathbf{C}'\mathbf{e} = [\sum_{\tau=t}^n (y_\tau - \bar{y})]$ contiene las sumas acumuladas parciales *descendentes* de residuos, a diferencia de su rotación vertical $\mathbf{C}\mathbf{e} = [\sum_{\tau=1}^t (y_\tau - \bar{y})]$ que contiene las sumas acumuladas parciales *ascendentes* de residuos. Como la suma de residuos es cero, $\sum_{t=1}^n (y_t - \bar{y}) = 0$, se cumple que $\sum_{t=1}^\tau (y_t - \bar{y}) = -\sum_{t=\tau+1}^n (y_t - \bar{y})$ y $\hat{\mathbf{e}}'\mathbf{C}\mathbf{C}'\hat{\mathbf{e}} = \hat{\mathbf{e}}'\mathbf{C}'\mathbf{C}\hat{\mathbf{e}}$, por lo que el estadístico \mathcal{NM}_n puede escribirse de dos formas equivalentes

$$\mathcal{NM}_n = \frac{\hat{\mathbf{e}}'\mathbf{C}\mathbf{C}'\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}} = \frac{\hat{\mathbf{e}}'\mathbf{C}'\mathbf{C}\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}}$$

o bien

$$\mathcal{NM}_{n} = \frac{\sum_{t=1}^{n} \left[\sum_{\tau=t}^{n} (y_{\tau} - \bar{y})\right]^{2}}{\sum_{t=1}^{n} (y_{t} - \bar{y})^{2}} = \frac{\sum_{t=1}^{n} \left[\sum_{\tau=1}^{t} (y_{\tau} - \bar{y})\right]^{2}}{\sum_{t=1}^{n} (y_{t} - \bar{y})^{2}}.$$
(3.9)

3.2.3. Distribución exacta

El procedimiento que Nyblom y Mäkeläinen (1983) siguieron para obtener la distribución exacta del estadístico \mathcal{NM}_n se simplifica aquí al advertir que la matriz de proyección $\mathbf{M} = \mathbf{I}_n - \mathbf{i}(\mathbf{i'i})^{-1}\mathbf{i'}$, que surge en la regresión sobre una constante, puede expresarse como

$$\mathbf{M} = \boldsymbol{\nabla}' (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-1} \boldsymbol{\nabla}, \tag{3.10}$$

en donde

$$\boldsymbol{\nabla} = \begin{pmatrix} -1 & 1 & & \\ & -1 & 1 & \mathbf{0} \\ & \mathbf{0} & \ddots & \ddots \\ & & & -1 & 1 \end{pmatrix}$$
(3.11)

es una matriz $(n - 1) \times n$ que transforma el vector de observaciones **y** en el vector de primeras diferencias $\mathbf{z} = (z_2 \dots z_n)'$ y $z_t = y_t - y_{t-1}$. Hay que advertir también que la matriz de primeras diferencias ∇ es la submatriz formada por las n - 1 primeras filas de

$$\mathbf{D} = egin{pmatrix} \mathbf{i}_{(1)} \ \mathbf{
abla} \end{pmatrix}$$

y que

$$\boldsymbol{\nabla}\mathbf{C} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} = (\mathbf{0} \mid \mathbf{I}_{n-1})$$

es una matriz $(n-1) \times n$ con unos en la primera superdiagonal y ceros en el resto. Así,

$$\boldsymbol{\nabla}\mathbf{C}\mathbf{C}'\boldsymbol{\nabla}' = \mathbf{I}_{n-1}.\tag{3.12}$$

En la derivación de la distribución exacta del estadístico \mathcal{NM}_n se siguen los siguientes pasos. En primer lugar, se parte de la forma del estadístico

$$\mathcal{NM}_n = \frac{\mathbf{y'MCC'My}}{\mathbf{y'My}},\tag{3.13}$$

que se obtiene reemplazando $\hat{\mathbf{e}} = \mathbf{M}\mathbf{y}$ en (3.8) y que, usando las identidades matriciales (3.10) y (3.12), puede expresarse como

$$\mathcal{NM}_n = \frac{\mathbf{y}' \mathbf{\nabla}' (\mathbf{\nabla} \mathbf{\nabla}')^{-2} \mathbf{\nabla} \mathbf{y}}{\mathbf{y}' \mathbf{\nabla}' (\mathbf{\nabla} \mathbf{\nabla}')^{-1} \mathbf{\nabla} \mathbf{y}}.$$
(3.14)

En segundo lugar, se define la transformación lineal de y

$$\boldsymbol{\epsilon} = (\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-1/2}\boldsymbol{\nabla}\mathbf{y} \sim N\left(0,\sigma_u^2\left(\mathbf{I}_{n-1} + \rho(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-1}\right)\right),$$

que permite escribir (3.14) como

$$\mathcal{NM}_n = \frac{\boldsymbol{\epsilon}' (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-1} \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}' \boldsymbol{\epsilon}}, \qquad (3.15)$$

en donde la matriz $({oldsymbol
abla} {oldsymbol
abla})^{-1}$ admite la descomposición espectral

$$(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-1} = \sum_{k=1}^{n-1} \lambda_k \mathbf{p}_k \mathbf{p}'_k, \qquad (3.16)$$

y en donde los autovalores λ_k (k = 1, ..., n-1) son todos positivos y sus autovectores asociados \mathbf{p}_k están normalizados ($\mathbf{p}'_k \mathbf{p}_k = 1$), son mutuamente ortogonales ($\mathbf{p}'_j \mathbf{p}_k = 0$ para todo $j \neq k$) y cumplen la propiedad

$$\sum_{k=1}^{n-1} \mathbf{p}_k \mathbf{p}'_k = \mathbf{I}_{n-1}.$$

Así, (3.15) puede expresarse como

$$\mathcal{NM}_n = rac{oldsymbol{\epsilon}'\left(\sum_{k=1}^{n-1}\lambda_k \mathbf{p}_k \mathbf{p}_k'
ight)oldsymbol{\epsilon}}{oldsymbol{\epsilon}'\left(\sum_{k=1}^{n-1}\mathbf{p}_k \mathbf{p}_k'
ight)oldsymbol{\epsilon}},$$

en donde $\mathbf{p}'_k \boldsymbol{\epsilon} \sim \operatorname{iid} N(0, \sigma_u^2(1 + \rho \lambda_k))$. Finalmente, definiendo la forma cuadrática $\boldsymbol{\xi}_k = \boldsymbol{\epsilon}' \mathbf{p}_k \mathbf{p}'_k \boldsymbol{\epsilon} / [\sigma_u^2(1 + \rho \lambda_k)] \sim \operatorname{iid} \chi_1^2$, se obtiene

$$\mathcal{NM}_n = \frac{\sum_{k=1}^{n-1} \lambda_k (1+\rho\lambda_k) \xi_k}{\sum_{k=1}^{n-1} (1+\rho\lambda_k) \xi_k},$$
(3.17)

en donde \mathcal{NM}_n aparece expresado como una ratio de dos combinaciones lineales de variables χ_1^2 mutuamente independientes.

Anderson (1992, p. 292) proporcionó¹ expresiones analíticas para los autovalores de la

¹El apéndice ?? describe el método de obtención de estos autovalores, pág. ??.

matriz $\boldsymbol{\nabla} \boldsymbol{\nabla}'$ dadas por

$$\lambda_k^{-1} = 2[1 - \cos(\frac{\pi k}{n})] = 4\sin^2\left(\frac{\pi k}{2n}\right), \quad k = 1, \dots, n-1,$$
(3.18)

en donde la segunda igualdad resulta de la identidad trigonométrica

$$\sin^2(x) = \frac{1}{2} - \frac{1}{2}\cos(2x).$$

Aquí, y de acuerdo con (2.23), se usan estas expresiones para obtener la media y varianza de \mathcal{NM}_n bajo H_0

$$E(\mathcal{NM}_n) = \overline{\lambda}_1 \quad \text{y} \quad \text{var}(\mathcal{NM}_n) = \frac{2}{m+2}[\overline{\lambda}_2 - \overline{\lambda}_1^2],$$

en donde $\bar{\lambda}_1$ y $\bar{\lambda}_2$ son la media y media cuadrática de los autovalores λ_k , respectivamente, que pueden calcularse usando las siguientes fórmulas propuestas por Gallego y Díaz (2011)

$$\bar{\lambda}_1 = \frac{n+1}{6} \quad \mathbf{y} \quad \bar{\lambda}_2 = \bar{\lambda}_1 \times \left(\frac{2n^2+7}{30}\right). \tag{3.19}$$

Estas fórmulas revelan la necesidad de corregir el estadístico \mathcal{NM}_n por un factor dependiente del tamaño muestral *n* para garantizar la existencia de los dos primeros momentos en muestras grandes. Nyblom y Mäkeläinen (1983) usaron como factor de corrección n - 1, el número de autovalores λ_k .

La función de distribución acumulada (cdf, en adelante) de $\mathcal{NM}_n/(n-1)$,

$$\Pr(\mathcal{NM}_n/(n-1) < \kappa) = \Pr\left[\sum_{k=1}^{n-1} \left(\frac{\lambda_k}{n-1} - \kappa\right) (1 + \rho \lambda_k) \xi_k < 0\right], \quad (3.20)$$

puede evaluarse mediante el procedimiento de Imhof (1961) descrito en el capítulo anterior, y cuyo coste computacionales se reduce notablemente al conocer la expresión exacta de los autovalores. La ecuación (3.20) puede usarse tanto para calcular valores críticos bajo H_0 como para evaluar la potencia del contraste bajo H_1 .

3.2.4. Distribución límite

Nyblom y Mäkeläinen (1983) encontraron que la distribución límite de \mathcal{NM}_n/n bajo

la secuencia de alternativas locales $H_1: \rho = c^2/n^2$ viene dada por

$$\mathcal{NM}_n/n \stackrel{a}{\sim} \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4} \right) \xi_k. \tag{3.21}$$

Para obtener esta distribución asintótica se escribe (3.17) como

$$\mathcal{NM}_n/n = rac{rac{1}{n^2}\sum_{k=1}^{n-1}\lambda_k(1+c^2n^{-2}\lambda_k)\xi_k}{rac{1}{n}\sum_{k=1}^{n-1}(1+c^2n^{-2}\lambda_k)\xi_k},$$

en donde se ha dividido el numerador y denominador por *n* y se ha reemplazado ρ por c^2/n^2 . De (3.18), y recordando que $\sin(x)/x \to 1$ cuando $x \to 0$, se tiene que $n^2 \lambda_k^{-1} = (\pi k)^2 \{ [(2n)/(\pi k)] \sin[\pi k/(2n)] \}^2 \to (\pi j)^2$ cuando $n \to \infty$. De aquí, el numerador

$$\frac{1}{n^2}\sum_{k=1}^{n-1}\lambda_k(1+c^2n^{-2}\lambda_k)\xi_k \xrightarrow{d} \sum_{k=1}^{\infty} \left[\frac{1}{\pi^2k^2}+\frac{c^2}{\pi^4k^4}\right]\xi_k$$

y el denominador

$$\frac{1}{n} \sum_{k=1}^{n-1} (1 + c^2 n^{-2} \lambda_k) \xi_k \xrightarrow{p} 1, \qquad (3.22)$$

porque $(c/n) \sum_{k=1}^{n-1} \xi_k / (\pi^2 k^2) \xrightarrow{p} 0$ y $\sum_{k=1}^{n-1} \xi_k / n \xrightarrow{p} 1$.

Nyblom y Mäkeläinen (1983) notaron que la distribución límite de \mathcal{NM}_n/n bajo H_0 es equivalente a la distribución límite del estadístico Cràmer-von Mises de bondad de ajuste, que puede evaluarse usando la serie dada por Anderson y Darling (1952)

$$F(\kappa) = \frac{1}{\pi\kappa} \sum_{j=0}^{\infty} (-1)^j \binom{-1/2}{j} (4j+1)^{1/2} \exp[-(4j+1)^{1/2}/(16\kappa)] K_{1/4}[(4j+1)^2/(16\kappa)],$$
(3.23)

en donde $K_{1/4}[z]$ es la función Bessel modificada de segunda clase. La serie converge muy rápidamente cuando $\kappa < 2$, siendo por tanto posible elegir un punto de truncamiento *J* a partir del cual los términos de la serie se anulan. En la práctica, *J* suele ser menor que 10.

La función de distribución (3.21) no puede evaluarse por Imhof (1961) al ser una forma cuadrática infinita. Además, hay que advertir que la aproximación tentativa de truncar

la serie asintótica eligiendo un tamaño muestral n grande,

$$F(\kappa) \simeq \Pr\left[\sum_{k=1}^{n-1} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4}\right) \xi_k < \kappa\right]$$

resulta bastante imprecisa porque se estarían aproximando bastantes autovalores. Aquí se ha tabulado la distribución asintótica mediante el método de Davies (1980) usando un tamaño muestral muy grande, n = 5000, y los autovalores exactos (3.18), obteniéndose unos valores críticos prácticamente idénticos a los de Anderson y Darling (1952).

3.2.5. El contraste invariante óptimo puntual

Aunque no lo derivaron explícitamente, Nyblom y Mäkeläinen (1983) consideraron también el contraste invariante más potente (MPI), también llamado contraste invariante óptimo puntual, para el problema de decisión H_0 : $\rho = 0$ frente a la alternativa simple H_1 : $\rho = \rho_1$, el cual es útil para evaluar la potencia de $\mathcal{NM}_n/(n-1)$. Del Teorema 1, el contraste MPI rechaza H_0 cuando

$$\mathcal{P}_n = rac{\mathbf{\tilde{e}}' \mathbf{\Omega}_e(\mathbf{\rho}_1)^{-1} \mathbf{\tilde{e}}}{\mathbf{\hat{e}}' \mathbf{\hat{e}}}$$

toma valores pequeños, en donde $\tilde{\mathbf{e}}$ y $\hat{\mathbf{e}}$ son los residuos de la estimación MCG y MCO, respectivamente. Para derivar la distribución exacta del estadístico \mathcal{P}_n cuando el parámetro verdadero es ρ , primero es necesario expresarlo en términos de **y**

$$\mathcal{P}_n = \frac{\mathbf{y}'_* \mathbf{M}_*(\rho_1) \mathbf{y}_*}{\mathbf{y}' \mathbf{M} \mathbf{y}} = \frac{\mathbf{y}' \mathbf{\Omega}_e(\rho_1)^{-1/2} \mathbf{M}_*(\rho_1) \mathbf{\Omega}_e(\rho_1)^{-1/2} \mathbf{y}}{\mathbf{y}' \mathbf{M} \mathbf{y}},$$

en donde hay que notar que el parámetro verdadero ρ no tiene que coincidir necesariamente con el valor ρ_1 fijado en H_1 . En segundo lugar, hay que notar también que la matriz idempotente $M_*(\rho_1)$ puede escribirse como

$$\boldsymbol{M}_{*}(\rho_{1}) = \boldsymbol{\Omega}_{e}(\rho_{1})^{1/2} \boldsymbol{\nabla}' (\boldsymbol{\nabla} \boldsymbol{\Omega}_{e}(\rho_{1}) \boldsymbol{\nabla}')^{-1} \boldsymbol{\nabla} \boldsymbol{\Omega}_{e}(\rho_{1})^{1/2},$$

en donde para simplificar la notación se supone que $\Omega_e(\rho_1)^{1/2}$ es una matriz simétrica². Así,

$$\mathcal{P}_n = rac{\mathbf{y}' \mathbf{\nabla}' (\mathbf{\nabla} \mathbf{\Omega}_e(
ho_1) \mathbf{\nabla}')^{-1} \mathbf{\nabla} \mathbf{y}}{\mathbf{y}' \mathbf{\nabla}' (\mathbf{\nabla} \mathbf{\nabla}')^{-1} \mathbf{\nabla} \mathbf{y}}.$$

Como $\nabla \mathbf{y} \sim N(0, \sigma^2 \nabla \Omega_e(\rho) \nabla')$, definiendo $\boldsymbol{\epsilon} = (\nabla \Omega_e(\rho) \nabla')^{-1/2} \nabla \mathbf{y} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n-1})$, se tiene que

$$\mathcal{P}_n = \frac{\epsilon' (\boldsymbol{\nabla} \boldsymbol{\Omega}_e(\rho) \boldsymbol{\nabla}')^{1/2} (\boldsymbol{\nabla} \boldsymbol{\Omega}_e(\rho_1) \boldsymbol{\nabla}')^{-1} (\boldsymbol{\nabla} \boldsymbol{\Omega}_e(\rho) \boldsymbol{\nabla}')^{1/2} \epsilon}{\epsilon' (\boldsymbol{\nabla} \boldsymbol{\Omega}_e(\rho) \boldsymbol{\nabla}')^{1/2} (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-1} (\boldsymbol{\nabla} \boldsymbol{\Omega}_e(\rho) \boldsymbol{\nabla}')^{1/2} \epsilon}$$

De (3.16) y $\nabla \Omega_e(\rho) \nabla' = \rho \mathbf{I}_{n-1} + \nabla \nabla'$, resultan las siguientes descomposiciones espectrales

$$\nabla \boldsymbol{\Omega}_{e}(\rho) \nabla' = \sum_{k=1}^{n-1} (\rho + \lambda_{k}^{-1}) \mathbf{p}_{k} \mathbf{p}_{k}',$$
$$(\nabla \boldsymbol{\Omega}_{e}(\rho_{1}) \nabla')^{-1} = \sum_{k=1}^{n-1} (\rho_{1} + \lambda_{k}^{-1})^{-1} \mathbf{p}_{k} \mathbf{p}_{k}',$$
$$(\nabla \boldsymbol{\Omega}_{e}(\rho_{1}) \nabla')^{-1} \nabla \boldsymbol{\Omega}_{e}(\rho) \nabla' = \sum_{k=1}^{n-1} [(1 + \rho\lambda_{k})/(1 + \rho_{1}\lambda_{k})] \mathbf{p}_{k} \mathbf{p}_{k}' \quad \mathbf{y}$$
$$(\nabla \nabla')^{-1} \nabla \boldsymbol{\Omega}_{e}(\rho) \nabla' = \sum_{k=1}^{n-1} (1 + \rho\lambda_{k}) \mathbf{p}_{k} \mathbf{p}_{k}',$$

que conducen a

$$\mathcal{P}_{n} = \frac{\sum_{k=1}^{n-1} [(1+\rho\lambda_{k})/(1+\rho_{1}\lambda_{k})]\xi_{k}}{\sum_{k=1}^{n-1} (1+\rho\lambda_{k})\xi_{k}},$$
(3.24)

en donde $\xi_k = \epsilon' \mathbf{p}_k \mathbf{p}'_k \epsilon / \sigma^2 \sim \chi_1^2$. La distribución \mathcal{P}_n en muestras finitas puede evaluarse usando el procedimiento de Imhof (1961).

En la derivación de la distribución límite de \mathcal{P}_n se considera el contraste de $H_0: \rho = 0$ frente a la secuencia de alternativas $H_1: \rho_1 = c_1^2/n^2$ cuando el parámetro verdadero es $\rho = c^2/n^2$. Remplazando estas expresiones de ρ y ρ_1 en (3.24), se obtiene

$$\mathcal{P}_n(c,c_1) = \frac{n^{-1} \sum_{k=1}^{n-1} \left[(1+c^2 n^{-2} \lambda_k) / (1+c_1^2 n^{-2} \lambda_k) \right] \xi_k}{n^{-1} \sum_{k=1}^{n-1} (1+c^2 n^{-2} \lambda_k) \xi_k} \xrightarrow{p} 1,$$

porque, por un lado, los pesos del numerador $(1 + c^2 n^{-2} \lambda_k) / (1 + c_1^2 n^{-2} \lambda_k) \rightarrow 1$ cuando $k \rightarrow \infty$ y, por otro, de acuerdo con (3.22) el denominador converge en probabilidad

²En una implementación práctica, $\Omega_e(\rho_1)^{1/2}$ sería el factor de Cholesky de $\Omega_e(\rho_1)$, que sería una matriz triangular inferior, y habría que distinguir entre $\Omega_e(\rho_1)^{1/2}$ y su traspuesta, $\Omega_e(\rho_1) = \Omega_e(\rho_1)^{1/2} \Omega_e(\rho_1)^{1/2}$.

a 1. Por tanto, la distribución límite de $\mathcal{P}_n(c,c_1)$ es degenerada y es necesario realizar la siguiente normalización

$$n[1-\mathcal{P}_n(c,c_1)] = \frac{\sum_{k=1}^{n-1} \left[\frac{(c_1^2 n^{-2} \lambda_k)(1+c^2 n^{-2} \lambda_k)}{(1+c_1^2 n^{-2} \lambda_k)} \right] \tilde{\xi}_k}{n^{-1} \sum_{k=1}^{n-1} (1+c^2 n^{-2} \lambda_k) \tilde{\xi}_k} \xrightarrow{d} \sum_{k=1}^{\infty} \frac{c_1^2 (\pi^2 k^2 + c^2)}{(\pi^2 k^2 + c_1^2) \pi^2 k^2} \tilde{\xi}_k.$$

Así, bajo $H_0: c = 0$,

$$n[1 - \mathcal{P}_n(0, c_1)] = \frac{\sum_{k=1}^{n-1} [c_1^2 n^{-2} \lambda_k / (1 + c_1^2 n^{-2} \lambda_k)] \xi_k}{n^{-1} \sum_{k=1}^{n-1} \xi_k} \xrightarrow{d} \sum_{k=1}^{\infty} \frac{c_1^2}{\pi^2 k^2 + c_1^2} \xi_k$$

Análogamente, bajo $H_1 : c = c_1$,

$$n[1 - \mathcal{P}_n(c_1, c_1)] = \frac{\sum_{k=1}^{n-1} (c_1^2 n^{-2} \lambda_k) \xi_k}{n^{-1} \sum_{k=1}^{n-1} (1 + c_1^2 n^{-2} \lambda_k) \xi_k} \xrightarrow{d} \sum_{k=1}^{\infty} \frac{c_1^2}{\pi^2 k^2} \xi_k$$

En este caso, el resultado $\mathcal{P}_{\infty}(c_1, c_1) \xrightarrow{p} 1$ puede justificarse como una consecuencia de la propiedad de consistencia del estimador MCO en el modelo clásico generalizado: como $|\hat{\boldsymbol{\beta}}_{MCO} - \tilde{\boldsymbol{\beta}}_{MCG}| \xrightarrow{p} 0$, resulta que $|\mathbf{\tilde{e}}' \mathbf{\Omega}_e(\rho_1)^{-1} \mathbf{\tilde{e}} - \mathbf{\hat{e}}' \mathbf{\hat{e}}| \xrightarrow{p} 0$.

La distribución muestral de $n[1 - \mathcal{P}(c, c_1)]$ se utiliza para calcular numéricamente la denominada potencia envolvente, que requiere una doble tabulación. En primer lugar, se tabula la distribución bajo H_0 : c = 0, que proporciona los valores críticos para el contraste MPI. Y, en segundo lugar, se tabula la distribución bajo una secuencia de alternativas H_1 : $c = c_1$.

3.2.6. Valores críticos y potencia del contraste

El cuadro 3.1 contiene algunos percentiles de la distribución exacta del estadístico $\mathcal{NM}_n/(n-1)$ para tamaños de muestra $n = \{11(10)101, \infty\}$, que se han obtenido usando una implementación propia en C++ del procedimiento de Davies (1973) e incorporada en el programa Empiricus. Puede comprobarse que los valores críticos para muestras de más de 61 observaciones son muy similares a los asíntoticos, de manera que la convergencia a la distribución asintótica es muy rápida. Además, como se deduce también de los momentos (3.19), se observa que la distribución muestral se desplaza ligeramente hacia la izquierda al aumentar el tamaño muestral. Estas dos características (convergencia rápida y desplazamiento hacia la izquierda) se aprecian más claramente

					α				
n	0.99	0.925	0.95	0.90	0.50	0.10	0.05	0.025	0.01
11	0.044	0.050	0.058	0.069	0.154	0.401	0.504	0.599	0.707
21	0.034	0.040	0.047	0.057	0.136	0.375	0.485	0.593	0.730
31	0.031	0.037	0.043	0.053	0.130	0.366	0.477	0.590	0.736
41	0.029	0.035	0.042	0.051	0.127	0.361	0.474	0.587	0.738
51	0.028	0.034	0.041	0.050	0.126	0.358	0.471	0.586	0.738
61	0.028	0.033	0.040	0.050	0.125	0.357	0.469	0.585	0.741
71	0.027	0.033	0.039	0.049	0.124	0.355	0.468	0.585	0.741
81	0.027	0.033	0.039	0.049	0.123	0.354	0.468	0.584	0.741
91	0.027	0.032	0.039	0.048	0.123	0.353	0.467	0.584	0.741
101	0.027	0.032	0.039	0.048	0.122	0.353	0.466	0.584	0.741
∞	0.025	0.030	0.037	0.046	0.119	0.347	0.461	0.580	0.744

Cuadro 3.1: Percentiles de la distribución $\Pr(\mathcal{NM}_n/(n-1) > \kappa) = \alpha$



Figura 3.1: Cdf y pdf del estadístico NM_n

en la figura 3.1, que muestra la cdf y la función de densidad de probabilidad (pdf, en adelante) del estadístico de contraste para tres tamaños muestrales. El programa Empiricus permite calcular valores críticos utilizando la sentencia "davies nm -n[] -p[]", en donde los modificadores -n[] y -p[] especifican los tamaños de muestra y las probabilidades, respectivamente. El programa también permite calcular *p*-valores introduciendo la sentencia "davies nm -n[] -x[]", en donde el modificador -x[] especifica los valores del estadístico. Opcionalmente, el comando davies puede reemplazarse por el de imhof.

El cuadro 3.2 muestra la potencia del contraste al nivel de significación $\alpha = 0.05$ para tamaños de muestra $n = \{11(10)101, \infty\}$ y valores de $\rho = \{0.0006, 0.003, 0.006, 0.011, 0.026, 0.05, 0.083, 0.129\}$. De (3.20), la potencia se calcula como

$$\Pr(\mathcal{NM}_n/(n-1) > \kappa_{\alpha,n}) = 1 - \Pr\left[\sum_{k=1}^{n-1} \left(\frac{\lambda_k}{n-1} - \kappa_{\alpha,n}\right) (1+\rho\lambda_k)\xi_k < 0\right],$$

en donde $\kappa_{\alpha,n}$ es el valor crítico que depende de α y *n*; en este caso, se corresponden

con los valores de la octava columna del cuadro 3.1. Para un ρ dado, se observa que la potencia del contraste es siempre mayor que 0.05 (insesgadez) y aumenta monótonamente con el tamaño muestral. De forma similar, dado un tamaño muestral *n*, se observa que la potencia aumenta monotonamente al alejarse ρ de cero. La potencia puede calcularse en Empiricus con la sentencia "davies nm -n[] -p[] -rho[]", en donde el modificador -rho[] especifica los valores de ρ bajo H_1 .

				ρ				
n	0.0006	0.003	0.006	0.011	0.018	0.026	0.037	0.05
11	0.051	0.053	0.057	0.062	0.069	0.079	0.090	0.103
21	0.053	0.062	0.078	0.102	0.132	0.168	0.208	0.250
31	0.057	0.078	0.115	0.166	0.224	0.286	0.346	0.401
41	0.062	0.101	0.164	0.242	0.322	0.397	0.464	0.523
51	0.069	0.129	0.221	0.320	0.412	0.492	0.560	0.618
61	0.078	0.163	0.279	0.393	0.491	0.572	0.638	0.692
71	0.088	0.200	0.337	0.459	0.558	0.637	0.700	0.750
81	0.100	0.238	0.391	0.517	0.616	0.692	0.751	0.797
91	0.114	0.277	0.441	0.569	0.665	0.737	0.792	0.840
101	0.128	0.316	0.487	0.615	0.708	0.775	0.833	0.866
∞	0.991	1	1	1	1	1	1	1

Cuadro 3.2: Potencia del estadístico $\mathcal{NM}_n/(n-1)$

El cuadro 3.3 contiene los valores críticos para el estadístico $n[1 - P_n]$ al nivel de significación $\alpha = 0.05$, que son necesarios para el cálculo de la potencia envolvente que se muestra en el cuadro 3.4. A diferencia de Nyblom y Mäkeläinen (1983) que consideraron valores de ρ bajo H_1 comprendidos entre 0.05 y 5, aquí se consideran valores comprendidos entre 6.41×10^{-4} y 0.05 porque son los que suelen encontrarse en aplicaciones prácticas. Comparando las columnas del cuadros 3.2 y 3.4, se comprueba que el estadístico POI siempre es más potente que el LBI, siendo prácticamente iguales en la proximidades de H_0 porque ambos son los más potentes. Los valores críticos del estadístico $n[1 - P_n]$ pueden calcularse en Empiricus con la sentencia "davies nmpoi -n[] -p[] -rho1[]", en donde el modificador <math>-rho1[] especifica los valores de ρ bajo H_1 . El modificador -p[] puede reemplazarse por -x[] para calcular p-valores. Mientras que la sentencia "davies nmpoi -n[] -x[] -rho[]" permitecalcular la potencia del contraste.

3.2.7. Inclusión de una deriva

Nyblom (1986) extendió el contraste de Nyblom y Mäkeläinen (1983) al modelo (3.1)

	ρ1								
n	0.0006	0.003	0.006	0.011	0.018	0.026	0.037	0.050	
11	0.032	0.129	0.288	0.500	0.754	1.038	1.342	1.651	
21	0.121	0.464	0.958	1.526	2.110	2.677	3.221	3.738	
31	0.262	0.930	1.768	2.609	3.400	4.144	4.851	5.528	
41	0.444	1.461	2.572	3.607	4.562	5.459	6.323	7.166	
51	0.661	2.003	3.330	4.525	5.633	6.688	7.714	8.719	
61	0.902	2.535	4.041	5.390	6.646	7.857	9.044	10.21	
71	1.160	3.047	4.708	6.208	7.621	8.988	10.33	11.66	
81	1.426	3.535	5.351	7.002	8.565	10.09	11.60	13.09	
91	1.697	4.003	5.968	7.767	9.487	11.17	12.84	14.49	
101	1.968	4.452	6.565	8.514	10.39	12.23	14.06	15.88	

Cuadro 3.3: Valores críticos para estadístico $n[1 - P_n]$ al nivel del 5 %

Cuadro 3.4: Potencia del estadístico $n[1 - P_n]$ con $\alpha = 0.05$

	ρ_1								
n	6.41e-04	0.003	0.006	0.011	0.018	0.026	0.037	0.050	
11	0.051	0.053	0.057	0.062	0.069	0.079	0.090	0.104	
21	0.053	0.062	0.078	0.102	0.132	0.169	0.210	0.255	
31	0.057	0.078	0.115	0.166	0.228	0.294	0.361	0.428	
41	0.062	0.101	0.165	0.247	0.334	0.422	0.504	0.579	
51	0.069	0.130	0.224	0.332	0.439	0.538	0.624	0.698	
61	0.078	0.164	0.287	0.416	0.535	0.636	0.720	0.788	
71	0.088	0.202	0.351	0.495	0.618	0.717	0.794	0.853	
81	0.100	0.242	0.413	0.566	0.689	0.782	0.850	0.899	
91	0.114	0.284	0.472	0.630	0.749	0.834	0.892	0.931	
101	0.129	0.327	0.528	0.686	0.798	0.874	0.923	0.953	

ampliado por la inclusión de un término constante δ en la ecuación del nivel

$$y_t = \mu_t + u_t,$$

 $\mu_t = \mu_{t-1} + \delta + v_t, \quad t = 1, \dots, n,$
(3.25)

en donde μ_t sigue ahora un paseo aleatorio con deriva fija. Bajo H_0 : $\rho = 0$, el nivel local se reduce a una tendencia lineal $\mu_t = \mu_0 + \delta t$.

En notación matricial, (3.25) puede escribirse como

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{u},$$

$$\mathbf{D}\boldsymbol{\mu} = \mathbf{i}_{(1)}\boldsymbol{\mu}_0 + \mathbf{i}\boldsymbol{\delta} + \mathbf{v},$$
(3.26)

que, reemplazando la ecuación de μ en la de y, conduce a

$$\mathbf{y} = \mathbf{i}\mu_0 + \mathbf{t}\delta + \mathbf{C}\mathbf{v} + \mathbf{u},\tag{3.27}$$

en donde $\mathbf{t} = (1 \ 2 \ \dots \ n)'$. Comparando (3.30) con (3.5) se observa que la inclusión de la deriva δ cambia la media de la distribución de \mathbf{y} , pero deja inalterada su matriz de covarianzas. De aquí, el contraste LBI para el problema de decisión (3.7) bajo (3.30) es el dado en (3.8) con la salvedad de que el vector $\hat{\mathbf{e}} = \mathbf{M}\mathbf{y}$ contiene ahora los residuos MCO en la regresión lineal de \mathbf{y} sobre la matriz de diseño $\mathbf{X} = (\mathbf{i} \ \mathbf{t})$, siendo $\mathbf{M} =$ $\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}$.

Para derivar la distribución exacta del estadístico de Nyblom (1986), definido por

$$\mathcal{N}_n = rac{\mathbf{y'MCC'My}}{\mathbf{y'My}},$$

se introduce aquí la matriz de segundas diferencias ∇^2 de orden $(n-2) \times (n-1)$, que puede expresarse como

$$\boldsymbol{\nabla}^2 = \boldsymbol{\nabla}_{n-2,n-1} \boldsymbol{\nabla}_{n-1,n},$$

en donde ∇_{orden} es la matriz de primeras diferencias (3.11) de orden indicado por el subíndice. Además, se proponen las identidades matriciales $\mathbf{M} = \nabla^{2'} (\nabla^2 \nabla^{2'})^{-1} \nabla^2 \mathbf{y}$ $\mathbf{MCC'}\mathbf{M} = \nabla^{2'} (\nabla^2 \nabla^{2'})^{-1} \nabla \nabla' (\nabla^2 \nabla^{2'})^{-1} \nabla^2$, que permiten escribir el estadístico de contraste como

$$\mathcal{N}_n = rac{\mathbf{y}' \mathbf{
abla}^{2'} (\mathbf{
abla}^2 \mathbf{
abla}^{2'})^{-1} \mathbf{
abla} \mathbf{
abla}' (\mathbf{
abla}^2 \mathbf{
abla}^{2'})^{-1} \mathbf{
abla}^2 \mathbf{y}}{\mathbf{y}' \mathbf{
abla}^{2'} (\mathbf{
abla}^2 \mathbf{
abla}^{2'})^{-1} \mathbf{
abla}^2 \mathbf{y}}.$$

Definiendo la transformación lineal

$$\boldsymbol{\epsilon} = (\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^{2'})^{-1/2} \boldsymbol{\nabla}^2 \mathbf{y} \sim N\left(0, \sigma_u^2 \left(\mathbf{I}_{n-2} + \rho(\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^{2'})^{-1/2} \boldsymbol{\nabla} \boldsymbol{\nabla}' (\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^{2'})^{-1/2}\right)\right),$$

se obtiene que

$$\mathcal{N}_n = \frac{\boldsymbol{\epsilon}'(\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^{2'})^{-1/2} \boldsymbol{\nabla} \boldsymbol{\nabla}'(\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^{2'})^{-1/2} \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}' \boldsymbol{\epsilon}}$$

Y siguiendo el argumento dado en (3.16)-(3.17), se llega a

$$\mathcal{N}_{n} = \frac{\sum_{k=1}^{n-2} \lambda_{k} (1 + \rho \lambda_{k}) \xi_{k}}{\sum_{k=1}^{n-2} (1 + \rho \lambda_{k}) \xi_{k}},$$
(3.28)

en donde λ_k son los autovalores de $(\nabla^2 \nabla^{2'})^{-1} \nabla \nabla' y \xi_k \sim iid\chi_1^2$. Por desgracia, solo los autovalores λ_k para $k \in 2, 4, ...$ son conocidos, por lo que no es posible expresar la distribución límite en términos de autovalores. No obstante, y de acuerdo con (2.23), la media y varianza de \mathcal{N}_n bajo H_0 ,

$$E(\mathcal{N}_n) = \overline{\lambda}_1 \quad \mathrm{y} \quad \mathrm{var}(\mathcal{N}_n) = \frac{2}{m+2} [\overline{\lambda}_2 - \overline{\lambda}_1^2],$$

pueden calcularse usando las siguientes fórmulas propuestas por Gallego y Díaz (2011)

$$\bar{\lambda}_1 = \frac{n+2}{15}$$
 y $\bar{\lambda}_2 = \bar{\lambda}_1 \times \left(\frac{11n^2 + 181}{840}\right)$, (3.29)

que revelan la necesidad de corregir el estadístico \mathcal{N}_n por un factor dependiente del tamaño muestral *n*. Nyblom (1986) usó como factor de corrección n - 2, el número de autovalores λ_k . En el cuadro 3.5 se presenta valores críticos del contraste $\mathcal{N}_n/(n-2)$ para diferentes tamaños muestrales n = 12(10)102, que se han obtenido con el programa Empiricus usando el listado de comandos 3.1.

Nyblom y Harvey (2001) derivaron también el contraste LBI de tendencia lineal determinista, pero considerando como hipótesis alternativa el paseo aleatorio integrado **Programa 3.1** Comandos de Empiricus para tabular la distribución de $N_n/(n-2)$

```
// Tamaño muestral
n = 12
// Percentiles
p = [0.01; 0.025; 0.05; 0.10; 0.50; 0.90; 0.95; 0.975; 0.99]
// Matriz de segundas diferencias
matrix D2 -n[n] -a[D2]
// Matriz de primeras diferencias
matrix D -n[n-1] -a[D]
// Factor de Cholescky de D2*D2'
T = chol(D2*D2')
// Inversa del factor de Cholescky
T = inv(T)
// Matriz de la forma cuadrática
A = T*D*D'*T'/(n-2)
// Procedimiento de Imhof (Precisión: 3)
imhof cdf -A[A] -B[I] -p[p] -P[3]
// Procedimiento de Davies
//davies cdf -A[A] -B[I] -p[p] -P[3]
```

α 0.005

Cuadro 3.5: Percenticles de la distribución $\Pr(N_n/(n-2) > \kappa) = \alpha$

n	0.99	0.925	0.95	0.90	0.50	0.10	0.05	0.025	0.01
12	0.0370243	0.041	0.044	0.050	0.083	0.152	0.179	0.204	0.232
22	0.0266474	0.030	0.033	0.038	0.069	0.136	0.164	0.192	0.227
32	0.0233958	0.027	0.030	0.035	0.065	0.131	0.159	0.187	0.224
42	0.0218124	0.025	0.028	0.033	0.062	0.128	0.156	0.185	0.223
52	0.0208767	0.024	0.027	0.032	0.061	0.126	0.155	0.183	0.222
62	0.0202628	0.023	0.027	0.031	0.060	0.125	0.154	0.182	0.221
72	0.0198225	0.023	0.026	0.031	0.059	0.124	0.153	0.182	0.221
82	0.0195008	0.023	0.026	0.030	0.059	0.124	0.152	0.181	0.220
92	0.0192467	0.022	0.026	0.030	0.059	0.123	0.152	0.181	0.220
102	0.0189800	0.022	0.025	0.030	0.058	0.123	0.151	0.180	0.220

 $\nabla^2 \mu_t = v_t$. Bajo esta especificación se obtiene el modelo de regresión generalizado

$$\mathbf{y} = \mathbf{i}\mu_0 + \mathbf{t}\delta + \mathbf{C}\mathbf{C}\mathbf{v} + \mathbf{u},\tag{3.30}$$

en donde el término de error v se acumula dos veces. El estadístico LBI toma la forma

$$\mathcal{NH}_n = rac{\mathbf{y'MCCC'C'My}}{\mathbf{y'My}},$$

que requiere una corrección n^3 . Su distribución exacta es la dada en (3.28), pero siendo λ_k (k = 1, ..., n - 2) los autovalores de ($\nabla^2 \nabla^{2'}$)⁻¹. Por desgracia, los autovalores son desconocidos y hay que calcularlos numéricamente. Aquí, se ha intentado sin éxito encontrar expresiones analíticas para λ_k utilizando las identidades matriciales

$$\boldsymbol{\nabla}^{2}\boldsymbol{\nabla}^{2'} = (\boldsymbol{\nabla}\boldsymbol{\nabla}')^{2} + {}^{1}\boldsymbol{0}_{1} \quad y \quad (\boldsymbol{\nabla}^{2}\boldsymbol{\nabla}^{2'})^{-1} = (\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-2} - {}^{1}\boldsymbol{0}_{1} \left[(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-2}(\boldsymbol{\nabla}^{2}\boldsymbol{\nabla}^{2'})^{-1} \right] {}^{1}\boldsymbol{0}_{1},$$

en donde ¹**0**₁ es la matriz nula con unos en las entradas (1,1) y (n - 2, n - 2).

3.3. Los contrastes de Tanaka y Saikkonen-Luukkonen

Nyblom y Mäkeläinen (1983) motivaron su contraste de hipótesis como un procedimiento para detectar estabilidad paramétrica en el modelo de nivel local y señalaron que también podría usarse para detectar no invertibilidad en un modelo IMA(1,1). Sin embargo, fueron Tanaka (1990) y Saikkonen y Luukkonen (1993) quienes derivaron directamente el contraste de raíz MA unitaria. Mientras que el primero consideró el modelo MA(1) regular como representación de interés, los segundos usaron el modelo IMA(1,1) regular, que puede interpretarse como un modelo MA(1) para la serie diferenciada. En consecuencia se han propuesto dos formas alternativas pero equivalentes del estadístico de contraste: una basada en la serie diferenciada; y la otra, en la serie original.

3.3.1. La aproximación de Tanaka

El modelo MA(1) regular (véase, p. ej., Box et al. 2008) se define como

$$z_t = (1 - \theta B)a_t, \quad t = 2, \dots, n,$$
 (3.31)

en donde a_t es un proceso de ruido blanco Gaussiano, $a_t \sim iidN(0, \sigma_a^2)$, y se consideran n - 1 observaciones de z_t para facilitar las comparaciones con (3.1). El modelo MA(1) admite la siguiente forma matricial

$$\mathbf{z} = \boldsymbol{\nabla}(\theta) \mathbf{a},\tag{3.32}$$

en donde $\mathbf{z} = (z_2 \ z_3 \ \dots \ z_n)'$ es un vector columna de n - 1 observaciones estacionarias, $\mathbf{a} = (a_1 \ a_2 \ \dots \ a_n)'$ es un vector columna de n errores aleatorios y $\nabla(\theta)$ es una matriz rectangular de orden $(n - 1) \times n$ con elementos iguales a $-\theta$ en la diagonal principal, a 1 en la primera superdiagonal, y a 0 en las restantes diagonales. Cuando $\theta = 1$, $\nabla(1)$ es la matriz de primeras diferencias ∇ definida en (3.11). Al igual que se ha hecho con ∇ , aquí se contempla $\nabla(\theta)$ como la submatriz formada por las últimas n - 1 filas de la matriz cuadrada $\mathbf{D}(\theta)$ de orden n formada por unos en la diagonal principal, $-\theta$ en la primera subdiagonal y ceros en el resto. Cuando $\theta = 1$, $\mathbf{D}(1)$ es la matriz \mathbf{D} definida en (3.3). Si \mathbf{B} es una matriz cuadrada de orden n con unos en la primera subdiagonal y ceros en el resto, entonces $\mathbf{D}(\theta) = \mathbf{I}_n - \theta \mathbf{B}$,

$$\mathbf{D}(\theta) = \begin{pmatrix} 1 & & \\ -\theta & 1 & \mathbf{0} \\ & \ddots & \ddots \\ \mathbf{0} & & -\theta & 1 \end{pmatrix} \quad \mathbf{y} \quad \mathbf{B} = \begin{pmatrix} 0 & & & \\ 1 & 0 & \mathbf{0} \\ & \ddots & \ddots \\ \mathbf{0} & & 1 & 0 \end{pmatrix}.$$

La función de autocovarianzas del proceso z_t es

$$\gamma_{z}(k) = \begin{cases} (1+\theta^{2})\sigma_{a}^{2} & k = 0\\ -\theta\sigma_{a}^{2} & |k| = 1\\ 0 & |k| > 1 \end{cases}$$

y la matriz de autocovarianzas de \mathbf{z} , $\mathbf{\Gamma}_z = [\gamma_z(|i-j|)]$, es una matriz tridiagonal con $\gamma_z(0)$ en la diagonal principal y $\gamma_z(1)$ en la primera subdiagonal y superdiagonal. Siguiendo la convención notacial establecida en (2.1), se expresa $\mathbf{\Gamma}_z = \sigma_a^2 \mathbf{\Omega}_z(\theta)$, en donde la matriz de cuasi-covarianzas

$$\mathbf{\Omega}_{\mathbf{z}}(\theta) = \mathbf{\nabla}(\theta)\mathbf{\nabla}(\theta)' = (1+\theta^2)\mathbf{I}_{n-1} - \theta(\mathbf{B}_{n-1} + \mathbf{B}'_{n-1})$$
(3.33)

tiene la estructura

$$\boldsymbol{\Omega}_{\mathbf{z}}(\boldsymbol{\theta}) = \begin{pmatrix} 1 + \theta^2 & -\theta & & \\ -\theta & 1 + \theta^2 & \ddots & \mathbf{0} \\ & \ddots & \ddots & -\theta \\ \mathbf{0} & & -\theta & 1 + \theta^2 \end{pmatrix}.$$

Conviene notar que la forma matricial del modelo MA(1) dada en (3.32) no es un caso especial del modelo clásico generalizado (2.1) porque no incluye regresores. No obstante, la metodología de los contrastes invariantes descrita en el capítulo 2 se aplica igualmente.

Tanaka (1990) consideró el problema estadístico de contrastar la hipótesis nula H_0 : $\theta = 1$ frente a la hipótesis alternativa unilateral H_1 : $\theta < 1$. Sin embargo, dado que las autocovarianzas $\gamma_z(0)$ y $\gamma_z(1)$ estás definidas para cualquier valor de θ , en esta tesis se enfatiza que H_0 se contrasta realmente frente a la alternativa bilaterial

$$H_0: \theta = 1 \quad versus \quad H_1: \theta \neq 1. \tag{3.34}$$

El contraste frente alternativa bilateral se justifica también porque el logaritmo de la función de verosimilitud concentrada de un proceso MA(1)

$$\ell(\theta|\mathbf{z},\tilde{\sigma}_a^2) = -\frac{n-1}{2}\ln\left(\frac{2\pi}{n-1}+1\right) - \frac{1}{2}\ln|\mathbf{\Omega}_{\mathbf{z}}(\theta)| - \frac{n-1}{2}\ln\left(\mathbf{z}'\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\mathbf{z}\right),$$

presenta una forma especial de simetría: $\ell(\theta | \mathbf{z}, \tilde{\sigma}_a^2) = \ell(1/\theta | \mathbf{z}, \tilde{\sigma}_a^2)$ porque $\Omega_{\mathbf{z}}(1/\theta) = (1/\theta^2)\Omega_{\mathbf{z}}(\theta)$. Esta forma de simetría implica que $\ell(\theta | \mathbf{z}, \tilde{\sigma}_a^2)$ debe tener un máximo o un mínimo en $\theta = 1$. Por tanto, parece razonable no rechazar H_0 cuando $\ell(\theta | \mathbf{z}, \tilde{\sigma}_a^2)$ alcanza un máximo en $\theta = 1$, es decir, cuando la segunda derivada de $\ell(\theta | \mathbf{z}, \tilde{\sigma}_a^2)$ evaluada en $\theta = 1$ es negativa. En el capítulo 2, se ha demostrado que el contraste LBIU de una hipótesis nula simple sobre un parámetro de la matriz de covarianzas frente a

una alternativa bilateral se basa precisamente en la segunda derivada de la función de verosimilitud concentrada.

La primera derivada de $\ell(\theta | \mathbf{z}, \tilde{\sigma}_a^2)$ respecto de θ es

$$\frac{d\ell(\theta|\mathbf{z},\tilde{\sigma}_a^2)}{d\theta} = -\frac{1}{2} \operatorname{tr}\left(\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta) \frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)}{d\theta}\right) - \frac{n-1}{2} \frac{\mathbf{z}'\left(\frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)^{-1}}{d\theta}\right)\mathbf{z}}{\mathbf{z}'\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\mathbf{z}},$$

que se anula cuando se evalúa bajo H_0 porque

$$\left. \frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)}{d\theta} \right|_{\theta=1} = \mathbf{\Omega}_{\mathbf{z}}(1)$$

у

$$\frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)^{-1}}{d\theta}\bigg|_{\theta=1} = -\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)}{d\theta}\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\bigg|_{\theta=1} = -\mathbf{\Omega}_{\mathbf{z}}(1)^{-1}.$$

A su vez, la segunda derivada de $\ell(\theta|\mathbf{z},\tilde{\sigma}_a^2)$ respecto de θ

$$\frac{\partial^{2}\ell(\theta|\mathbf{z},\tilde{\sigma}_{a}^{2})}{\partial\theta^{2}} = -\frac{1}{2} \operatorname{tr}\left(\frac{d\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)}{d\theta}\frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)}{d\theta} + \mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\frac{d^{2}\mathbf{\Omega}_{\mathbf{z}}(\theta)}{d\theta^{2}}\right)$$
$$-\frac{n-1}{2}\frac{\mathbf{z}'\left(\frac{d^{2}\mathbf{\Omega}_{\mathbf{z}}(\theta)^{-1}}{d\theta^{2}}\right)\mathbf{z}}{\mathbf{z}'\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\mathbf{z}} + \frac{n-1}{2}\left[\frac{\mathbf{z}'\left(\frac{d\mathbf{\Omega}_{\mathbf{z}}(\theta)^{-1}}{d\theta}\right)\mathbf{z}}{\mathbf{z}'\mathbf{\Omega}_{\mathbf{z}}^{-1}(\theta)\mathbf{z}}\right]^{2}$$

evaluada bajo H_0 conduce a

$$\frac{\partial^2 \ell(\theta | \mathbf{z}, \tilde{\sigma}_a^2)}{\partial \theta^2} \bigg|_{\theta=1} = -\frac{1}{2} \operatorname{tr} \left(-\mathbf{I}_{n-1} + 2\mathbf{\Omega}_{\mathbf{z}}^{-1}(1) \right) \\ -\frac{n-1}{2} \frac{\mathbf{z}' \left(2\mathbf{\Omega}_{\mathbf{z}}^{-1}(1) - 2\mathbf{\Omega}_{\mathbf{z}}^{-2}(1) \right) \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}^{-1}(1) \mathbf{z}} - \frac{n-1}{2} \mathbf{z}' \mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}^{-1}(1) \mathbf{z}}$$

que será positiva cuando la ratio

$$\mathcal{T}_n = \frac{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-2} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-1} \mathbf{z}}$$
(3.35)

tome valores grandes. Así, la región crítica del contraste LBIU de no-invertibilidad en un modelo MA(1) viene dada por $T_n > \kappa$.

Tanaka (1990) notó³ que $\Omega_z(1)^{-1} = \mathbf{H}'\mathbf{H}$, en donde

$$\mathbf{H} = \begin{pmatrix} \frac{1}{\sqrt{2}} & & & \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \mathbf{0} \\ \vdots & \vdots & \ddots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{2}{\sqrt{n(n-1)}} & \cdots & \frac{n-1}{\sqrt{n(n-1)}} \end{pmatrix}.$$

Usando esta descomposición matricial expresó el estadístico T_n como

$$\mathcal{T}_{n} = \frac{\sum_{\tau=1}^{n-1} \left(\sum_{t=1}^{\tau-1} (\tau - t) z_{t} - \frac{\tau}{n} \sum_{t=1}^{n-1} (n - t) z_{t} \right)^{2}}{\sum_{\tau=1}^{n-1} \left(\frac{1}{\tau(\tau+1)} \left(\sum_{t=1}^{\tau} t z_{t} \right) \right)^{2}},$$
(3.36)

en donde se ha adaptado aquí la fórmula original al tamaño muestral n - 1.

3.3.2. La aproximación de Saikkonen-Luukkonen

El modelo IMA(1,1) regular

$$\nabla y_t = (1 - \theta B)a_t, \quad t = 1, \dots, n, \tag{3.37}$$

puede escribirse en forma matricial como

$$\mathbf{D}(1)\mathbf{y} = \mathbf{i}_{(1)}y_0 + \mathbf{D}(\theta)\mathbf{a},\tag{3.38}$$

en donde y_0 es un valor premuestral que se supone no estocástico y el resto de símbolos se han definido previamente. Premultiplicando (3.38) por la inversa **C** de **D**(1), se obtiene que

$$\mathbf{y} = \mathbf{i}y_0 + \mathbf{C}\mathbf{D}(\theta)\mathbf{a},\tag{3.39}$$

que es un caso especial del modelo de regresión lineal generalizado (2.1) con $\mathbf{X} = \mathbf{i}_n$, $\boldsymbol{\beta} = y_0$ y $\mathbf{e} = \mathbf{CD}(\theta)\mathbf{a}$. Consecuentemente, la distribución muestral de y es

$$\mathbf{y} \sim N\left(\mathbf{i}y_0, \sigma_a^2 \mathbf{\Omega}_e(\theta)\right) \quad \text{con} \quad \mathbf{\Omega}_e(\theta) = \mathbf{C}\mathbf{D}(\theta)\mathbf{D}(\theta)'\mathbf{C}'.$$
 (3.40)

³En el apéndice ?? de esta tesis se presenta una demostración de esta descomposición, pág. ??.

Al igual que Tanaka (1990), Saikkonen y Luukkonen (1993) consideraron el contraste de $H_0: \theta = 1$ frente a la alternativa $H_1: |\theta| < 1$. Sin embargo, como se ha justificado en el apartado anterior, H_0 se contrasta realmente frente a la alternativa bilateral H_1 : $\theta \neq 1$. Bajo H_0 , $\Omega_e(\theta) = \mathbf{I}_n$; y de acuerdo con el Teorema 3, el contraste LBIU rechaza H_0 frente a H_1 cuando

$$S\mathcal{L}_n = \frac{\hat{\mathbf{e}}'\mathbf{C}\mathbf{B}\mathbf{B}'\mathbf{C}'\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}} > \kappa, \qquad (3.41)$$

en donde $\hat{\mathbf{e}} = \mathbf{M}\mathbf{y}$ son los residuos minimocuadráticos en la regresión de y sobre una constante, es decir, los datos centrados, $\hat{e}_t = y_t - \bar{y}$. La obtención del estadístico SL_n requiere el cálculo de la segunda derivada

$$\frac{d^2 \mathbf{\Omega}_e(\theta)}{d\theta^2} = \mathbf{C} \frac{d^2 \mathbf{D}(\theta) \mathbf{D}(\theta)'}{d\theta^2} \mathbf{C}',$$

,

en donde

$$\mathbf{D}(\theta)\mathbf{D}(\theta)' = \begin{pmatrix} 1 & -\theta & 0 & \dots & 0 \\ -\theta & 1+\theta^2 & -\theta & & \\ 0 & -\theta & 1+\theta^2 & \ddots & \\ \vdots & & \ddots & \ddots & -\theta \\ 0 & & & -\theta & 1+\theta^2 \end{pmatrix}$$

у

$$\frac{d^2 \left[\mathbf{D}(\theta) \mathbf{D}(\theta)' \right]}{d\theta^2} = 2 \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & \ddots & \\ \vdots & \ddots & \ddots & 0 \\ 0 & & 0 & 1 \end{pmatrix} = 2\mathbf{B}\mathbf{B}'.$$

3.3.3. La equivalencia entre los tres estadísticos

Los estadísticos de Nyblom y Mäkeläinen (1983), Tanaka (1990) y Saikkonen y Luukkonen (1993) pueden escribirse como

$$\mathcal{NM}_n = \frac{\mathbf{y'MCC'My}}{\mathbf{y'My}}, \quad \mathcal{T}_n = \frac{\mathbf{z'}(\boldsymbol{\nabla\nabla'})^{-2}\mathbf{z}}{\mathbf{z'}(\boldsymbol{\nabla\nabla'})^{-1}\mathbf{z}}, \quad \mathcal{SL}_n = \frac{\mathbf{y'MCBB'C'My}}{\mathbf{y'My}}.$$

Para probar la equivalencia entre \mathcal{NM}_n y \mathcal{T}_n se propone aquí una expresión alternativa

para este último considerablemente más simple e intuitiva. El modelo MA(1) bajo H_0 puede escribirse como $\mathbf{z} = \nabla(1)\mathbf{\tilde{a}}$, en donde $\mathbf{\tilde{a}}$ son los denominados residuos exactos, que se definen como la esperanza matemática de **a** condicionada a **z**

$$\tilde{\mathbf{a}} = E(\mathbf{a}|\mathbf{z}) = E(\mathbf{a}\mathbf{z}')E(\mathbf{z}\mathbf{z}')^{-1}\mathbf{z} = \boldsymbol{\nabla}(1)'\boldsymbol{\Omega}_{\mathbf{z}}(1)^{-1}\mathbf{z}.$$

Las identidades matriciales (3.10) y (3.12) propuestas para simplificar la derivación de la distribución límite de \mathcal{NM}_n se usan ahora para expresar el numerador y denominador de \mathcal{T}_n como

$$\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-2} \mathbf{z} = \mathbf{\tilde{a}}' \mathbf{C} \mathbf{C}' \mathbf{\tilde{a}} \quad \mathbf{y} \quad \mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-1} \mathbf{z} = \mathbf{\tilde{a}}' \mathbf{\tilde{a}},$$

que conducen a dos formas equivalentes de T_n

$$\mathcal{T}_{n} = \frac{\sum_{t=1}^{n} \left(\sum_{\tau=t}^{n} \tilde{a}_{\tau}\right)^{2}}{\sum_{t=1}^{n} \tilde{a}_{t}^{2}} = \frac{\sum_{t=1}^{n} \left(\sum_{\tau=1}^{t} \tilde{a}_{\tau}\right)^{2}}{\sum_{t=1}^{n} \tilde{a}_{t}^{2}}.$$
(3.42)

Además, se cumple que $\mathbf{i}' \mathbf{\tilde{a}} = 0$ porque $\nabla \mathbf{i} = \mathbf{0}$. Así, cuando los datos de z_t se obtengan diferenciando una serie y_t , $\mathbf{z} = \nabla \mathbf{y}$, los residuos exactos del modelo MA(1) estrictamente no invertible, $\mathbf{\tilde{a}} = \nabla (1)' \Omega_{\mathbf{z}} (1)^{-1} \mathbf{D} \mathbf{y}$, coincidirán con los residuos minimocuadráticos en la regresión de \mathbf{y} sobre una constante, $\mathbf{\hat{e}} = \mathbf{M} \mathbf{y}$. En otras palabras, el estadístico $\mathcal{N}\mathcal{M}_n$ basado en y_t es equivalente al estadístico \mathcal{T}_n basado en $z_t = y_t - y_{t-1}$. En esta tesis se propone el cálculo de \mathcal{T}_n basado en los residuos exactos del modelo MA(1) estrictamente no invertible, resultado que, como se mostrará más adelante, es especialmente útil para extender el contraste de no invertibilidad al modelo ARMA(p, q + 1).

Para probar la equivalencia entre \mathcal{NM}_n y \mathcal{SL}_n , se utiliza la propiedad de suma residual nula asociada a la estimación minimocuadrática, que impone la siguiente estructura para el vector fila **yMC** = $(0 \sum_{t=2}^{n} \hat{a}_t \dots \hat{a}_n)$, junto con la definición de la matriz de desplazamiento horizontal **B**, que conduce a **yMCB** = $(\sum_{t=2}^{n} \hat{a}_t \dots \hat{a}_n 0)$. Así, **y'MCC'My** = **y'MCBB'C'My** = $\sum_{j=2}^{n} (\sum_{t=j}^{n} \hat{a}_t)^2$.

La equivalencia entre los tres estadísticos es una consecuencia de la equivalencia entre los tres modelos de interés. La denominada forma reducida del modelo estructural de nivel local (3.1) viene dada por

$$z_t = v_t + (1 - B)u_t, \quad t = 2, \dots, n,$$
 (3.43)

en donde $z_t = y_t - y_{t-1}$ es un proceso estacionario. La función de autocovarianzas de z_t sólo tiene dos coeficientes no nulos

$$\gamma_z(k) = \begin{cases} \sigma_u^2 (2 + \sigma_v^2 / \sigma_u^2) & k = 0\\ -\sigma_u^2 & k = 1 \end{cases}$$

y se corresponde con la de un proceso MA(1) con parámetro positivo, $\theta > 0$. De aquí, la ratio señal-ruido $\rho = \sigma_v^2 / \sigma_u^2$ puede expresarse en términos del parámetro MA

$$\rho = \frac{(1-\theta)^2}{\theta}$$

y viceversa

$$\theta = \frac{\rho + 2 + \sqrt{\rho^2 + 4\rho}}{2},$$

en donde se observa que el contraste frente a alternativa unilateral (3.7) en el modelo de nivel local es equivalente al contraste frente alternativa bilateral (3.34) en el modelo MA(1). Para comprender mejor por qué el estadístico \mathcal{NM}_n es LBI, mientras que el estadístico \mathcal{T}_n es LBIU, conviene notar que la función de verosimilitud en ambos modelos es equivalente $L(\rho | \mathbf{z}) = L(\theta | \mathbf{z})$, pero

$$\frac{d^2 L(\rho | \mathbf{z})}{d\theta^2} = \frac{d^2 L(\rho | \mathbf{z})}{d\rho^2} \left(\frac{d\rho}{d\theta}\right)^2 + \frac{d L(\rho | \mathbf{z})}{d\rho} \frac{d^2 \rho}{d\theta^2}$$

у

$$\frac{d^2 L(\rho | \mathbf{z})}{d\theta^2} \Big|_{\theta=1} = 2 \left. \frac{d L(\rho | \mathbf{z})}{d\rho} \right|_{\rho=0},$$

en donde se observa que el contraste de no invertibilidad en el modelo estructural se basa en la primera derivada de la función de verosimilitud; mientras que en el marco ARIMA, en la segunda derivada.

La aproximación de Tanaka puede aplicarse a la forma reducida (3.43) para derivar el estadístico LBI para el contraste de H_0 : $\rho = 0$ frente H_0 : $\rho > 0$. El modelo reducido en

notación matricial es

$$\mathbf{z} = \boldsymbol{\nabla} \mathbf{C} \mathbf{v} + \boldsymbol{\nabla} \mathbf{u},\tag{3.44}$$

en donde $\mathbf{z} = \nabla \mathbf{y} \mathbf{y} \nabla \mathbf{C} \mathbf{v} = (v_2 \dots v_n)'$ son vectores $(n-1) \times 1$. De aquí,

$$\mathbf{z} \sim N\left(0, \sigma_u^2(\boldsymbol{\nabla}\boldsymbol{\nabla}' + \rho \mathbf{I}_{n-1})\right).$$

Se comprueba fácilmente que la primera derivada de la función de versosimilitud concentrada $L(\rho | \mathbf{z}, \tilde{\sigma}_u^2)$ respecto de ρ será negativa bajo H_0 cuando

$$\frac{d\ln[\mathbf{z}'(\boldsymbol{\nabla}\boldsymbol{\nabla}'+\rho\mathbf{I}_{n-1})^{-1}\mathbf{z}]}{d\rho}\Big|_{\rho=0}$$

tome valores grandes, es decir, cuando

$$\mathcal{T}_n = rac{\mathbf{z}'(\mathbf{\nabla}\mathbf{\nabla}')^{-2}\mathbf{z}}{\mathbf{z}'(\mathbf{\nabla}\mathbf{\nabla}')^{-1}\mathbf{z}} > \kappa.$$

3.3.4. Distribución exacta

Establecida la equivalencia entre los tres estadísticos \mathcal{NM}_n , \mathcal{T}_n y \mathcal{SL}_n , la distribución exacta para los dos últimos puede obtenerse de la distribución del primero y de la relación $\rho = (1 - \theta)^2 / \theta$. No obstante, se proporciona una derivación detallada usando el modelo MA(1) como proceso generador de datos, que permitirá posteriormente obtener la distribución límite en términos de movimientos Brownianos.

Los pasos que se siguen en la derivación de la distribución exacta del estadístico T_n son similares a los descritos para NM_n . En primer lugar, se expresa T_n en términos de **a** sustituyendo (3.32) en (3.35)

$$\mathcal{T}_{n} = \frac{\mathbf{a}' \nabla'(\theta) \mathbf{\Omega}_{\mathbf{z}}(1)^{-2} \nabla(\theta) \mathbf{a}}{\mathbf{a}' \nabla(\theta)' \mathbf{\Omega}_{\mathbf{z}}(1)^{-1} \nabla(\theta) \mathbf{a}}.$$
(3.45)

En segundo lugar, se calcula la descomposición espectral de las matrices que aparecen en el numerador y denominador de la ratio. Para ello hay que tener en cuenta que la matriz $\nabla(\theta)' \Omega_z(1)^{-2} \nabla(\theta)$ de orden $n \times n$ tiene un autovalor nulo y n -1 autovalores no nulos que son idénticos a los de la matriz $\nabla(\theta) \nabla(\theta)' \Omega_z(1)^{-2} =$ $\Omega_{\mathbf{z}}(\theta)\Omega_{\mathbf{z}}(1)^{-2}$ de orden $(n-1) \times (n-1)^4$. Análogamente, la matriz del denominador $\nabla(\theta)'\Omega_{\mathbf{z}}(1)^{-1}\nabla(\theta)$ tiene los mismos autovalores no nulos que $\Omega_{\mathbf{z}}(\theta)\Omega_{\mathbf{z}}(1)^{-1}$. Si $\Lambda(\theta) = \operatorname{diag}(\lambda_1(\theta), \dots, \lambda_{n-1}(\theta))$ es la matriz diagonal que contiene los n-1 autovalores de $\Omega_{\mathbf{z}}(\theta)^{-1}$, entonces se tiene que

$$\mathbf{\Omega}_{\mathbf{z}}(\theta) = \mathbf{P}\mathbf{\Lambda}(\theta)^{-1}\mathbf{P}' \quad \mathbf{y} \quad \mathbf{\Omega}_{\mathbf{z}}(\theta)\mathbf{\Omega}_{\mathbf{z}}(1)^{-r} = \mathbf{P}\mathbf{\Lambda}(\theta)^{-1}\mathbf{\Lambda}(1)^{r}\mathbf{P}',$$

en donde *r* es 1 o 2, y **P** es la matriz ortonormal de autovectores que no depende de θ , **P**'**P** = **I**_{*n*-1}. Así, pues, la descomposición espectral para cada matriz de la ratio viene dada por

$$\mathbf{\nabla}'(\theta)\mathbf{\Omega}_{\mathbf{z}}(1)^{-r}\mathbf{\nabla}(\theta) = \sum_{k=1}^{n-1} \frac{\lambda_k^r(1)}{\lambda_k(\theta)} \mathbf{q}_k \mathbf{q}'_k$$

en donde \mathbf{q}_k es el autovector de $\nabla'(\theta) \mathbf{\Omega}_{\mathbf{z}}(1)^{-r} \nabla(\theta)$ asociado al autovalor $\lambda_k^r(1) / \lambda_k(\theta)$. Finalmente, se define la forma cuadrática $\xi_k = \mathbf{a}' \mathbf{q}_k \mathbf{q}'_k \mathbf{a} / \sigma_a^2 \sim \mathrm{iid}\chi_1^2$ que conduce a

$$\mathcal{T}_n = \frac{\sum_{k=1}^{n-1} [\lambda_k(1)^2 / \lambda_k(\theta)] \xi_k}{\sum_{k=1}^m [\lambda_k(1) / \lambda_k(\theta)] \xi_k},$$
(3.46)

en donde los autovalores $\lambda_k(\theta)^{-1} = (1 + \theta^2) - 2\theta \cos(\frac{\pi k}{n})$ de (3.33) pueden escribirse como

$$\lambda_k(\theta)^{-1} = (1-\theta)^2 + \theta \lambda_k(1)^{-1}.$$

Comparando (3.46) con (3.17), se observa que la distribución de T_n es idéntica a la de \mathcal{NM}_n porque

$$\frac{\lambda_k(1)}{\lambda_k(\theta)} = \theta[1 + \rho \lambda_k(1)]. \tag{3.47}$$

Por tanto, como sucede con el estadístico \mathcal{NM}_n , es necesario corregir \mathcal{T}_n por un factor dependiente del tamaño muestral *n* para conseguir que los dos primeros momentos sean finitos en muestras grandes. Se obtiene así la cdf de $\mathcal{T}_n/(n-1)$,

$$\Pr(\mathcal{T}_n/(n-1) < \kappa) = \Pr\left[\sum_{k=1}^{n-1} \frac{\lambda_k(1)}{\lambda_k(\theta)} \left(\frac{\lambda_k(1)}{n-1} - \kappa\right) \xi_k < 0\right], \quad (3.48)$$

que es equivalente a (3.20), cuya evaluación por Imhof (1961) permite obtener los mismos cuadros 3.1 y 3.2 proporcionados para $NM_n/(n-1)$, con una modificación me-

⁴Si $\nabla(\theta)' \Omega_{\mathbf{z}}(1)^{-2} \nabla(\theta) \mathbf{x} = \lambda \mathbf{x}$, entonces $\nabla(\theta) \nabla(\theta)' \Omega_{\mathbf{z}}(1)^{-2} \mathbf{y} = \lambda \mathbf{y} \operatorname{con} \mathbf{y} = \nabla(\theta) \mathbf{x}$.

nor: en el cuadro 3.2 los valores de ρ bajo H_1 corresponden a los valores de $\theta = \{0.975, 0.95, 0.925, 0.9, 0.875, 0.85, 0.825, 0.8\}.$

3.3.5. Distribución límite

Sin citar el trabajo de Nyblom y Mäkeläinen (1983), Tanaka (1990) encontró que la distribución límite de $T_n/(n-1)$ bajo la secuencia de alternativas locales $H_1: \theta = 1 - c/n$ viene dada por

$$\mathcal{T}_n/(n-1) \stackrel{a}{\sim} \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4} \right) \xi_k,$$
 (3.49)

que, como cabía esperar, es idéntica a (3.21). La demostración, por tanto, se basa en los resultados descritos en el apartado (3.2.4) y en (3.47). Recordando que $n^2 \lambda_k(1)^{-1} \rightarrow (\pi k)^2$ cuando $n \rightarrow \infty$, se tiene que

$$\frac{1}{n^2} \sum_{k=1}^{n-1} \left[\frac{\lambda_k(1)^2}{\lambda_k(1-c/n)} \right] \xi_k = \sum_{k=1}^{n-1} \left[\frac{1-c/n}{n^2 \lambda_k(1)^{-1}} + \frac{c^2}{n^4 \lambda_k(1)^{-2}} \right] \xi_k \xrightarrow{d} \sum_{k=1}^{\infty} \left[\frac{1}{\pi^2 k^2} + \frac{c^2}{\pi^4 k^4} \right] \xi_k$$

у

$$\frac{1}{n}\sum_{k=1}^{n-1} \left[\frac{\lambda_k(1)}{\lambda_k(1-c/n)}\right] \xi_k = \frac{1}{n}\sum_{k=1}^{n-1} \left[\frac{n-c}{n} + \frac{c^2}{n^2\lambda_k(1)^{-1}}\right] \xi_k \xrightarrow{p} 1,$$

porque $\sum_{k=1}^{n-1} \xi_k / (n\pi^2 k^2) \xrightarrow{p} 0$ y $\sum_{k=1}^{n-1} \xi_k / n \xrightarrow{p} 1$.

En la extensión estacional del estadístico T_n , Tam y Reinsel (1997) expresaron también la distribución límite en términos de un funcional de movimientos brownianos

$$\mathcal{T}_n/n \xrightarrow{d} \int_0^1 \left[V(r) + c V^*(r) \right]^2 dr, \qquad (3.50)$$

en donde V(r) = W(r) - rW(1) es un puente browniano y $V^*(r) = \int_0^1 W(s)ds - r \int_0^1 W(s)ds$ es un movimiento browniano integrado. Su demostración es larga y complicada, pero puede simplificarse considerablemente usando las identidades matriciales (3.10) y (3.12) y notando que el modelo MA(1) bajo la alternativa local $\theta = 1 - c/n$, $z_t = a_t - a_{t-1} + (c/n)a_{t-1}$, puede escribirse en forma matricial como

$$\mathbf{z} = \boldsymbol{\nabla} [\mathbf{I}_n - \frac{c}{n} \mathbf{C}'] \mathbf{a},$$

en donde $\nabla \mathbf{C} \mathbf{a} = -(a_1 \dots a_{n-1} 0)'$. Así, el denominador de \mathcal{T}_n/n dividido por *n*

$$\frac{1}{n}\mathbf{z}'\mathbf{\Omega}_{\mathbf{z}}(1)^{-1}\mathbf{z} = \frac{1}{n}\mathbf{a}'[\mathbf{I}_n - \frac{c}{n}\mathbf{C}]\mathbf{M}[\mathbf{I}_n - \frac{c}{n}\mathbf{C}']\mathbf{a} \xrightarrow{p} \sigma_a^2$$

porque $E(\mathbf{a}'\mathbf{M}\mathbf{a}) = (n-1)\sigma_a^2$, $E(\mathbf{a}'\mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{a}) = n(n+1)\sigma_a^2/6$ y $E(\mathbf{a}'\mathbf{M}\mathbf{C}'\mathbf{a}) = (n-1)\sigma_a^2/2$ implican que

$$\frac{1}{n}\mathbf{a}'\mathbf{M}\mathbf{a} \xrightarrow{p} \sigma_a^2, \quad \frac{c^2}{n^3}\mathbf{a}'\mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{a} \xrightarrow{p} 0 \quad \mathbf{y} \quad \frac{c}{n^2}\mathbf{a}'\mathbf{M}\mathbf{C}'\mathbf{a} \xrightarrow{p} 0.$$

Del mismo modo, el numerador de T_n/n dividido por *n*

$$\frac{1}{n^2}\mathbf{z}'\mathbf{\Omega}_{\mathbf{z}}(1)^{-2}\mathbf{z} = \frac{1}{n^2}\mathbf{a}'[\mathbf{I}_n - \frac{c}{n}\mathbf{C}]\mathbf{M}\mathbf{C}'\mathbf{C}\mathbf{M}[\mathbf{I}_n - \frac{c}{n}\mathbf{C}']\mathbf{a},$$

en donde el *t*-ésimo elemento del vector $\mathbf{CM}[\mathbf{I}_n - (c/n)\mathbf{C}']\mathbf{a}$ es

$$\sum_{\tau=1}^{t} (a_{\tau} - \bar{a}) - \frac{c}{n} \sum_{\tau=1}^{t} \left(\sum_{i=\tau}^{n} a_{i} - \frac{1}{n} \sum_{h=1}^{n} \sum_{i=h}^{n} a_{i} \right),$$

en donde \bar{a} es la media de los errores del vector **a**. Ahora sólo hay que tener en cuenta las siguientes leyes de convergencia relativas a la suma acumulada parcial de un proceso de ruido blanco (véase, p. ej., Hamilton 1994):

1.
$$n^{-1/2} \sum_{\tau=1}^{t} a_{\tau} \stackrel{d}{\rightarrow} \sigma_{a} W(r),$$

2. $n^{-1/2} t \bar{a} = (t/n) n^{-1/2} \sum_{\tau=1}^{n} a_{\tau} \stackrel{d}{\rightarrow} \sigma_{a} r W(1),$
3. $n^{-3/2} \sum_{\tau=1}^{t} \sum_{i=\tau}^{n} a_{i} \stackrel{d}{\rightarrow} \sigma_{a} \int_{0}^{r} W(s) ds,$
4. $(t/n) n^{-3/2} \sum_{h=1}^{n} \sum_{i=h}^{n} a_{i} \stackrel{d}{\rightarrow} \sigma_{a} r \int_{0}^{1} W(s) ds,$

en donde W(r) es el proceso de Wiener y r es la división entera $[t/n] \in [0, 1]$. Así, pues, el numerador de \mathcal{T}_n/n dividido por n converge en distribución a

$$\sigma_a^2 \int_0^1 \left\{ [W(r) - rW(1)] - c \left[\int_0^r W(s) ds - r \int_0^1 W(1) ds \right] \right\}^2 dr,$$

que al dividirlo por σ_a^2 completa la demostración. No obstante, la siguiente observación es necesaria . El movimiento browniano integrado $V^*(r)$ en (3.50) aparece multiplicado por *c*; mientras que en la demostración aquí presentada, por -c. La equivalencia surge por la independencia de V(r) y $V^*(r)$, $E(\mathbf{a'CMC'CMa}/n^3) \xrightarrow{p} 0$. De aquí, (3.50) puede escribirse como

$$\mathcal{T}_n/n \xrightarrow{d} \int_0^1 \left[V(r)^2 + c^2 V^*(r)^2 \right] dr, \qquad (3.51)$$

cuya equivalencia con (3.49) es ahora más clara.

3.3.6. Inclusión de una deriva

Sin citar a Nyblom (1986), Tam y Reinsel (1998) consideraron el modelo IMA(1,1) con término constante

$$\nabla y_t = \mu + (1 - \theta B)a_t,$$

que es equivalente al modelo de nivel local con deriva (3.25) cuando $0 < \theta \le 1$, y que puede contemplarse como un modelo IMA(2,2) no invertible

$$\nabla^2 y_t = (1-B)(1-\theta B)a_t.$$

Así, el contraste LBIU de H_0 : $\theta = 1$ frente a H_1 : $\theta \neq 1$ plantea la cuestión de si un proceso MA(2) tienen dos raíces unitarias o solamente una, y se basa en $N_n/(n-2) > \kappa$. Usando los resultados de MacNeill (1978), Tam y Reinsel (1998) encontraron que la distribución límite bajo la secuencia de alternativas $\rho = 1 - c/n$ viene dada por

$$N_n/n \xrightarrow{d} \int_0^1 V^2(r) dr,$$

en donde V(r) is un puente browniano generalizado de primer orden

$$V(r) = W(r) - rW(1) + 3r(1-r)\left(W(1) - 2\int_0^1 W(u)du\right), \quad 0 < r < 1.$$

Resulta trivial derivar el contraste de Nyblom y Harvey (2001) como el contraste de la hipótesis de raíz MA unitaria doble frente a la alternativa de invertibilidad en el modelo $\nabla^2 y_t = (1 - \theta B)^2 a_t$.

3.4. El contraste de Nyblom y Mäkeläinen corregido

Nyblom y Mäkeläinen (1983) derivaron el contraste LBI de no invertibilidad en el mar-

co del modelo más simple en el que puede plantearse este problema: el modelo de nivel local o modelo IMA(1,1). El contraste ha sido extendido al modelo IMA(1,1) impuro cuyo término de error presenta correlación serial, o bien de tipo ARMA(p,q), o bien no especificada. En el primer caso se habla de correciones paramétricas; en el segundo, de correcciones no paramétricas.

3.4.1. La corrección no paramétrica de KPSS

En la derivación de la distribución asintótica de \mathcal{NM}_n , Kwiatkowski et al. (1992) relajaron el supuesto $u_t \sim iidN(0, \sigma_u^2)$ para permitir cierta heterogeneidad y correlación serial débil en los errores. Siguiendo a Phillips y Perron (1988), suponen que u_t cumple los siguientes supuestos:

- (i) $E(u_t) = 0 \ \forall t$;
- (ii) $\sup_t E|u_t|^{\beta+\epsilon} < \infty$ para algún $\beta > 2$ y $\epsilon > 0$;
- (iii) $n^{-1} \lim_{n \to \infty} E(u_1 + \dots + u_n)^2 \to \sigma^2 \in (0, \infty);$

(iv) u_t es una secuencia α -mixing con coeficientes α_m tales que $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$.

Bajo los supuestos (i)-(iv), demostraron que el numerador de \mathcal{NM}_n dividido por n^2 tiene la distribución límite

$$\frac{1}{n^2}\sum_{t=1}^T \left(\sum_{\tau=1}^t \hat{u}_t\right)^2 \Rightarrow \sigma \int_0^1 V(r) dr,$$

en donde σ^2 es la denominada varianza a largo plazo definida en (iii) y el símbolo \Rightarrow denota convergencia débil. De aquí concluyen que, cuando los errores no son iid, entonces el denominador de \mathcal{NM}_n dividido por *n* no debe ser un estimador consistente de σ_u^2 , sino de σ^2 , el cual está dado por

$$\hat{\sigma}^{2}(l) = n^{-1} \sum_{t=1}^{n} \hat{u}_{t}^{2} + 2n^{-1} \sum_{s=1}^{l} w(s, l) \sum_{t=s+1}^{n} \hat{u}_{t} \hat{u}_{t-s},$$

en donde w(s, l) es una función kernel y l es el parámetro ancho de banda. Siguiendo a Newey y West (1987), usaron el kernel de Bartlett w(s, l) = 1 - s/(l+1) que garantiza que $\hat{\sigma}^2(l) > 0$. Además, para asegurar su consistencia, es necesario que $l \to \infty$ cuando $n \to \infty$. El estadístico \mathcal{NM}_n/n corregido es

$$\mathcal{KPSS}_n/n = \frac{\hat{\sigma}_u^2}{\hat{\sigma}^2(l)} \mathcal{NM}_n/n = \frac{1}{n^2} \frac{\sum_{t=1}^n (\sum_{\tau=1}^n \hat{u}_{\tau})^2}{\hat{\sigma}^2(l)},$$

en donde $\hat{u}_t = (y_t - \bar{y}) \text{ y } \hat{\sigma}_u^2 = \sum_{t=1}^n \hat{u}_t^2 / n$. Similarmente, el estadístico \mathcal{N}_n / n corregido

$$\mathcal{KPSS}_n^{(t)}/n = \frac{\hat{\sigma}_u^2}{\hat{\sigma}^2(l)} \mathcal{N}_n/n = \frac{1}{n^2} \frac{\sum_{t=1}^n (\sum_{\tau=1}^n \hat{u}_{\tau})^2}{\hat{\sigma}^2(l)},$$

en donde el superíndice (t) indica que \hat{u}_t son los residuos en la regresión y_t sobre una tendencia lineal.

3.4.2. La corrección no paramétrica de Tanaka

Tanaka (1990) modificó el estadístico T_n (e implicitamente \mathcal{NM}_n) para contrastar no invertibilidad en un modelo MA(1) impuro

$$z_t = (1 - \theta B)b_t, \quad t = 2, \dots, n,$$

$$b_t = \psi(B)a_t, \qquad (3.52)$$

en donde se supone que b_t es un proceso lineal general estacionario e invertible con

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j, \quad \psi_0 = 1, \quad \sum_{j=0}^{\infty} j |\psi_j| < \infty, \quad y \quad \psi(1) = \sum_{j=0}^{\infty} \psi_j \neq 0.$$
(3.53)

De acuerdo con el teorema 3 de Tanaka (1990), en el modelo MA(1) impuro y bajo la secuencia de alternativas H_1 : $\theta = 1 - c/n$, el estadístico corregido $\varphi \mathcal{N} \mathcal{M}_n/n$ sigue asintóticamente la distribución límite dada en (3.50), siendo el factor de corrección

$$arphi = rac{\sum_{j=0}^{\infty}\psi_j^2}{\left(\sum_{j=0}^{\infty}\psi_j
ight)^2}.$$

La demostración se basa en una descomposición tipo Beveridge y Nelson (1981) de la suma acumulada parcial de los errores b_t^5

$$\sum_{t=1}^{n} b_t = \sum_{t=1}^{n} \psi(B) a_t = \sum_{t=1}^{n} \sum_{j=0}^{\infty} \psi_j B^j a_t = \sum_{j=0}^{\infty} \psi_j \sum_{t=1}^{n} a_{t-j},$$

⁵Una demostración similar puede encontrarse en el apéndice 17.A de Hamilton (1994, p. 534).

en donde

$$\sum_{t=1}^{n} a_{t-j} = \sum_{t=1}^{n} a_t + \sum_{t=1}^{j} (a_{t-j} - a_{n+t-j}).$$

Así,

$$\sum_{t=1}^{n} b_t = \psi(1) \sum_{t=1}^{n} a_t + \sum_{j=1}^{\infty} \psi_j^*(a_{1-j} - a_{n+1-j})$$

en donde los pesos $\psi_i^* = \psi_j + \psi_{j+1} + \dots$ son una serie *absolutamente* convergente

$$\sum_{j=1}^{\infty} |\psi_j^*| = \sum_{j=1}^{\infty} j |\psi_j| < \infty$$

por los supuestos (3.53). La principal implicación es

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\infty}\psi_j^*(a_{1-j}-a_{n+1-j})\stackrel{p}{\to}0,$$

de manera que

$$\frac{1}{\sqrt{n}}\sum_{\tau=1}^t b_\tau \xrightarrow{d} \frac{1}{\sqrt{n}}\psi(1)\sum_{\tau=1}^t a_t.$$

El modelo MA(1) impuro bajo la alternativa local $\theta = 1 - c/n$ puede escribirse en forma matricial como

$$\mathbf{z} = \boldsymbol{\nabla}(1) [\mathbf{I}_n - \frac{c}{n} \mathbf{C}'] \mathbf{b},$$

en donde $\mathbf{b} = (b_1 \dots b_n)$. De aquí, el estadístico de contraste

$$\frac{\mathcal{N}\mathcal{M}_n}{n} = \frac{1}{n} \frac{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-2} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)_1^{-1} \mathbf{z}} = \frac{\frac{1}{n^2} \mathbf{b}' [\mathbf{I}_n - \frac{c}{n} \mathbf{C}] \mathbf{M} \mathbf{C}' \mathbf{C} \mathbf{M} [\mathbf{I}_n - \frac{c}{n} \mathbf{C}'] \mathbf{b}}{\frac{1}{n} \mathbf{b}' [\mathbf{I}_n - \frac{c}{n} \mathbf{C}] \mathbf{M} [\mathbf{I}_n - \frac{c}{n} \mathbf{C}'] \mathbf{b}}.$$
(3.54)

El denominador de (3.54) converge en probabilidad a la varianza de b_t

$$\frac{1}{n}\mathbf{b'Mb} = \frac{1}{n}\sum_{t=1}^{n}(b_t - \bar{b})^2 \xrightarrow{p} \sigma_b^2$$

que, por la condición de estacionariedad del proceso lineal general, viene dada por

$$\sigma_b^2 = \sigma_a^2 \sum_{j=1}^\infty \psi_j^2.$$

Para estudiar la convergencia en distribución del numerador de (3.54), hay que notar que se cumplen las siguientes leyes de convergencia:

1.
$$n^{-1/2} \sum_{\tau=1}^{t} b_{\tau} \stackrel{d}{\rightarrow} \sigma_a \psi(1) W(r),$$

2. $n^{-1/2} t \overline{b}_j = (t/n) n^{-1/2} \sum_{\tau=1}^{n} b_{\tau} \stackrel{d}{\rightarrow} \sigma_a r \psi(1) W(1),$
3. $n^{-3/2} \sum_{\tau=1}^{t} \sum_{i=\tau}^{n} b_{\tau} \stackrel{d}{\rightarrow} \sigma_a \psi(1) \int_0^r W(s) ds,$
4. $(t/n) n^{-3/2} \sum_{h=1}^{n} \sum_{i=h}^{n} b_{\tau} \stackrel{d}{\rightarrow} \sigma_a r \psi(1) \int_0^1 W(s) ds.$

Es claro ahora que, en el modelo MA(1) impuro bajo la secuencia de alternativas H_1 : $\theta = 1 - c/n$,

$$\varphi \frac{\mathcal{N}\mathcal{M}_n}{n} \stackrel{d}{\to} \int_0^1 \left(V(r)^2 + c^2 V^*(r)^2 \right) dr.$$
(3.55)

Para el proceso estacionario b_t se cumple que

$$2\pi f_b(0) = \sigma_a^2 \left(\sum_{k=0}^{\infty} \psi_k\right)^2 = \gamma_b(0) + 2\sum_{k=1}^{\infty} \gamma_b(k) \quad \text{y} \quad \gamma_b(k) = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k},$$

en donde $\gamma_b(k) = E(b_t b_{t-k})$ es la autocovarianza de b_t en el retardo k y $f_b(0)$ es la función de densidad espectral de b_t evaluada en 0. Tanaka (1990) propuso el siguiente estimador consistente de φ

$$\hat{\varphi} = \frac{\hat{\gamma}_b(0)}{\hat{\sigma}_b^2(l)}$$

con

$$\hat{\sigma}_b^2(l) = 2\pi \hat{f}_b(0) = \hat{\gamma}_u(0) + 2\sum_{k=1}^l (1 - \frac{k}{l+1})\hat{\gamma}_b(k) \quad \text{y} \quad \hat{\gamma}_b(k) = \frac{1}{n}\sum_{t=k+1} \hat{b}_t \hat{b}_{t-k},$$

en donde

$$\hat{b}_t = \sum_{k=1}^t \frac{k}{\sqrt{t(t+1)}} z_t$$

es el *t*-ésimo elemento del vector columna $\hat{\mathbf{b}} = \mathbf{H}\mathbf{z}$. Tanaka (1990) justificó el uso de \hat{b}_t porque los denominados *residuos condicionados* $\hat{b}_t^{(0)} = z_t + \hat{b}_{t-1}^{(0)}$ obtenidos suponiendo que $\hat{b}_1^{(0)} = 0$ son inconsistentes. Sin embargo, aquí se propone el uso de los *residuos exactos* $\tilde{\mathbf{b}} = \nabla(1)\mathbf{\Omega}_{\mathbf{z}}(1)^{-1}\mathbf{z}$ para estimar las autocovarianzas $\gamma_b(k)$. Se comprueba fácilmente que estos residuos, $\tilde{b}_t = b_t - \bar{b}$, son consistentes y pueden calcularse recursivamente de la relación

$$\tilde{b}_t = z_t + \tilde{b}_{t-1}, \quad t = 2, \dots, n,$$

usando como condición inicial

$$\tilde{b}_1 = -\frac{\sum_{t=2}^n (n-t) z_t}{n}.$$

Dada la equivalencia entre $T_n = NM_n$, la notación T_n se usa, en adelante, para el estadístico NM_n corregido con un factor no paramétrico

$$\mathcal{T}_n/n = \hat{arphi} rac{\mathcal{N}\mathcal{M}_n}{n} = rac{1}{n^2} rac{\sum_{t=1}^n \left(\sum_{ au=1}^t ilde{b}_t
ight)^2}{\hat{\sigma}_b^2(l)}.$$

Cuando $\mathbf{z} = \nabla \mathbf{y}$, entonces $\mathbf{\tilde{b}} = \nabla(1)\mathbf{\Omega}_{\mathbf{z}}(1)^{-1}\nabla \mathbf{y} = \mathbf{M}\mathbf{y}$ y el estadístico \mathcal{T}_n/n coincide con \mathcal{KPSS}_n/n . La especificación de una estructura ARMA(p,q) conocida para el error b_t tiene la ventaja de permitir tabular la distribución exacta del estadístico \mathcal{T}_n en muestras finitas y evaluar la medida en que tal distribución se aproxima a la obtenida en el caso simple.

3.4.3. La corrección paramétrica de Leybourne y McCabe

Leybourne y McCabe (1994) extendieron el contraste de Nyblom y Mäkeläinen (1983) al modelo de nivel local (3.1) generalizado

$$\phi(B)y_t = \mu_t + u_t,$$

 $\mu_t = \mu_{t-1} + v_t, \quad t = 1, \dots, T,$
(3.56)

en donde $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ es un polinomio AR de orden p estacionario. Su procedimiento para contrastar (3.7) en este marco se basa en tratar la serie filtrada $y_t^* = \phi(B)y_t$ como un proceso de nivel local. En una aplicación práctica, sin embargo, los coeficientes ϕ_i ($i = 1, \dots, p$) son desconocidos y deben estimarse de forma consistente. Citando a Pötscher (1991), advierten de que la aproximación tentativa de estimar el modelo bajo H_0 ,

$$y_t = \mu + \phi_1 y_{t-1} + \cdots + \phi_{t-p} + u_t,$$

conduce a estimaciones inconsistentes cuando H_1 es la hipótesis verdadera, porque el término de error u_t sería un paseo aleatorio correlacionado con los retardos de y_t . En cambio, la estimación por máxima verosimilitud exacta del modelo sin restringir produce estimaciones $\tilde{\phi}_i$ (i = 1, ..., p) consistentes de los parámetros autorregresivos. En lugar de estimar la forma estructural (3.56), proponen estimar su forma reducida ARIMA(p, 1, 1). Su procedimiento de contraste puede resumirse como sigue:

1. se estima por máxima verosimilitud exacta el modelo ARIMA(p, 1, 1) equivalente

$$\phi(B)\nabla y_t = (1-\theta B)a_t,$$

2. se construye la serie filtrada

$$ilde{y}_t^* = y_t - \sum_{j=1}^p ilde{\phi}_j y_{t-j}, \quad t = 1, \ldots, n,$$

3. se genera la serie filtrada centrada

$$ilde{y}_t^* = ilde{y}_t^* - ilde{y}^*, \quad t = 1, \ldots, n_t$$

4. y, finalmente, se calcula el estadístico \mathcal{NM}_n corregido

$$\mathcal{LM}_n/(n-1) = \frac{1}{n-1} \frac{\sum_{t=1}^n \left(\sum_{\tau=1}^t \tilde{y}_t^*\right)^2}{\sum_{t=1}^n \tilde{y}_t^{*^2}},$$

que, bajo la secuencia de alternativas $\rho = c^2/n$, sigue la distribución límite (3.21).

Para probar que la corrección paramétrica no cambia distribución límite del estadístico, derivaron, en primer lugar, la distribución de la suma acumula parcial

$$\frac{1}{\sqrt{n}}\sum_{\tau=1}^t \tilde{\tilde{y}}_t^*,$$

teniendo en cuenta que la serie filtrada centrada $\tilde{\tilde{y}}_t^*$ puede expresarse como

$$\tilde{y}_t^* - \tilde{y}^* = (\phi_1 - \tilde{\phi}_1)(y_{t-1} - \bar{y}_{-1}) + \dots + (\phi_p - \tilde{\phi}_p)(y_{t-p} - \bar{y}_{-p}) + (\mu_t - \bar{\mu}) + (u_t - \bar{u}).$$

Se aplican los siguientes resultados que ya han aparecido en los apartados anteriores

1.
$$n^{-1/2} \sum_{\tau=1}^{t} (u_{\tau} - \bar{u}) \stackrel{d}{\to} \sigma_{u}^{2} V(r),$$

2. $n^{-1/2} \sum_{\tau=1}^{t} (\mu_{\tau} - \bar{\mu}) \stackrel{d}{\to} c^{2} \sigma_{u}^{2} r V(1),$
3. $n^{-1/2} \sum_{\tau=1}^{t} (y_{\tau-j} - \bar{y}_{-j}) \stackrel{d}{\to} \sigma_{u}^{2} \phi(1)^{-1} [V(r) + c^{2} r V(1)],$
4. $(\phi_{j} - \tilde{\phi}_{j}) n^{-1/2} \sum_{\tau=1}^{t} (y_{\tau-j} - \bar{y}_{-j}) \stackrel{p}{\to} 0,$

en donde el último resultado sigue de la consistencia de las estimaciones. Realmente, la demostración anterior parece innecesaria en tanto en cuanto la consistencia de las estimaciones asegura que la serie filtrada y_t^* converge a un proceso de nivel local.

Hay tres inconvenientes en esta aproximación. En primer lugar, la ampliación del modelo con un polinomio AR(p) en lugar de una estructura ARMA(p,q) puede conducir a un orden p muy grande, sobre todo cuando el modelo subyacente tiene una estructura MA adicional casi no invertible. En segundo lugar, y relacionado con el anterior, se plantea el problema de cómo elegir p. Finalmente, el filtrado de la serie y_t por el polinomio AR(p) ignora p condiciones iniciales.

3.4.4. La corrección paramétrica de Saikkonen y Luukkonen

Saikkonen y Luukkonen (1993) extendieron el contraste LBIU de noinvertibilidad al modelo IMA(1,1) impuro

$$\nabla y_t = (1 - \theta B)b_t, \quad t = 1, \dots, n,$$

$$\phi(B)b_t = \theta(B)a_t, \tag{3.57}$$

en donde los polinomios $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ y $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ no comparten raíces comunes y cumplen las condiciones de estacionariedad e invertibilidad, respectivamente. Siguiendo los pasos descritos para el modelo puro (3.37), se llega a la forma matricial de (3.57)

$$\mathbf{y} = \mathbf{i}\mathbf{y}_0 + \mathbf{e}, \quad \mathbf{e} = \mathbf{C}\mathbf{D}(\theta)\mathbf{b}, \tag{3.58}$$

en donde $\mathbf{b} = (b_1 \ b_2 \ \dots \ b_n)' \sim N(\mathbf{0}, \mathbf{\Gamma}_b), \mathbf{\Gamma}_b = [\gamma_b(|i-j|)]$ y cada autocovarianza $\gamma_b(k) = E(b_t b_{t-k})$ es una función conocida de σ_a^2 y del cojunto de parámetros $\boldsymbol{\varphi} =$
$\{\phi_1, \ldots, \phi_p, \vartheta_1, \ldots, \vartheta_q\}$. Así, y usando la notación convencional $\Gamma_b = \sigma_a^2 \Omega_b$, se tiene que la distribución muestral de **y** es ahora

$$\mathbf{y} \sim N\left(\mathbf{i}y_0, \sigma_a^2 \mathbf{\Omega}_e(\theta)\right) \quad \text{con} \quad \mathbf{\Omega}_e(\theta) = \mathbf{C}\mathbf{D}(\theta)\mathbf{\Omega}_b\mathbf{D}(\theta)'\mathbf{C}',$$
 (3.59)

en donde $\mathbf{\Omega}_b = [\sigma_a^{-2} \gamma_b (|i-j|)]$ sólo depende de $oldsymbol{arphi}$.

Por el teorema 3, el contraste LBIU para elegir entre H_0 : $\theta = 1$ frente a la alternativa bilateral H_1 : $\theta \neq 1$ bajo (3.59) viene dado por

$$\mathcal{SL}_{n}^{*} = \frac{\tilde{\mathbf{e}}_{*}^{\prime} \mathbf{C}_{*} \mathbf{B} \Omega_{b} \mathbf{B}^{\prime} \mathbf{C}_{*}^{\prime} \tilde{\mathbf{e}}_{*}}{\tilde{\mathbf{e}}_{*}^{\prime} \tilde{\mathbf{e}}_{*}} > \kappa, \qquad (3.60)$$

en donde $\tilde{\mathbf{e}}_* = \mathbf{\Omega}_b^{-1/2} \tilde{\mathbf{e}}, \mathbf{C}'_* = \mathbf{\Omega}_b^{-1/2} \mathbf{C}$ y $\tilde{\mathbf{e}}$ son los residuos de la estimación MCG de (3.58). Obviamente, cuando $p = q = 0, \mathbf{\Omega}_b = \mathbf{I}_n$ y \mathcal{SL}_n^* coincide con el estadístico sin corregir \mathcal{SL}_n .

La obtención de SL_n^* requiere el cálculo de la segunda derivada

$$\frac{d^2 \mathbf{\Omega}_e(\theta)}{d\theta^2} = \mathbf{C} \frac{d^2 \mathbf{D}(\theta) \mathbf{\Omega}_b \mathbf{D}(\theta)'}{d\theta^2} \mathbf{C}'$$

en donde la matriz cuadrada $\mathbf{D}(\theta)\mathbf{\Omega}_{b}\mathbf{D}(\theta)'$ puede particionarse como

$$\sigma_a^{-2} \left(\begin{array}{c|c} \gamma_b(0) & [\gamma_b(i) - \theta \gamma_b(i-1)] \\ \hline [\gamma_b(i) - \theta \gamma_b(i-1)] & [(1+\theta^2)\gamma_b(|i-j|) - \theta(\gamma_b(|i-j|-1) + \gamma_b(|i-j|-1)] \end{array} \right),$$

siendo $[\gamma_b(i) - \theta \gamma_b(i-1)]$ un vector columna de orden n-1 con elemento genérico $\gamma_b(i) - \theta \gamma_b(i-1)$ para i = 1, ..., n-1 y siendo $[(1+\theta^2)\gamma_b(|i-j|) - \theta(\gamma_b(|i-j| - 1) + \gamma_b(|i-j| - 1))]$ una matriz cuadrada de orden n-1 con elemento genérico $(1 + \theta^2)\gamma_b(|i-j|) - \theta(\gamma_b(|i-j| - 1) + \gamma_b(|i-j| - 1))$ para i, j = 1, ..., n-1. Así,

$$\frac{d^2 \left[\mathbf{D}(\theta) \mathbf{\Omega}_b \mathbf{D}(\theta)'\right]}{d\theta^2} = 2\sigma_a^{-2} \left(\frac{0 \quad \mathbf{0}'}{\mathbf{0} \mid \left[\gamma_b(|i-j|)\right]} \right) = 2\mathbf{B}\mathbf{\Omega}_b \mathbf{B}'.$$

Saikkonen y Luukkonen (1993) demostraron que, bajo H_0 , SL_n^*/n sigue la distribución límite derivada para SL_n/n en el modelo IMA(1,1) puro. Su demostración es tediosa y

es objeto de simplificación en la sección siguiente. No obstante, se destacan a continuación los principales resultados asintóticos bajo los que obtienen la distribución límite.

- Bajo los supuestos dados en (3.57), los valores premuestrales b₀, b₋₁, ..., b_{-p+1}, a₀, a₋₁, ..., a_{-q+1} pueden ignorarse en el cálculo del estadístico porque no afectarán a b_t cuando t → ∞. Con valores premuestrales nulos, el proceso b_t puede escribirse en notación matricial como Φb = Θa, en donde Φ = I_n φ₁B ... φ_pB^p, Θ = I_n θ₁B ... θ_qB^q y B^k es una matriz nula con unos en la subdiagonal k > 0. La matriz de cuasicovarianzas Ω_b puede recemplazarse por ΨΨ', en donde Ψ = Φ⁻¹Θ = I_n + ψ₁B + ... + ψ_{n-1}Bⁿ⁻¹ y los pesos ψ_j se obtienen de la relación φ(B)ψ(B) = θ(B). Claramente, Ω_b⁻¹ = ΠΠ', en donde Π = Θ⁻¹Φ = I_n + π₁B + ... + π_{n-1}Bⁿ⁻¹ y los pesos π_j se obtienen de la relación φ(B)π(B) = φ(B). Además, se cumple que Ω_b^{1/2} = Ψ y Ω_b^{-1/2} = Π.
- 2. Las matrices Π y **CB** conmutan por tener estructura Toeplitz triangular inferior.
- Asintóticamente, los residuos MCG, ẽ, y MCO, ê coinciden, por la consistencia del estimador MCO de y₀.
- 4. La descomposición Beveridge y Nelson (1981) de $\pi(B)$ es $\pi(B) = \pi(1) + (1 B)\pi^*(B)$, en donde $\pi_*(B) = \sum_{i=0}^{\infty} (\sum_{j=i+1}^{\infty} \pi_j^*)B^i$. De aquí, $\Pi = \pi(1)\mathbf{I} + \mathbf{D}\Pi^*$.
- 5. $\hat{\mathbf{e}}_* = \mathbf{\Pi}\hat{\mathbf{e}} = \pi(1)\hat{\mathbf{e}} + \mathbf{D}\mathbf{\Pi}_*\hat{\mathbf{e}}$ y $\hat{\mathbf{e}}\mathbf{\Pi}'_*\mathbf{D}'\mathbf{D}\mathbf{\Pi}_*\hat{\mathbf{e}}/n^2 \xrightarrow{p} 0$ porque $\sum_{k=0}^{\infty} |\pi_k| < \infty$.
- 6. $\mathcal{SL}_n^*/n = \pi(1)^2 \hat{\mathbf{e}}' \mathbf{CC}' \hat{\mathbf{e}}/n^2 \sigma_a^2$, porque el denominador $\tilde{\mathbf{e}}'_* \tilde{\mathbf{e}}_*/n \xrightarrow{p} \sigma_a^2$.
- 7. $S\mathcal{L}_{n}^{*}/n = \pi(1)^{2}\mathbf{b}'\mathbf{M}'\mathbf{C}\mathbf{C}'\mathbf{M}\mathbf{b}/n^{2}\sigma_{a}^{2}$ y, por Tanaka (1990),

$$\mathbf{b'M'CC'Mb}/n^2\sigma_a^2 \xrightarrow{d} \psi(1)^2 \sum_{k=1}^{\infty} \frac{1}{\pi^2 k^2} \xi_k, \quad \xi_k \sim \chi_1^2.$$

En una aplicación práctica, la matriz de cuasicovarianzas $\Omega_b(\boldsymbol{\varphi})$ es desconocida y debe reemplazarse por una estimación consistente $\tilde{\Omega}_b(\tilde{\boldsymbol{\varphi}})$, en donde $\tilde{\boldsymbol{\varphi}}$ es un estimador consistente de $\boldsymbol{\varphi}$. Saikkonen y Luukkonen (1993) proponen usar estimaciones de máxima verosimilitud y, como Leybourne y McCabe (1994), advierten de que, cuando p > 0, la estimación del modelo bajo H_0 , cuando H_1 es cierta, produce estimaciones inconsistentes; mientras que en el caso dual (estimación del modelo bajo H_1 cuando H_0 es cierta), las estimaciones son consistentes. En cambio, cuando p = 0, señalan que las estimaciones son consistentes en cualquier caso, y sugieren estimar el modelo bajo H_0 ,

$$y_t = \mu + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q},$$

porque la estimación de μ coincidiría con la de y_0 en el modelo de regresión generalizado y la estimación de σ_a^2 proporcionaría el denominador del estadístico. Aquí se advierte de que, bajo H_1 , las estimaciones de los parámetros MA de este modelo serían inconsistentes porque a_t seguiría un proceso IMA(1,1) y estaría correlacionado con sus retardos. Obviando este tema, el cálculo del estadístico requiere invertir la matriz $\tilde{\Omega}_b$ que sugieren calcular usando los algoritmos de estimación de máxima verosimilitud de Dent (1977) y Ansley (1979).

3.4.5. La corrección paramétrica de Tam y Reinsel

En la extensión estacional de los contrastes de Tanaka (1990) y Saikkonen y Luukkonen (1993), Tam y Reinsel (1997) propusieron una corrección paramétrica similar a la de Leybourne y McCabe (1994), pero diseñada para el proceso ARIMA(p, 1, q + 1) que, para facilitar los desarrollos posteriores, se escribe aquí como

$$\phi(B)\nabla y_t = \theta(B)b_t, \quad t = 1, \dots, n,$$

$$b_t = (1 - \theta B)a_t,$$
(3.61)

en donde los polinomios $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ y $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ no comparten raíces comunes y cumplen las condiciones de estacionariedad e invertibilidad, respectivamente. El modelo (3.61) puede escribirse en forma matricial como

$$\Phi \mathbf{z} = \Theta \mathbf{b} + \Phi_0 \mathbf{z}_0 - \Theta_0 \mathbf{b}_0,$$
$$\mathbf{b} = \nabla_{n-1,n}(\theta) \mathbf{a},$$

en donde $\mathbf{z} = (z_2 \ z_3 \ \dots \ z_n)', z_t = y_t - y_{t-1}, \mathbf{b} = (b_2 \ b_3 \ \dots \ b_n)', \mathbf{z}_0 = (z_{-p+2} \ \dots \ z_0 \ z_1)',$ $\mathbf{b}_0 = (b_{-q+2} \ \dots \ b_0 \ b_1)' \ \mathbf{y} \ \mathbf{a} = (a_1 \ a_2 \ \dots \ a_n)'.$ Además, $\mathbf{b}_0 = \nabla_{q,q+1}(\theta)\mathbf{a}_0$ con $\mathbf{a}_0 = (a_{-q+1} \ \dots \ a_0 \ a_1)'$ Las matrices $\mathbf{\Phi} \ \mathbf{y} \ \mathbf{\Theta}$ son cuadradas de orden n-1 y tienen una estructura especial, que simplifica el cálculo de sus inversas: son matrices trianguales inferiores, con diagonal principal unitaria y subdiagonales escalares formadas por $-\phi_j (j = 1, ..., p)$ o $-\theta_j (j = 1, ..., q + 1)$, respectivamente. Usando la matriz retardo **B**, y notando que **B**^{*j*} tiene unos en la *j*-ésima subdiagonal y ceros en el resto de subdiagonales, se cumple que $\mathbf{\Phi} = \mathbf{I}_{n-1} - \phi_1 \mathbf{B} - \cdots - \phi_p \mathbf{B}^p$ y $\mathbf{\Theta} = \mathbf{I}_{n-1} - \theta_1 \mathbf{B} - \cdots - \theta_q \mathbf{B}^q$. Las matrices $\mathbf{\Phi}_0$ y $\mathbf{\Theta}_0$ asociadas a los vectores de valores premuestrales \mathbf{z}_0 y \mathbf{b}_0 son de orden $n \times p$ y $n \times q$, respectivamente, y tienen la siguiente estructura

$$\boldsymbol{\Phi}_{0} = \begin{pmatrix} \phi_{p} & \dots & \phi_{1} \\ & \ddots & \vdots \\ 0 & & \phi_{p} \\ \hline & 0 \end{pmatrix} \quad \mathbf{y} \quad \boldsymbol{\Theta}_{0} = \begin{pmatrix} \theta_{q} & \dots & \theta_{1} \\ & \ddots & \vdots \\ 0 & & \theta_{q} \\ \hline & 0 \end{pmatrix},$$

en donde se observa que el bloque no nulo es una matriz triangular superior. Como el vector $\mathbf{b} = \mathbf{\Theta}^{-1}[\mathbf{\Phi}\mathbf{z} + \mathbf{\Phi}_0\mathbf{z}_0 - \mathbf{\Theta}_0\nabla(\theta)\mathbf{a}_0]$ es un proceso MA(1), Tam y Reinsel (1997) propusieron dos modificaciones del estadístico de Tanaka (1990)

$$\mathcal{T}_n = \frac{1}{n-1} \frac{\mathbf{z}'(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-2}\mathbf{z}}{\mathbf{z}'(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-1}\mathbf{z}},$$

dadas por

$$\mathcal{TR}_n^* = \frac{1}{n-1} \frac{\tilde{\mathbf{z}}'_* (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-2} \tilde{\mathbf{z}}_*}{\tilde{\mathbf{z}}'_* (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-1} \tilde{\mathbf{z}}_*} \quad \text{y} \quad \mathcal{TR}_n^{**} = \frac{1}{n-1} \frac{\tilde{\mathbf{z}}'_{**} (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-2} \tilde{\mathbf{z}}_{**}}{\tilde{\mathbf{z}}'_{**} (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-1} \tilde{\mathbf{z}}_{**}},$$

en donde $\tilde{\mathbf{z}}_* = \tilde{\boldsymbol{\Theta}}^{-1} \tilde{\boldsymbol{\Phi}} \mathbf{z}$ y $\tilde{\mathbf{z}}_{**} = \tilde{\boldsymbol{\Theta}}^{-1} [\tilde{\boldsymbol{\Phi}} \mathbf{z} + \tilde{\boldsymbol{\Phi}}_0 \tilde{\mathbf{z}}_0 - \tilde{\boldsymbol{\Theta}}_0 \boldsymbol{\nabla}(1) \tilde{\mathbf{a}}_0]$ son las correcciones propuestas para el vector de observaciones estacionarias basadas en la expresión anterior de **b** y en estimaciones consistentes de los parámetros desconocidos.

Varios comentarios son oportunos. En primer lugar, el estadístico \mathcal{TR}_n^* es más simple de calcular que \mathcal{TR}_n^{**} porque no depende de los valores premuestales \mathbf{z}_0 y \mathbf{a}_0 . En segundo lugar, los dos estadísticos \mathcal{TR}_n^* y \mathcal{TR}_n^{**} requieren la estimación previa y sin restricciones del modelo ARIMA(p, 1, q + 1) por un método que proporciones estimaciones consistentes, por ejemplo, máxima verosimilitud exacta. En tercer lugar, \mathcal{TR}_n^* es equivalente a \mathcal{LM}_n^* cuando la correlación serial sea de tipo puramente AR, q = 0, y el primero puede entenderse como una mejora del segundo cuando q > 0. En cuarto lugar, una vez estimado libremente el modelo, los valores premuestrales $\tilde{\mathbf{z}}_0$ y $\tilde{\mathbf{a}}_0$ se obtienen fijando θ en 1. Finalmente, es claro que en muestras grandes, y siempre y cuando el proceso b_t sea estrictamente invertible, el efecto de las condiciones iniciales se irá anulando a largo plazo y los dos estadísticos serán asintóticamente iguales. En este caso, los dos procesos filtrados $\tilde{\mathbf{z}}_*$ y $\tilde{\mathbf{z}}_{**}$ convergerán a un proceso MA(1) y su distribución asintótica común coincidirá con la de \mathcal{NM}_n .

Por tanto, el procedimiento de Tam y Reinsel (1997) para contrastar H_0 : $\theta = 1$ frente a H_1 : $\theta \neq 1$ en el modelo (3.61) es el siguiente:

 se estima por máxima verosimilitud exacta el modelo ARIMA(*p*, 1, *q* + 1) definido en (3.61)

$$\phi(B)\nabla y_t = \theta(B)(1-\theta B)a_t,$$

en donde la raíz MA real más próxima a la unidad se considera como $\tilde{\theta}$;

- 2. se fija $\tilde{\theta}$ en 1, y se obtienen los valores premuestrales⁶ \tilde{z}_0 y \tilde{a}_0 ;
- 3. se construye la serie filtrada $\tilde{\mathbf{z}}_{**} = \tilde{\boldsymbol{\Theta}}^{-1}[\tilde{\boldsymbol{\Phi}}\mathbf{z} + \tilde{\boldsymbol{\Phi}}_0 \tilde{\mathbf{z}}_0 \tilde{\boldsymbol{\Theta}}_0 \boldsymbol{\nabla}(1) \tilde{\mathbf{a}}_0]$ usando esta fórmula matricial o la ecuación en diferencias

$$z_{t}^{**} = \tilde{\theta}_{1} z_{t-1}^{**} + \dots + \tilde{\theta}_{1} z_{t-1}^{**} + z_{t} - \tilde{\phi}_{1} z_{t-1} - \dots - \tilde{\phi}_{p} z_{t-p} + \tilde{u}_{t}$$

en donde los valores iniciales de z_t^{**} se fijan en cero, u_t es cero para t > r = máx(p,q+1) y los primeros r = máx(p,q+1) valores de u_t están dados por

$$\tilde{u}_t = \tilde{\phi}_t z_0 + \tilde{\phi}_{t+1} z_{-1} + \dots + \phi_p z_{-p+t} + \tilde{\theta}_t \tilde{b}_0 + \theta_{t+1} \tilde{b}_{-1} + \dots + \theta_q \tilde{b}_{-q+t};$$

4. se calcula el estadístico usando la fórmula escalar de Tanaka (1990)

$$\mathcal{TR}_{n}^{**} = \frac{\sum_{\tau=1}^{n-1} \left(\sum_{t=1}^{\tau-1} (\tau-t) \tilde{z}_{t}^{**} - \frac{\tau}{n} \sum_{t=1}^{n-1} (n-t) \tilde{z}_{t}^{**} \right)^{2}}{\sum_{\tau=1}^{n-1} \left(\frac{1}{\tau(\tau+1)} \left(\sum_{t=1}^{\tau} t \tilde{z}_{t}^{**} \right) \right)^{2}},$$

⁶En el apéndice A se describe un algoritmo para calcular una forma más conveniente del estadístico \mathcal{TR}_n^{**} que se propone más adelante. El algoritmo permite también la estimación de los valores premuestrales.

o la matricial

$$\mathcal{TR}_n^{**} = rac{1}{n-1} rac{ ilde{\mathbf{z}}_{**}' (\mathbf{\nabla} \mathbf{\nabla}')^{-2} ilde{\mathbf{z}}_{**}}{ ilde{\mathbf{z}}_{**}' (\mathbf{\nabla} \mathbf{\nabla}')^{-1} ilde{\mathbf{z}}_{**}};$$

5. finalmente, se compara el valor del estadístico \mathcal{TR}_n^{**} con el correspondiente valor crítico del cuadro (3.1).

3.5. Contraste de no invertibilidad en un proceso ARIMA

En las secciones anteriores, se ha demostrado la equivalencia entre las formas básicas de los estadísticos de Nyblom y Mäkeläinen (1983), Tanaka (1990) y Saikkonen y Luukkonen (1993)

$$\mathcal{NM}_n = \frac{1}{n-1} \frac{\mathbf{y'MCC'My}}{\mathbf{y'My}}, \ \mathcal{T}_n = \frac{1}{n-1} \frac{\mathbf{z'}(\nabla \nabla')^{-2}\mathbf{z}}{\mathbf{z'}(\nabla \nabla')^{-1}\mathbf{z}}, \ \mathcal{SL}_n = \frac{1}{n-1} \frac{\mathbf{y'MCBB'C'My}}{\mathbf{y'My}}.$$

Se ha destacado que el estadístico \mathcal{NM}_n se obtiene en el modelo estructural de nivel local, mientras que los estadísticos \mathcal{T}_n y \mathcal{SL}_n se obtienen el marco del modelo IMA(1,1). Además, se ha resaltado que los estadísticos \mathcal{NM}_n y \mathcal{SL}_n se expresan en términos de la serie integrada, mientras que el estadístico \mathcal{T}_n se calcula a partir de la serie diferenciada. Se ha mostrado también que la obtención del estadístico \mathcal{NM}_n es mucho más directa que la de los estadísticos \mathcal{T}_n y \mathcal{SL} : el primero se obtiene de la derivada primera de la función de verosimilitud, mientras que los dos últimos se obtienen de la derivada segunda. Finalmente, se han revisado cuatro contrastes para detectar la presencia de una raíz MA unitaria en un modelo ARIMA general invertible, los propuestos Tanaka (1990), Saikkonen y Luukkonen (1993), Leybourne y McCabe (1994) y Tam y Reinsel (1997). Los tres últimos pueden contemplarse como tres correcciones paramétricas alternativas del contraste simple de no invertibilidad de Nyblom y Mäkeläinen (1983); el primero, como una corrección no paramétrica.

En esta sección se propone un contraste de raíz MA unitaria en el marco de un modelo ARIMA general no invertible. La principal aportación radica, por tanto, en permitir la presencia de otras raíces MA unitarias. El modelo ARIMA con múltiples raíces MA unitarias es especialmente conveniente por dos razones. En primer lugar, se trata de una especificación que puede presentarse en la descripción de series estacionales con modelos ARIMA multiplicativos, en los que pueden surgir dos formas de no invertibilidad: regular y estacional. En segundo lugar, la inclusión en el modelo ARIMA de componentes deterministas (deriva, tendencia lineal o ciclos estacionales) puede realizarse con estructuras de tipo IMA no invertibles. De este modo, el modelo ARIMA general no invertible proporciona un marco común que permite formular una gran diversidad de contrastes de no invertibilidad o estabilidad paramétrica.

Aunque los contrastes de no invertibilidad pueden obtenerse a partir de la segunda derivada de la función de verosimilitud del modelo ARIMA, los resultados establecidos en las secciones anteriores sugieren encontrar representaciones estructurales equivalentes. Por esta razón el primer problema que se plantea en esta sección es el de extender el contraste de Nyblom y Mäkeläinen (1983) al modelo de nivel local generalizado, en el que los errores presentan correlación serial. El estadístico resultante es equivalente al de Saikkonen y Luukkonen (1993), SL_n^* , pero su obtención es más intuitiva e inmediata. Las identidades matriciales propuestas en las secciones anteriores, y que se han usado para establecer la equivalencia entre las tres formas básicas del estadístico de no invertibilidad, se usan ahora para encontrar una forma conveniente del estadístico $S\mathcal{L}_n^*$ que permita evaluar su distribución muestral tanto bajo H_0 como bajo H_1 . De este modo, no solo se pueden calcular los valores críticos del contraste, sino que además se puede evaluar la potencia del contraste. Más aún, se podrían calcular estadísticos POI. Las nuevas expresiones del estadístico SL_n^* son también interesantes en el análisis práctico de series temporales. Por un lado, se proporciona una forma conveniente para calcular el estadístico a partir de los residuos exactos del modelo ARIMA. Por otro lado, se proporciona otra forma conveniente para calcular los valores críticos y *p*-valores con los procedimientos de integración numérica de Imhof (1961) o Davies (1973). También se demuestra que los estadísticos corregidos propuestos por Leybourne y McCabe (1994) y Tam y Reinsel (1997) pueden contemplarse como aproximaciones al estadístico SL_n^* basadas en los residuos condicionados a valores premuestrales nulos o estimados con retrovisiones. Este resultado sugiere considerar un estadístico aproximado basado en los residuos exactos, que denotamos por \mathcal{GM}_n^* .

Finalmente, se evalúan los diferentes estadísticos realizando un doble estudio. En primer lugar, se supone que los parámetros del modelo ARIMA son conocidos y se tabulan las distribuciones de los estadísticos en muestras finitas por integración numérica. Con este ejercicio se evalúa el efecto que la correlación serial tiene sobre la distribución muestral del estadístico. En segundo lugar, se supone que los parámetros son desconocidos y se estiman por máxima verosimilitud exacta. Para ello se realiza un estudio de Monte Carlo y se evalúa, además del efecto de la correlación, el impacto que tiene la estimación en la distribución muestral.

3.5.1. El modelo de nivel local generalizado

El modelo de nivel local (3.1) considerado por Nyblom y Mäkeläinen (1983) se generaliza aquí del siguiente modo:

$$y_{t} = \mu_{t} + u_{t}, \quad \phi(B)u_{t} = \theta(B)w_{1t},$$

$$\mu_{t} = \mu_{t-1} + v_{t}, \quad \phi(B)v_{t} = \theta(B)w_{2t}, \quad t = 1, \dots, n,$$
(3.62)

en donde los errores u_t y v_t son dos procesos ARMA(p,q) independientes, pero con estructura ARMA común dada por $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ y $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$. Se supone que las raíces de $\phi(B)$ caen fuera del círculo unitario, pero se permite que las raíces de $\theta(B)$ caigan fuera o sobre el círculo unitrario. Además, se supone que ambos polinomios no tienen raíces comunes, que indicarían la presencia de factores comunes redundantes. Como en (3.1), se supone que el valor premuestral μ_0 es no estocástico, $w_{1t} \sim iidN(0, \sigma_1^2), w_{2t} \sim iidN(0, \sigma_2^2)$ y $E(w_{1t}w_{2,t-k}) = 0 \forall k$.

La generalización (3.62) es especialmente interesante porque su forma reducida asociada pertenece a la clase ARIMA(p, 1, q + 1)

$$\phi(B)\nabla y_t = \theta(B)(1 - \theta B)a_t. \tag{3.63}$$

Para comprobarlo simplemente se reescribe (3.62) como

$$\nabla y_t = v_t + u_t - u_{t-1}, \quad t = 2, \dots, n,$$
 (3.64)

0

$$\phi(B)\nabla y_t = \theta(B)[w_{2t} + (1-B)w_{1t}],$$

en donde $w_{2t} + (1 - B)w_{1t}$ sigue un proceso MA(1) con parámetro θ tal que $\rho = \sigma_2^2/\sigma_1^2 = (1 - \theta)^2/\theta$ y con varianza del error $\sigma_a^2 = \sigma_1^2/\theta$. La equivalencia se ha definido en térmi-

nos de la función de autocovarianzas del proceso estacionario $z_t = \nabla y_t$. De aquí, el contraste de no invertibilidad H_0 : $\theta = 1$ frente H_1 : $\theta \neq 1$ en (3.63) puede obtenerse como un contraste de nivel local determinista H_0 : $\rho = 0$ frente H_1 : $\rho > 0$ en (3.62).

El supuesto de que los errores u_t y v_t en (3.62) tienen una esctructura ARMA(p,q)común no es tan restrictivo como puede parecer en principio. En efecto, si se permiten procesos distintos del tipo $u_t \sim \text{ARMA}(p_1,q_1)$ y $v_t \sim \text{ARMA}(p_2,q_2)$, entonces $\nabla y_t \sim \text{ARMA}(p,q+1)$ con $p = p_1 + p_2$ y $q = q_1p_2 + q_2p_1$, de donde se deduce que y_t tendrá asociada una forma estructural como la definida en (3.62). Esta es otra manifestación del denominado problema de identificación referente a que dos modelos estructurales aparentemente distintos pueden compartir la misma forma reducida.

Siguiendo el desarrollo descrito para el modelo simple (3.1), (3.62) puede escribirse como un modelo de regresión generalizado

$$\mathbf{y} = \mathbf{i}\mu_0 + \mathbf{e}, \quad \mathbf{e} = \mathbf{C}\mathbf{v} + \mathbf{u}, \tag{3.65}$$

en donde ahora $\mathbf{u} \sim N[\mathbf{0}, \sigma_1^2 \mathbf{\Omega}_{\mathbf{u}}(\boldsymbol{\varphi})], \mathbf{v} \sim N[\mathbf{0}, \sigma_2^2 \mathbf{\Omega}_{\mathbf{u}}(\boldsymbol{\varphi})]$ y la matriz $\mathbf{\Omega}_{\mathbf{u}}(\boldsymbol{\varphi})$ depende del vector de parámetros $\boldsymbol{\varphi} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$ a través de las autocovarianzas $\gamma_u(k)/\sigma_1^2$. Por tanto, la distribución muestral de **y** viene dada por

$$\mathbf{y} \sim N\left(\mathbf{i}\mu_0, \,\sigma_1^2 \mathbf{\Omega}_{\mathbf{y}}(\rho, \boldsymbol{\varphi})\right) \quad \text{con} \quad \mathbf{\Omega}_{\mathbf{y}}(\rho, \boldsymbol{\varphi}) = \rho \mathbf{C} \mathbf{\Omega}_{\mathbf{u}}(\boldsymbol{\varphi}) \mathbf{C}' + \mathbf{\Omega}_{\mathbf{u}}(\boldsymbol{\varphi}), \tag{3.66}$$

en donde las matrices **C** y **C**' son intercambiables por ser $\Omega_{\mathbf{u}}(\boldsymbol{\varphi})$ una matriz Toeplitz, de manera que $\Omega_{\mathbf{y}}(\rho, \boldsymbol{\varphi}) = \rho \mathbf{C}' \Omega_{\mathbf{u}}(\boldsymbol{\varphi}) \mathbf{C} + \Omega_{\mathbf{u}}(\boldsymbol{\varphi})$. Para simplificar la notación se escribe la matriz $\Omega_{\mathbf{y}}(\rho, \boldsymbol{\varphi})$ de orden $n \times n$ sólo en términos del parámetro de interés ρ , $\Omega_{\mathbf{y}}(\rho)$. Alternativamente, y siguiendo una aproximación similar a la de Tanaka (1990), se multiplica (3.65) por la matriz de diferencias ∇ , se obtiene la forma matricial de (3.64)

$$\mathbf{z} = \boldsymbol{\nabla} \mathbf{e}_{\prime} \quad \boldsymbol{\nabla} \mathbf{e} = \boldsymbol{\nabla} \mathbf{C} \mathbf{v} + \boldsymbol{\nabla} \mathbf{u}_{\prime}$$

cuya distribución muestral, recordando que $\nabla C = [0|I_{n-1}]$, es

$$\mathbf{z} \sim N\left[\mathbf{0}, \ \sigma_1^2 \mathbf{\Omega}_{\mathbf{z}}(\rho, \boldsymbol{\varphi})\right] \quad \text{con} \quad \mathbf{\Omega}_{\mathbf{z}}(\rho, \boldsymbol{\varphi}) = \rho \mathbf{\Omega}_{\mathbf{u}, n-1}(\boldsymbol{\varphi}) + \boldsymbol{\nabla} \mathbf{\Omega}_{\mathbf{u}, n}(\boldsymbol{\varphi}) \boldsymbol{\nabla}', \tag{3.67}$$

o, en términos de θ y σ_a^2 ,

$$\mathbf{z} \sim N\left[\mathbf{0}, \ \sigma_a^2 \mathbf{\Omega}_{\mathbf{z}}(\theta, \boldsymbol{\varphi})\right] \quad \text{con} \quad \mathbf{\Omega}_{\mathbf{z}}(\theta, \boldsymbol{\varphi}) = \boldsymbol{\nabla}(\theta) \mathbf{\Omega}_{\mathbf{u},n}(\boldsymbol{\varphi}) \boldsymbol{\nabla}(\theta)'.$$
 (3.68)

De nuevo, para simplificar la notación, la matriz cuadrada $\Omega_z(\theta, \varphi)$ de orden n - 1 se escribe sólo en términos del parámetro de interés ρ , si bien hay que recordar su dependencia de los restantes parámetros de la estructura ARMA adicional. Comparando (3.67) y (3.68) es claro que resulta más inmediato calcular la primera derivada de $\Omega_z(\rho)$ respecto de ρ que la segunda derivada de $\Omega_z(\theta)$ respecto de θ . Por esta razón, en esta tesis, se prefiere el uso de la representación estructural en la obtención de los estadísticos. Una vez obtenido el estadístico de esta forma, los resultados intermedios servirán de guía para su obtención en la representación ARIMA.

3.5.2. Estadísticos de contraste

Por el teorema 2, el contraste LBI rechaza la hipótesis nula H_0 : $\rho = 0$ frente a la alternativa unilateral H_1 : $\rho > 0$ bajo (3.66) cuando

$$\mathcal{NM}_{n}^{*} = \frac{1}{n-1} \frac{\tilde{\mathbf{e}}_{*}^{\prime} \mathbf{C}_{*}^{\prime} \mathbf{\Omega}_{\mathbf{u},n} \mathbf{C}_{*} \tilde{\mathbf{e}}_{*}}{\tilde{\mathbf{e}}_{*}^{\prime} \tilde{\mathbf{e}}_{*}} > \kappa$$
(3.69)

en donde $\tilde{\mathbf{e}}_*$ y \mathbf{C}_* se han definido previamente para \mathcal{SL}_n^* en (3.60). Comparando estos dos estadísticos, se observa que la diferencia entre ambos se debe a la presencia en el último de las matrices de desplazamiento **B** y **B**'. Sin embargo, hay que notar que en el modelo de regresión transformado

$$\mathbf{y}_* = \mathbf{i}_* y_0 + \mathbf{e}_*, \quad \mathbf{e}_* \sim N(0, \sigma_a^2 \mathbf{I}_n), \tag{3.70}$$

que se usa en la estimación MCG de la constante y_0 , $\hat{y}_0^* = (\mathbf{i}'_* \mathbf{i}_*)^{-1} \mathbf{i}'_* \mathbf{y}_*$. Por la propiedad de ortogonalidad $\mathbf{i}'_* \mathbf{\tilde{e}}_* = 0$, y notando que la primera fila de \mathbf{C}'_* es \mathbf{i}'_* , resulta que el primer elemento del vector columna $\mathbf{C}'_* \mathbf{\hat{e}}_*$ es nulo. De aquí, la matriz de desplazamiento vertical \mathbf{B}' , simplemente mueve el cero a la última posición. Dado que $\Omega_{\mathbf{u}}(\boldsymbol{\varphi})$ es una matriz Toeplitz simétrica, se cumple que $\mathbf{\tilde{e}}'_* \mathbf{C}_* \mathbf{B} \Omega_{\mathbf{u}}(\boldsymbol{\varphi}) \mathbf{B}' \mathbf{C}'_* \mathbf{\tilde{e}}_* = \mathbf{\tilde{e}}'_* \mathbf{C}_* \Omega_{\mathbf{u}}(\boldsymbol{\varphi}) \mathbf{C}'_* \mathbf{\tilde{e}}_*,$ por lo que \mathbf{B} y \mathbf{B}' no juegan ningún papel en el cálculo de \mathcal{SL}_n^* y pueden omitirse. En definitiva, $\mathcal{NM}_n^* = \mathcal{SL}_n^*$. Ahora se propone una forma alternativa del estadístico SL_n^* que es computacionalmene más conveniente porque permite reutilizar los procedimientos usados en la estimación de un modelo ARIMA. Con este fin, se aplica la aproximación de Tanaka (1990) a la forma reducida (3.66) para obtener el contraste LBI en términos de la serie diferenciada z. Así, por el corolario 2.17 se obtiene

$$\mathcal{SL}_{n}^{*} = \frac{1}{n-1} \frac{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(\rho_{0})^{-1} \mathbf{\Omega}_{\mathbf{u},n-1} \mathbf{\Omega}_{\mathbf{z}}(\rho_{0})^{-1} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(\rho_{0})^{-1} \mathbf{z}}.$$
(3.71)

La comparación directa de (3.69) con (3.71) no permite apreciar la igualdad que se esperaría encontrar entre estos dos estadísticos. Para probar la equivalencia, se propone aquí definir la matriz de diferencias transformada $\nabla_* = \nabla \Omega_y(\rho_0)^{1/2}$ y la matriz de proyección $\mathbf{M}_* = \nabla'_* (\nabla_* \nabla'_*)^{-1} \nabla_*$, que anula al vector \mathbf{i}_* porque $\nabla_* \mathbf{i}_* =$ $\nabla \Omega_y(\rho_0)^{1/2} \Omega_y(\rho_0)^{-1/2} \mathbf{i} = \mathbf{0}$. Además, como $\mathbf{z} = \nabla_* \mathbf{y}_*$, $\Omega_z(\rho_0) = \nabla' \Omega_y(\rho_0) \nabla =$ $\nabla'_* \nabla_* y \Omega_{\mathbf{u},n-1} = \nabla'_* \mathbf{C}'_* \Omega_{\mathbf{u},n} \mathbf{C}_* \nabla_*$, se tiene que

$$\mathcal{SL}_n^* = \frac{\mathbf{y}_*' \mathbf{\nabla}_*' (\mathbf{\nabla}_* \mathbf{\nabla}_*')^{-1} \mathbf{\nabla}_* \mathbf{C}_* \mathbf{\Omega}_{\mathbf{u}}(\boldsymbol{\varphi}) \mathbf{C}_*' \mathbf{\nabla}_*' (\mathbf{\nabla}_* \mathbf{\nabla}_*')^{-1} \mathbf{\nabla}_* \mathbf{y}_*}{\mathbf{y}_*' \mathbf{\nabla}_*' (\mathbf{\nabla}_* \mathbf{\nabla}_*')^{-1} \mathbf{\nabla}_* \mathbf{y}_*} = \mathcal{NM}_n^*.$$

El estadístico (3.71) para contrastar $H_0: \theta = 1$ frente $H_1: \theta \neq 1$ tiene la forma

$$\mathcal{SL}_{n-1}^* = \frac{1}{n-1} \frac{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-1} \mathbf{\Omega}_{\mathbf{z}}(0) \mathbf{\Omega}_{\mathbf{z}}(1)^{-1} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(1)^{-1} \mathbf{z}},$$
(3.72)

en donde $\Omega_{\mathbf{z}}(0)$ y $\Omega_{\mathbf{z}}(1)$ se obtienen de $\Omega_{\mathbf{z}}(\theta)$.

3.5.3. Distribución exacta

Puesto que $\mathbf{z}_* = \mathbf{\Omega}_{\mathbf{u},n-1}^{-1/2}(\boldsymbol{\varphi})\mathbf{z}$, se tiene que $\mathbf{z}_* \sim N(\mathbf{0}, \sigma_1^2(\boldsymbol{\rho}\mathbf{I}_{n-1} + \boldsymbol{\nabla}_{\#}\boldsymbol{\nabla}'_{\#}))$, en donde $\boldsymbol{\nabla}_{\#} = \mathbf{\Omega}_{\mathbf{u},n-1}^{-1/2}(\boldsymbol{\varphi})\boldsymbol{\nabla}\mathbf{\Omega}_{\mathbf{u},n}^{1/2}(\boldsymbol{\varphi})$. Además, notando que $\mathbf{\Omega}_{\mathbf{u},n-1}^{1/2}(\boldsymbol{\varphi})(\boldsymbol{\nabla}\mathbf{\Omega}_{\mathbf{u},n}\boldsymbol{\varphi})\boldsymbol{\nabla}')^{-1}\mathbf{\Omega}_{\mathbf{u},n-1}^{1/2}(\boldsymbol{\varphi}) = (\boldsymbol{\nabla}_{\#}\boldsymbol{\nabla}'_{\#})^{-1}, (3.71)$ puede escribirse como

$$\mathcal{SL}_{n}^{*} = \frac{\epsilon' (\nabla_{\#} \nabla_{\#}')^{-1} \epsilon}{\epsilon' \epsilon}, \qquad (3.73)$$

en donde

$$\boldsymbol{\epsilon} = (\boldsymbol{\nabla}_{\#} \boldsymbol{\nabla}'_{\#})^{-1/2} \mathbf{z}'_{*} \sim N \left(\mathbf{0}, \ \sigma_{1}^{2} (\mathbf{I}_{n-1} + \rho \boldsymbol{\nabla}_{\#} \boldsymbol{\nabla}'_{\#}) \right).$$

Comparando (3.73) con (3.15) y siguiendo un desarrollo similar se obtiene

$$S\mathcal{L}_{n}^{*} = \frac{\sum_{k=1}^{n-1} \lambda_{k} (1+\rho\lambda_{k})\xi_{k}}{\sum_{k=1}^{n-1} (1+\rho\lambda_{k})\xi_{k}},$$
(3.74)

en donde λ_k (k = 1, ..., n - 1) son ahora los recíprocos de los autovalores de $\nabla_{\#} \nabla'_{\#}$, que también pueden obtenerse de la matriz

$$\boldsymbol{\Omega}_{\mathbf{u},n-1}^{-1}\boldsymbol{\nabla}\boldsymbol{\Omega}_{\mathbf{u},n}(\boldsymbol{\varphi})\boldsymbol{\nabla}' = 2\mathbf{I}_{n-1} - (\mathbf{B} + \mathbf{B}') - ([\boldsymbol{\varphi}_u|\mathbf{0}] + [\mathbf{0}|\boldsymbol{\varphi}_u]),$$

en donde $[\boldsymbol{\phi}_u|\mathbf{0}]$ es una matriz nula salvo por la primera columna $\boldsymbol{\phi}_u = \boldsymbol{\Omega}_{\mathbf{u},n-1}^{-1}(\boldsymbol{\varphi})\gamma_u$ con $\gamma_u = (\gamma_u(1) \dots \gamma_u(n-1))'$, y $[\mathbf{0}|\boldsymbol{\phi}_u]$ es una matriz nula salvo por la última columna $\boldsymbol{\phi}_u = \boldsymbol{\Omega}_{\mathbf{u},n-1}^{-1}(\boldsymbol{\varphi})\boldsymbol{\gamma}_u$ con $\boldsymbol{\gamma}_u = (\gamma_u(n-1) \dots \gamma_u(1))'$. Conviene notar que $[\mathbf{0}|\boldsymbol{\phi}_u]$ es la matriz $[\boldsymbol{\phi}_u|\mathbf{0}]$ rotada 180°.

Si el error u_t sigue un proceso ARMA invertible, la mayoría de coeficientes de autoregresión parcial son asintóticamente nulos y los autovalores vienen determinados por la matriz tridiagonal $\nabla \nabla'$. Consecuentemente, se obtiene la distribución límite

$$\mathcal{GM}_{n}^{*}/n \sim \sum_{k=1}^{\infty} \left(\frac{1}{k^{2}\pi^{2}} + \frac{c^{2}}{k^{4}\pi^{4}}\right) \xi_{k}.$$
 (3.75)

Ahora bien, si el proceso u_t es no invertible, entonces esta distribución asintótica se desplazará a la izquierda por la presencia de la matriz $[\boldsymbol{\phi}_u | \mathbf{0}] + [\mathbf{0} | \boldsymbol{\ddot{\phi}}_u]$

Un desarrollo similar puede seguirse para obtener la distribución límite del estadístico $S\mathcal{L}_n^*$ bajo la secuencia de alternativas $H_1: \theta = 1 - c/n$. Para ello se escribe (3.60) como

$$\mathcal{SL}_n^* = rac{\mathbf{y}_*' \mathbf{
abla}_*' (\mathbf{
abla}_* \mathbf{
abla}_*')^{-1} \mathbf{
abla}_* \mathbf{C}_* \mathbf{\Omega}_{\mathbf{b}} \mathbf{C}_*' \mathbf{
abla}_*' (\mathbf{
abla}_* \mathbf{
abla}_*')^{-1} \mathbf{
abla}_* \mathbf{y}_*}{\mathbf{y}_*' \mathbf{
abla}_*' (\mathbf{
abla}_* \mathbf{
abla}_*')^{-1} \mathbf{
abla}_* \mathbf{y}_*},$$

o bien como

$$\mathcal{SL}_n^* = \frac{\mathbf{z}'(\boldsymbol{\nabla}_*\boldsymbol{\nabla}_*')^{-1}\boldsymbol{\Omega}_{\mathbf{b},n-1}(\boldsymbol{\nabla}_*\boldsymbol{\nabla}_*')^{-1}\mathbf{z}}{\mathbf{z}'(\boldsymbol{\nabla}_*\boldsymbol{\nabla}_*')^{-1}\mathbf{z}}$$

en donde $\mathbf{z} = \nabla_* \mathbf{y}_* = \nabla \mathbf{y}$ es el proceso MA(1) impuro $\mathbf{z} = \nabla(\theta)\mathbf{b}$. Definiendo $\nabla_*(\theta) = \nabla(\theta) \mathbf{\Omega}_{\mathbf{b}}^{1/2} \mathbf{y} \mathbf{b}_* = \mathbf{\Omega}_{\mathbf{b}}^{-1/2} \mathbf{b} \sim N(\mathbf{0}, \sigma_b^2 \mathbf{I}_n)$, se tiene que

$$\mathcal{SL}_n^* = \frac{\mathbf{b}_*' \nabla_*(\theta)' (\nabla_* \nabla_*')^{-1} \Omega_{\mathbf{b},n-1} (\nabla_* \nabla_*')^{-1} \nabla_*(\theta) \mathbf{b}_*}{\mathbf{b}_*' \nabla_*(\theta)' (\nabla_* \nabla_*')^{-1} \nabla_*(\theta) \mathbf{b}_*},$$

que puede escribirse como

$$\mathcal{SL}_{n}^{*} = \frac{\mathbf{b}_{*}^{\prime} \nabla_{\#}(\theta)^{\prime} (\nabla_{\#} \nabla_{\#}^{\prime})^{-2} \nabla_{\#}(\theta) \mathbf{b}_{*}}{\mathbf{b}_{*}^{\prime} \nabla_{\#}(\theta)^{\prime} (\nabla_{\#} \nabla_{\#}^{\prime})^{-1} \nabla_{\#}(\theta) \mathbf{b}_{*}},$$
(3.76)

en donde $\nabla_{\#}(\theta) = \Omega_{\mathbf{b},n-1}^{-1/2} \nabla(\theta) \Omega_{\mathbf{b},n}^{1/2}$ y $\nabla_{\#} = \nabla_{\#}(1)$. Comparando (3.76) con (3.45), y siguiendo los mismos pasos que en la derivación de la distribución de \mathcal{T}_n , se obtiene

$$\mathcal{SL}_{n}^{*} = \frac{\sum_{k=1}^{n-1} [\lambda_{k}^{\#}(\theta) / \lambda_{k}^{\#}(1)^{2}] \xi_{k}}{\sum_{k=1}^{m} [\lambda_{k}^{\#}(\theta) / \lambda_{k}^{\#}(1)] \xi_{k}},$$
(3.77)

en donde $\lambda_k^{\#}(\theta)$ son los recíprocos de los autovalores de la matriz $\nabla_{\#}(\theta)\nabla_{\#}(\theta)'$, que también pueden obtenerse de la matriz

$$\boldsymbol{\Omega}_{\mathbf{b},n-1}^{-1}\boldsymbol{\nabla}(\theta)\boldsymbol{\Omega}_{\mathbf{b},n}\boldsymbol{\nabla}(\theta)' = (1+\theta^2)\mathbf{I}_{n-1} - \theta(\mathbf{B}+\mathbf{B}') - \theta([\boldsymbol{\phi}_b|\mathbf{0}] + [\mathbf{0}|\boldsymbol{\phi}_b]).$$

Si el proceso b_t es invertible, la mayoría de coeficientes de autoregresión parcial son asintoticamente nulos y los autovalores vienen determinados por la matriz tridiagonal $\nabla \nabla'$. Consecuentemente, se obtiene la distribución límite

$$\mathcal{SL}_{n}^{*}/n \stackrel{a}{\sim} \sum_{k=1}^{\infty} \left(\frac{1}{k^{2}\pi^{2}} + \frac{c^{2}}{k^{4}\pi^{4}}\right) \xi_{k}.$$
 (3.78)

3.5.4. Estadísticos alternativos

Como se ha descrito en la subsección 3.4.5, Tam y Reinsel (1997) especificaron el modelo ARIMA(p, 1, d + 1) como

$$\phi(B)\nabla y_t = \theta(B)b_t, \quad t = 1, \dots, n,$$

$$b_t = (1 - \theta B)a_t.$$
(3.79)

La primera ecuación se usa para filtrar la serie y_t y transformarla en el ruido b_t , y la segunda sirve para calcular el estadístico de Tanaka como

$$\mathcal{TR}_n^* = \frac{1}{n-1} \frac{\tilde{\mathbf{b}}(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-2}\tilde{\mathbf{b}}}{\tilde{\mathbf{b}}(\boldsymbol{\nabla}\boldsymbol{\nabla}')^{-2}\tilde{\mathbf{b}}}.$$

La serie filtrada \tilde{b}_t se obtiene en la primera ecuación, $\phi(B)\nabla y_t = \theta(B)b_t$, fijando los valores premuestrales $(y_0, \ldots, y_{-p}, b_0, \ldots, b_{-q})$ en cero o estimándolos con el método de retrovisión. Aquí se propone una forma alternativa de estimar el ruido b_t usando la segunda ecuación, $b_t = (1 - \theta B)a_t$, que permite obtener \tilde{b}_t a partir de los residuos \tilde{a}_t . Bajo H_0 , $\tilde{\mathbf{b}} = \nabla \tilde{\mathbf{a}}$; por tanto, el estadístico de Tanaka puede expresarse como

$$\mathcal{GM}_n^* = \frac{1}{n-1} \frac{\tilde{\mathbf{a}}' \boldsymbol{\nabla}' (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-2} \boldsymbol{\nabla} \tilde{\mathbf{a}}}{\tilde{\mathbf{a}}' \boldsymbol{\nabla}' (\boldsymbol{\nabla} \boldsymbol{\nabla}')^{-1} \boldsymbol{\nabla} \tilde{\mathbf{a}}};$$

y usando las identidades matriciales $\nabla' (\nabla \nabla')^{-2} \nabla = \mathbf{M} \mathbf{C}' \mathbf{C} \mathbf{M} \ \mathbf{y} \ \nabla' (\nabla \nabla')^{-1} \nabla = \mathbf{M}$, se tiene que

$$\mathcal{GM}_n^* = \frac{1}{n-1} \frac{\tilde{\mathbf{a}}' \mathbf{MC}' \mathbf{CM} \tilde{\mathbf{a}}}{\tilde{\mathbf{a}}' \mathbf{M} \tilde{\mathbf{a}}} = \frac{\sum_{t=1}^n [\sum_{\tau}^t (\tilde{a}_{\tau} - \tilde{a})]^2}{\sum_{t=1}^n (\tilde{a}_t - \tilde{a})^2}.$$

Ahora se ve claramente que distintos supuestos sobre el cálculo de los residuos conducen a distintos estadísticos de contraste. Leybourne y McCabe (1994) fijaron los valores premuestrales en cero, lo que conduce a un estadístico basado en los residuos condicionales $\tilde{a}_t^{(0)}$; Tam y Reinsel (1997) consideraron también la posibilidad de estimar los valores premuestrales con retovisiones, que conduce a un estadístico basado en los residuos condicionados a las retrovisiones $\tilde{a}_t^{(b)}$. Aquí, se propone el uso de los residuos exactos del modelo (3.79).

Siguiendo a Gallego (2009), para calcular los residuos exactos se escribe el modelo ARI-MA en forma matricial como

$$\mathbf{z} = \mathbf{\Phi}^{-1}[\mathbf{\Theta}\mathbf{a} + \mathbf{F}\mathbf{u}_0]$$

donde $\mathbf{\Phi} = \mathbf{I}_{n-1} - \phi_1 \mathbf{B} - \dots - \phi_p \mathbf{B}^p$ y $\mathbf{\Theta} = \mathbf{I}_{n-1} - \theta_1 \mathbf{B} - \dots - \theta_{q+1} \mathbf{B}^{q+1}$ son matrices $(n-1) \times (n-1)$ con estructura triangular inferior, $z = (z_2, \dots, z_n)$ y $a = (a_2, \dots, a_n)$ son vectores de dimensión (n-1), $\mathbf{u}_0 = (z_1, z_0, \dots, z_{p-2}, a_1, a_0, \dots, a_{q-1})$ es un vector con p + q + 1 valores premuestrales, y **F** es una matriz $(n-1) \times (p+q+1)$. Los residuos exactos **ã** son

$$\tilde{\mathbf{a}} = E(\mathbf{a}|\mathbf{z}') = E(\tilde{\mathbf{a}}\mathbf{z}')E(\mathbf{z}\mathbf{z}')^{-1}\mathbf{z} = \mathbf{\Psi}'\mathbf{\Omega}_{\mathbf{z}}^{-1}\mathbf{z}.$$

donde $\Psi = \Phi^{-1}\Theta = \mathbf{I}_{n-1} + \psi_1 \mathbf{B} + \cdots + \psi_{n-1} \mathbf{B}^{n-1}$ es una matriz triangular inferior

que contiene el peso ψ_i en la subdiagonal *j*-ésima.

3.5.5. Múltiples raíces MA unitarias

El procedimiento de contraste propuesto resulta también válido cuando la estructura ARMA(p,q) es homogéneamente no invertible. Como ilustración se considera de nuevo el contraste de Nyblom (1986) descrito en la subsección 3.2.7. El modelo de nivel local con una deriva es equivalente al proceso IMA(2,2) no invertible

$$\nabla^2 y_t = (1-B)(1-\theta B)a_t,$$

en donde el problema de contrastar H_0 : $\theta = 1$ frente a H_1 : $\theta \neq 1$ plantea la cuestión de si el proceso MA(2) tiene dos raíces MA unitarias o solamente una. Se trata, por tanto, de contrastar la presencia de una raíz MA unitaria adicional. El modelo IMA(1,2) para la primera diferencia de y_t , ∇y_t , puede contemplarse como un caso especial de (3.64)

$$abla y_t = \mu_t + u_t, \quad u_t = (1 - B)\omega_{1t},$$

 $\mu_t = \mu_{t-1} + v_t, \quad v_t = (1 - B)\omega_{2t}.$

Notando que $\Omega_{\mathbf{u},n} = \nabla_{n-1,n} \nabla'_{n-1,n} \operatorname{y} \nabla_{n-1,n} \Omega_{\mathbf{u},n} \nabla_{n-1,n} = \nabla^2 \nabla^{2'}$, se comprueba fácilmente que (3.71) se reduce al estadístico \mathcal{N}_n .

3.5.6. Evaluación de los contrastes

Teóricamente, las correcciones al estadístico de contraste de Nyblom y Mäkeläinen (1983) propuestas por Tanaka (1990), Saikkonen y Luukkonen (1993), Leybourne y Mc-Cabe (1994) y Tam y Reinsel (1997) consiguen que la presencia de correlación serial no altere su distribución muestral asintótica bajo la hipótesis nula de no invertibilidad. Sin embargo, en muestras finitas, la distribución de estos cuatros estadísticos es desconocida, por lo que es interesante evaluar cómo se ve afectada por diferentes formas de correlación serial. Con este propósito se consideran dos escenarios: (1) los parámetros son conocidos y no es necesario estimar el modelo y (2) los parámetros son desconocidos y se estiman por máxima verosimilitud exacta.

Parámetros conocidos

Una característica interesante de los estadísticos de no invertibilidad es que pueden expresarse como ratios de formas cuadráticas en variables normales. Estas expresiones permiten tabular las distribuciones muestrales mediante procedimientos de integración numérica, sin necesidad de recurrir a simulación. En este apartado se tabulan las distribuciones muestrales de los estadísticos SL_n , GM_n , TR_n , LM_n y T_n suponiendo que el modelo teórico es un MA(2) no invertible

$$z_t = (1-B)(1-\theta B)a_t,$$

donde el primer factor MA, 1 - B, se corresponde con la hipótesis nula de no invertibilidad, y el segundo permite controlar la presencia de correlación serial extra. Cuando el parámetro de control θ es menor que la unidad, las distribuciones muestrales de todos los estadísticos deberían coincidir con las del estadístico \mathcal{NM}_n , véase Cuadro 3.1; en cambio, cuando $\theta = 1$, la distribución de referencia sería la del estadístico \mathcal{N}_n . Cuadro 3.5, que se encuentra desplazada a la izquierda en relación a la de \mathcal{NM}_n . Cabría esperar que, a medida que θ se acerca a la unidad, la distribución muestral de cada estadístico se desplace hacia la de \mathcal{N}_n . De este modo, el modelo MA(2) permite evaluar la capacidad de cada estadístico para contrastar la presencia de múltiples raíces. Por otro lado, como el MA(2) invertible puede aproximarse por un AR de orden finito,

$$(1+\theta B+\theta^2 B^2+\cdots+\theta^p B^p)z_t=(1-B)a_t,$$

el modelo permite también evaluar el efecto de este tipo correlación serial.

Para el modelo MA(2) se definen las siguientes matrices para el proceso estacionario z_t : Ω_z , matriz de covarianzas; L_z , factor de Cholesky de Ω_z ; Ψ_z , matriz triangular con los pesos ψ_j del MA(2), $\psi_0 = 1$, $\psi_1 = -(1 + \theta)$, θ ; Π_z , matriz triangular con los pesos π_j del MA(2), $\pi_j = (1 + \theta)\pi_{j-1} - \theta\pi_{j-2}$. Análogamente, se definen las correspondientes matrices para el proceso auxiliar $u_t = (1 - \theta B)\omega_t$.

Usando esta notación, los estadísticos de no invertibilidad pueden expresarse en términos de un vector $\boldsymbol{\epsilon} \sim N(0, \sigma^2 \mathbf{I})$:

1. Saikkonen-Luukkonen

$$\mathcal{SL}_{n}^{*} = \frac{1}{n-1} \frac{\tilde{\mathbf{e}}_{u,n}^{\prime} \mathbf{L}_{u,n}^{-1} \mathbf{C}^{\prime} \boldsymbol{\Omega}_{u,n} \mathbf{CL}_{u,n}^{\prime-1} \tilde{\mathbf{e}}_{*}}{\tilde{\mathbf{e}}_{*}^{\prime} \tilde{\mathbf{e}}_{*}} \frac{1}{n-1} \frac{\boldsymbol{\epsilon}^{\prime} \mathbf{L}_{u,n-1}^{\prime} \boldsymbol{\Omega}_{z,n-1}^{-1} \mathbf{L}_{u,n-1} \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}^{\prime} \boldsymbol{\epsilon}}.$$

2. Gallego-Mazas

$$\mathcal{GM}_n^* = \frac{1}{n-1} \frac{\tilde{\mathbf{a}}' \mathbf{M} \mathbf{C}' \mathbf{C} \mathbf{M} \tilde{\mathbf{a}}}{\tilde{\mathbf{a}}' \mathbf{M} \tilde{\mathbf{a}}} = \frac{1}{n-1} \frac{\epsilon' \mathbf{\Psi}_{\mathbf{z}} \mathbf{L}_{\mathbf{u}}^{-1} \mathbf{M} \mathbf{C}' \mathbf{C} \mathbf{M} \mathbf{\Psi}_{\mathbf{z}} \mathbf{L}_{\mathbf{u}}' \epsilon}{\epsilon' \mathbf{\Psi}_{\mathbf{z}} \mathbf{L}_{\mathbf{u}}^{-1} \mathbf{M} \mathbf{\Psi}_{\mathbf{z}} \mathbf{L}_{\mathbf{u}}' \epsilon}.$$

3. Leybourne-McCabe/Tam-Reinsel

$$\mathcal{LM}_{n} = \frac{1}{n-1} \frac{\tilde{\mathbf{a}}_{0}' \mathbf{MC}' \mathbf{CM} \tilde{\mathbf{a}}_{0}}{\tilde{\mathbf{a}}_{0}' \mathbf{M} \tilde{\mathbf{a}}_{0}} = \frac{1}{n-1} \frac{\epsilon' \mathbf{L}_{\mathbf{z}}' \mathbf{\Pi}_{\mathbf{u}}' \mathbf{MC}' \mathbf{CM} \mathbf{\Pi}_{\mathbf{u}} \mathbf{L}_{\mathbf{z}} \epsilon}{\epsilon' \mathbf{L}_{\mathbf{z}}' \mathbf{\Pi}_{\mathbf{u}}' \mathbf{M} \mathbf{\Pi}_{\mathbf{u}} \mathbf{L}_{\mathbf{z}} \epsilon}$$

4. Tanaka

$$\mathcal{T}_n = arphi rac{\mathbf{z}'(\mathbf{
abla}\mathbf{
abla}')^{-2}\mathbf{z}}{\mathbf{z}'(\mathbf{
abla}\mathbf{
abla}')^{-2}\mathbf{z}} = arphi rac{\mathbf{\epsilon}' \mathbf{L}_z'(\mathbf{
abla}\mathbf{
abla}')^{-2}\mathbf{L}_z \mathbf{\epsilon}}{\mathbf{\epsilon}' \mathbf{L}_z'(\mathbf{
abla}\mathbf{
abla}')^{-1}\mathbf{L}_z \mathbf{\epsilon}}.$$

El Cuadro 3.6 muestra los valores críticos de los cuatro estadísticos al nivel de significación del 5%, para distintos tamaños muestrales ($n = 10, 20, \dots, 100$) y diferentes valores del parámetro de control ($\theta = -1, -0.75, \dots, 1$). Los cuatro estadísticos deberían coincidir cuando $\theta = 0$ porque no es necesario aplicar ninguna corrección. Este resultado se observa en la columna correspondiente a $\theta = 0$, que se repite en las cuatro subtablas, en donde el valor crítico es aproximadamente 0.47 para n > 30 y ligeramente superior, 0.50, para n = 10. Se observa también que cuando la raíz MA extra está alejada de la unidad, θ < 0.5, los cuatro estadísticos se comportan razonablemente bien puesto que sus valores críticos están próximos a 0.47. Sin embargo, a medida que θ se aproxima a uno se aprecian dos comportamientos diferentes: mientras que las distribuciones de SL_n y GM_n se desplazan hacia la izquierda acercándose a la distribución de \mathcal{N}_n , los estadísticos de \mathcal{LM}_n y \mathcal{T}_n se desplazan en dirección opuesta, alejándose de la distribución de referencia. Estos desplazamientos en la distribución muestral repercuten en el tamaño empírico del contraste, es decir, en el nivel de significación resultante cuando se utilizan los valores críticos del estadístico \mathcal{NM}_n . Estos tamaños empíricos se muestran en el Cuadro 3.7, salvo la columa $\theta = 1$, que muestra el tamaño empírico calculado con los valores críticos de N_n . De nuevo se observa que, cuando θ está lejos de la unida, los tamaños son próximos al 5 %. En cambio, cuando θ se acerca a uno, los tamaños de SL_n y GM_n son menores que 5 % por el desplazamiento a la izquierda de sus respectivas distribuciones, mientras que los tamaños de SL_n y GM_n están muy por encima del 5 %.

Las principales conclusiones que pueden extraerse de esta evidencia empírica son las siguientes:

- 1. El empleo de los residuos exactos en el cálculos de los estadísticos de no invertibilidad, como en SL_n y GM_n , proporciona mejores resultados que otros residuos o correcciones no paramétricas, como en \mathcal{LM}_n y \mathcal{T}_n .
- 2. En presencia de otras raíces MA positivas se deberían usar los valores críticos correctos, en lugar de los de \mathcal{NM}_n . En este sentido, la forma propuesta para el estadístico de \mathcal{SL}_n proporciona un método conveniente para el cálculo de valores críticos específicos para cada modelo.

Parámetros desconocidos

En una aplicación real, los parámetros del modelo son desconocidos y tienen que ser estimados. La estimación introduce una fuente de aleatoriedad adicional que puede distorsionar aún mas la distribución muestral de los estadísticos de no invertibilidad. Para evaluar este efecto se vuelve a considerar el proceso generador de datos

$$z_t = (1-B)(1-\theta B)a_t,$$

y se estima por máxima verosimilitud exacta el modelo

$$z_t = (1 - \theta_0 B)(1 - \theta B)a_t,$$

donde la estimación de θ_0 debería ser próxima a 1 y la de θ , al valor fijado para el parámetro de control. Se cumple, por tanto, que $\theta_0 \ge \theta$. Esta condición implica que, una vez estimado el modelo, la estimación más alta se identifica con θ_0 y se fija en 1. Se obtiene así el modelo bajo H_0 con el que se calculan los estadísticos descritos en el apartado anterior. Además, se considera el estadístico de $T\mathcal{R}_n$ basado en los residuos calculados con el método de retrovisión.

El Cuadro 3.8 muestra el percentil del 95% de la distribución muestral de cada estadístico, que se ha obtenido por simulación generando 10000 muestras aleatorias con diferentes tamaños muestrales y distintos valores de θ . En términos generales, los percentiles obtenidos por simulación con un $\alpha = 5\%$ son muy similares a los valores críticos cuando $\theta = 0$, siendo el percentil para 100 observaciones igual a 0.47. Cuando $\theta < 0.5$, los percentiles de la distribución muestral de los estadísticos SL_n , GM_n , TR_n y \mathcal{LM}_n son muy cercanos al valor del percentil cuando $\theta = 0$. Esto supone que los cuatro estadísticos objeto de estudio se comportan en conjunto de la misma forma, y además funcionan bastante bien cuando la raíz MA extra está alejada de la unidad. Esta situación se modifica cuando el valor de la raíz MA extra se aproxima a 1. Al igual que sucede en el Cuadro 3.6 se obervan dos comportamientos diferentes, por una parte, la distribución muestral de los estadísticos SL_n y GM_n se desplaza hacia la izquierda y sus valores de aproximan a los de la distribución de N_n . Como ejemplo, el percentil de la distribución simulada del estadístico SL_n para $\theta = 1$ y n = 100 coincide de forma exacta con el valor crítico reportado en el Cuadro 3.6. Por el contrario, y coincidiendo también con los resultados del Cuadro 3.6, la distribución muestral simulada de los estadísticos TR_n y LM_n se desplaza en sentido contrario alejándose de los valores para una raíz unitaria doble de N_n

Con respecto al efecto de la estimación del parámetro θ_0 sobre la distribución muestral se aprecia una distorsión sobre los valores críticos en el Cuadro (3.6), y de forma más severa cuando n < 50. La cuantía de las diferencias entre el valor del percentil obtenido mediante simulación y el valor crítico es variable, tanto en cuantía como en signo, y afecta por igual a todos los estadísticos.

En síntesis, los principales resultados que se pueden reportar son las siguientes:

- La simulación de la distribución muestral de los estadísticos SL_n, GM_n, LM_n y T_n refleja unos resultados muy similares a los recogidos en el Cuadro 3.6. De nuevo, los mejores resultados son los que proporcionan los estadísticos SL_n y GM_n.
- Se aprecia una ligera distorsión entre los valores de la distribución simulada en el Cuadro 3.8 y los valores críticos del Cuadro 3.6, y de una forma más evidente cuando *n* < 50.

θ 1 -1 -0.75 -0.5 0 0.25 0.5 0.75 п -0.25 10 0.591 0.583 0.566 0.540 0.504 0.450 0.362 0.241 0.179 20 0.530 0.525 0.516 0.503 0.485 0.456 0.403 0.290 0.164 0.508 0.504 30 0.498 0.490 0.477 0.4580.420 0.328 0.159 40 0.496 0.493 0.489 0.483 0.474 0.459 0.430 0.355 0.156 50 0.489 0.487 0.484 0.479 0.471 0.459 0.436 0.373 0.155 0.485 0.483 0.480 0.476 0.469 0.460 0.440 0.386 0.154 60 70 0.482 0.480 0.477 0.4740.468 0.460 0.443 0.395 0.153 80 0.479 0.4770.4750.472 0.468 0.460 0.445 0.403 0.152 90 0.4770.4760.4740.4710.467 0.4600.447 0.409 0.152 100 0.476 0.4740.472 0.470 0.466 0.460 0.448 0.414 0.151 (b) \mathcal{GM}_n 0.540 0.513 0.505 0.504 0.504 0.440 10 0.504 0.497 0.308 0.485 0.485 0.479 0.422 20 0.506 0.490 0.485 0.485 0.220 30 0.492 0.481 0.478 0.477 0.477 0.477 0.473 0.427 0.192 0.474 0.4740.470 0.432 40 0.484 0.476 0.4740.474 0.178 0.480 0.473 0.471 0.437 50 0.471 0.471 0.471 0.469 0.171 60 0.477 0.471 0.470 0.470 0.470 0.470 0.467 0.440 0.166 70 0.475 0.470 0.469 0.469 0.469 0.468 0.467 0.443 0.163 80 0.473 0.469 0.468 0.468 0.468 0.467 0.466 0.445 0.161 90 0.472 0.468 0.467 0.467 0.467 0.467 0.465 0.447 0.159 0.471 0.466 0.466 100 0.467 0.466 0.466 0.465 0.4480.158 (c) \mathcal{LM}_n 10 0.420 0.481 0.494 0.498 0.504 0.538 0.678 1.003 1.210 0.394 20 0.471 0.479 0.481 0.485 0.511 0.672 1.390 2.204 0.384 0.468 0.473 0.474 0.477 0.497 0.642 1.552 3.203 30 0.380 0.617 40 0.466 0.470 0.471 0.474 0.489 1.615 4.203 0.597 0.377 0.468 0.469 0.471 0.4841.632 5.203 50 0.465 0.375 0.464 0.467 0.468 0.470 0.481 0.581 1.628 6.203 60 70 0.374 0.4640.466 0.467 0.468 0.478 0.568 1.612 7.203 0.373 80 0.464 0.466 0.466 0.468 0.476 0.557 1.589 8.203 90 0.372 0.463 0.465 0.466 0.467 0.475 0.549 1.564 9.203 100 0.371 0.463 0.465 0.465 0.466 0.473 0.541 1.537 10.203 (d) T_n 0.385 0.387 0.427 0.504 1.274 4.835 10 0.396 0.694 135.40 0.58420 0.422 0.423 0.427 0.443 0.485 0.867 2.572 65.30 0.435 0.449 0.729 1.833 42.71 30 0.435 0.438 0.477 0.54440 0.441 0.442 0.444 0.452 0.474 0.524 0.660 1.472 31.73 50 0.445 0.446 0.447 0.454 0.471 0.511 0.620 1.260 25.25 0.448 0.4480.449 0.455 0.470 0.503 0.593 1.121 20.99 60 70 0.450 0.4500.451 0.456 0.468 0.497 0.5741.022 17.98 0.451 0.451 0.452 0.456 0.493 0.559 0.949 15.73 80 0.468 0.452 0.453 0.489 0.548 0.893 13.99 90 0.452 0.457 0.467 0.453 0.453 0.454 0.457 0.487 0.539 12.61 100 0.466 0.848

Cuadro 3.6: Valores críticos al 5 % bajo $z_t = (1 - B)(1 - \theta B)a_t$ (a) SL_n

				Ç.	, - · - n				
					θ				
п	-1	-0.75	-0.5	-0.2	5 (0.25	0.5	0.75	1*
10	8.24	7.95	7.26	6.3	1 5.01	1 3.20	0.97	0.00	0.00
20	6.46	6.29	5.99	5.5	8 4.99	9 4.10	2.63	0.48	0.00
30	5.96	5.84	5.65	5.3	8 5.00	0 4.40	3.34	1.25	0.05
40	5.71	5.62	5.48	5.2	8 5.00) 4.55	3.73	1.88	0.16
50	5.56	5.48	5.37	5.2	2 5.00) 4.64	3.96	2.34	0.34
60	5.47	5.40	5.31	5.1	9 5.00	4.70	4.13	2.70	0.57
70	5.39	5.34	5.26	5.1	5 4.99	9 4.73	4.24	2.97	0.82
80	5.34	5.29	5.22	5.1	3 4.99	9 4.76	4.33	3.19	1.06
90	5.30	5.26	5.20	5.1	2 5.00	4.80	4.41	3.37	1.30
100	5.28	5.25	5.19	5.12	2 5.01	4.82	4.47	3.52	1.52
				(b)	\mathcal{GM}_n				
10	6.34	5.29	5.03	5.00	5.01	5.00	4.73	3.02	23.20
20	5.67	5.14	5.00	4.99	4.99	4.99	4.81	3.16	13.94
30	5.45	5.10	5.01	5.00	5.00	5.00	4.87	3.53	10.15
40	5.34	5.08	5.01	5.00	5.00	5.00	4.90	3.80	8.33
50	5.27	5.06	5.00	5.00	5.00	4.99	4.91	4.00	7.35
60	5.23	5.05	5.00	5.00	5.00	5.00	4.93	4.15	6.76
70	5.19	5.03	4.99	4.99	4.99	4.99	4.93	4.25	6.39
80	5.16	5.03	4.99	4.99	4.99	4.99	4.94	4.33	6.15
90	5.15	5.03	5.00	5.00	5.00	5.00	4.95	4.41	5.98
100	5.14	5.04	5.01	5.01	5.01	5.00	4.96	4.47	5.85
				(c)	\mathcal{LM}_n				
10	2.72	4.66	4.98	5.01	5.01	5.04	7.20	45.58	100
20	2.66	4.80	4.98	4.99	4.99	5.01	6.64	51.99	100
30	2.65	4.87	4.99	5.00	5.00	5.02	6.25	50.84	100
40	2.64	4.90	4.99	5.00	5.00	5.01	5.99	48.50	100
50	2.63	4.91	4.99	5.00	5.00	5.00	5.81	46.03	100
60	2.63	4.93	4.99	5.00	5.00	5.01	5.69	43.72	100
70	2.62	4.93	4.99	4.99	4.99	5.00	5.59	41.55	100
80	2.62	4.94	4.99	4.99	4.99	5.00	5.52	39.58	100
90	2.62	4.95	4.99	5.00	5.00	5.00	5.47	37.80	100
100	2.63	4.96	5.00	5.01	5.01	5.01	5.44	36.18	100
				(d) \mathcal{T}_n				
10	0.51	0.59	0.98	2.17	5.01	11.69	38.52	100.0	0 10
20	2.85	2.88	3.06	3.63	4.99	8.17	19.97	98.8	4 10
30	3.62	3.64	3.74	4.10	5.00	7.09	14.22	84.2	.4 10
40	3.98	3.99	4.07	4.33	5.00	6.55	11.57	64.6	9 10
50	4.19	4.20	4.25	4.46	5.00	6.23	10.06	50.1	1 10
60	4.33	4.34	4.38	4.55	5.00	6.02	9.12	40.3	0 10
70	4.42	4.42	4.46	4.61	4.99	5.86	8.45	33.5	51 10
80	4.49	4.50	4.53	4.65	4.99	5.75	7.97	28.7	0 10
90	4.55	4.56	4.59	4.70	5.00	5.67	7.62	25.1	7 10
			1 (1	4 7 4	E 01		7 24	22 5	0 10

Cuadro 3.7: Tamaño empírico al 5 % bajo $z_t = (1 - B)(1 - \theta B)a_t$ (a) \mathcal{SL}_n

		(11			
			θ			
п	0	0.5	0.8	0.9	1	
20	0.500	0.408	0.256	0.191	0.163	
30	0.472	0.419	0.298	0.210	0.161	
50	0.494	0.436	0.345	0.252	0.157	
100	0.479	0.445	0.405	0.323	0.151	

Cuadro 3.8: Percentiles al 5 % bajo $z_t = (1 - B)(1 - \hat{\theta}B)a_t$ (a) \mathcal{SL}_n

(b) \mathcal{GM}_n									
		θ							
п	0	0.5	0.8	0.9	1				
20	0.480	0.467	0.373	0.289	0.214				
30	0.463	0.451	0.379	0.287	0.188				
50	0.481	0.455	0.403	0.317	0.170				
100	0.471	0.456	0.429	0.362	0.157				

(c)	Τ	\mathcal{R}_n
(\mathbf{c})	'	, <i>v</i> _n

			θ		
п	0	0.5	0.8	0.9	1
20	0.480	0.480	1.563	1.886	2.004
30	0.463	0.469	1.935	2.645	3.003
50	0.481	0.477	2.263	3.732	5.003
100	0.471	0.468	2.415	4.935	10.003

(d) $\mathcal{L}\mathcal{N}$	1 n
------------------------------	------------

	(a) \mathcal{LM}_n									
		θ								
п	0	0.5	0.8	0.9	1					
20	0.480	0.521	1.481	1.863	2.003					
30	0.463	0.506	1.801	2.612	3.003					
50	0.481	0.485	2.039	3.683	5.003					
100	0.471	0.470	2.028	4.798	10.003					

Capítulo 4

Contrastes óptimos de no invertibilidad estacional

4.1. Introducción

En el capítulo anterior se desarrolla la teoría óptima para contrastar una raíz unitaria regular en un modelo general. Con este fin se plantean distintos contrastes óptimos de no invertibilidad como el de Saikkonen y Luukkonen (1993) que desarrolla el estadístico exacto, el de Leybourne y McCabe (1994) que emplea los residuos condicionados, o el de Tam y Reinsel (1997) que utiliza los residuos obtenidos por el método de retrovisión. En la revisión que se realiza de estos estadísticos se proponen formulaciones alternativas novedosas que permiten encontrar las similitudes entre estos contrastes que han sido derivados en modelos ARIMA y estructural, aparentemente distintos, a través de formulaciones matriciales convenientes que se proponen en esta tesis. Además, se propone un estadístico alternativo a los anteriores basado en los residuos exactos (\mathcal{GM}). El estadístico \mathcal{GM} tiene algunas caraterísticas que hacen recomendable su uso. Por una parte, su fórmula requiere menor coste computacional que otros estadísticos; en segundo lugar se puede aplicar en un modelo generalizado en el que se permite la existencia de otras raices unitarias. Finalmente su funcionamiento es significativamente mejor que otros estadísticos que satisfacen la misma hipótesis de contraste.

Una lógica extensión de estos contrastes son los modelos ARIMA estacionales. Por este motivo, en este capítulo se derivan los contraste para raíces unitarias en modelos estacionales, en primer lugar para la hipótesis nula de una raíz unitaria estacional y después, para raíces unitarias en las frecuencias estacionales.

Estos contrastes de raíces unitarias en modelos estacionales también han sido ampliamente desarrollados en la literatura, uno de los referentes es el propuesto por Tam y Reinsel (1997) que siguieron las aproximaciones de Tanaka (1990) y Saikkonen y Luukkonen (1993) para derivar el contraste LBIU de no invertibilidad en un modelo $MA(1)_s$ puro y que denominaron contraste LBIU de raíz MA estacional unitaria, donde ahora

$$\theta_0(B) = (1 - \theta_0 B^s).$$

También extendieron su contraste LBIU de no invertibilidad al modelo MA(1)_s impuro que puede contemplarse como un modelo ARMA(p, q + s), en donde el orden q + s indica que el modelo contiene un factor MA(1)_s. En este contexto sólo consideraron una corrección paramétrica para el estadístico de contraste algo más simple que la propuesta por Saikkonen y Luukkonen (1993). En un trabajo posterior, Tam y Reinsel (1998) derivaron el contraste LBIU de no invertibilidad en un modelo MA(1)_s estacional con término constante. Los resultados de Tam y Reinsel (1997, 1998) para el proceso MA(1)_s estacional son especialmente interesantes porque pueden particularizarse de forma trivial para el proceso MA(1) regular y pueden extenderse a otros casos especiales que se tratan en este capítulo. Además, una aportación interesante de este capítulo es que realiza la extensión, a un modelo estcional, de contrastes tan relevantes como Saikkonen y Luukkonen (1993) o Leybourne y McCabe (1994) que sin embargo aún sido tratados de forma específica en la literatura existente.

A pesar que el contraste de Tam y Reinsel (1997) es referente dentro de los contrastes para una raíz MA estacional ha sido criticado puesto que supone el contraste conjunto de *s* raíces unitarias al mismo tiempo. Sin embargo, si el interés es realizar un contraste de hipótesis sobre uno de los factores que contiene la descomposición de $(1 - \theta_0 B^s)$, es necesario aplicar un contraste distinto. En este sentido, en los modelos estructurales y teniendo en cuenta la mayor simplicidad que ofrece la aditividad de sus componentes se comtempla la posibilidad de contrastar que algunos ciclos estacionales, o frecuencias armónicas sean estocásticos. Siguiendo este argumento, en este capítulo se derivan los contraste para detectar raíces unitarias sobre dos polinomios MA;

$$\theta_0(B) = (1 + \theta B + \dots + \theta^{1/s} B^{1/s}),$$
(4.1)

donde s-1 raíces dependen de θ y

$$\theta_0(B) = (1 - 2c_k \theta^{1/2} B + \theta B^2), \tag{4.2}$$

con dos raíces MA complejas. Estos polinomios provienen de la factorización de $(1 - \theta_0 B^s)$ desarrollada ampliamente en Gallego y Treadway (1995)

$$(1-\theta_0 B^s) = (1-\theta^{1/s}B)(1+\theta^{1/s}B) \prod_{k=1}^{[s/2]-1} (1-2c_k \theta^{1/s}B + \theta^{2/s}B^2).$$

En el marco de los modelos estructurales se han realizado diversos trabajos de referencia como Canova y Hansen (1995), Caner (1998), Harvey y Busetti (2003) o Taylor (2003) en los que desarrollan estos contrastes para detectar raíces unitarias en polinomios como los descritos en (4.1) y (4.2), incluyendo diversas correcciones paramétricas y no paramétricas que a su vez están consideradas las extensiones a un modelo estacional de los trabajos de Kwiatkowski et al. (1992) y Leybourne y McCabe (1994). El primer trabajo de referencia es el de Canova y Hansen (1995) en el que se desarrolla un contraste para las frecuencias estacionales y una corrección no paramétrica para recoger autocorrelación. En Harvey y Busetti (2003) se da un paso más incorporando una tendencia lineal en el modelo, una estimación de la varianza de la perturbación no paramétrica y dos correcciones paramétricas basadas en los residuos suavizados y en los de predicción. Taylor (2003) completa los trabajos anteriores incorporando un estadístico de contraste y su correspondiente estimación de la varianza consistente en un modelo que contiene tendencias lineales estacionales.

Sin embargo, el contraste para determinar la existencia de raíces unitarias en una o varias frecuencias en un modelo estructural requiere que el resto de componentes estacionales del modelo de referencia, que no son objeto del contraste de hipótesis, deben ser deterministas. Esto plantea la conveniencia de aplicar estos contrastes puesto que la existencia de otra raíz unitaria que no se incluya en en la hipótesis nula a contrastar puede distorsionar de forma clara la potencia del contraste. Para evitar este hecho, la representación ARIMA puede resultar conveniente para realizar un contrate en un contexto estacional puesto que no exige que el resto de componentes sean deterministas satisfaciendo la misma hipótesis de contraste. A pesar de que no se han desarrollado este tipo de contrastes para frecuencias estacionales de forma explícita en el contexto de los modelos ARIMA, esto es, el desarrollo específico de contrastes para raíces unitarias en (4.1) y (4.2), sí se pueden encontrar referencias en trabajos como el de Taylor (2003) que desarrolla de forma pormenorizada los contrastes para cada una de las frecuencias. Por este motivo es de interés derivar contrastes para el ciclo estacional en el contexto de los modelos ARIMA, que sobre la base de los artículos de referencia que deben ser Tam y Reinsel (1997) y Tam y Reinsel (1998) permitan contrastar una hipótesis nula de estacionalidad determinista sobre una o varias de las frecuencias estacionales.

Por tanto, en este capítulo se desarrollan los estadísticos de contraste basados en los residuos exactos (\mathcal{GM}) para el ciclo estacional en modelos ARIMA así como sus distribuciones límite sin las restricciones que se impone en los modelos estructurales. De nuevo se destaca la oportunidad de la utilización de este estadístico por varios motivos; no require de un cálculo complejo, se puede aplicar en presencia una raíz MA extra y su funcionamiento en razonable.

El capítulo se organiza de la siguiente forma, en la primera parte se desarrollan las extensiones de los contrastes del apartado anterior al marco del MA(1)_s, fundamentalmente para los tres contrastes de Nyblom y Mäkeläinen (1983), Saikkonen y Luukkonen (1993) Tam y Reinsel (1997) y sus correciones para recoger autocorrelación y un estudio sobre la potencia del contraste de Tam y Reinsel (1997). A continuación, se derivan los contrastes para el ciclo estacional, en primer lugar en el marco ARIMA, y en segundo lugar empleando una representación estructural. A continuación se propone la extensión a un modelo ARIMA general, y finalmente se presenta una evaluacion de los contrastes SL y GM para una raíz unitaria estacional en presencia de otra raíz MA extra.

4.2. Contrastes de no invertibilidad en un modelo MA(1)_s

Como se ha señalado en la introducción del presente capítulo, una de las extensiones de los contrastes desarrollados en el capitulo anterior es considerar un modelo MA(1) estacional. Siguiendo el enfoque del capítulo previo en este apartado, y de forma muy resumida dado que muchos de los desarrollos son los mismos que para los contrastes en de raíces unitarias en un modelo regular, se derivan los contrastes de referencia para un MA(1) estacional. La referencia en este caso el estadístico de Tam y Reinsel (1997) que siguió las aproximaciones de Tanaka (1990) y Saikkonen y Luukkonen (1993), descritas en las dos secciones anteriores en el marco de los modelos ARIMA. También se deriva la extensión del contraste de Nyblom y Mäkeläinen (1983) a un modelo de nivel local estacional.

A pesar de que estos contrastes son referentes, y muy citados en artículos posteriores, no tienen extensión a un contraste para una raíz unitaria en un modelo MA estacional. Por esto, se propone las extensiones de los estadísticos de Nyblom y Mäkeläinen (1983) y Saikkonen y Luukkonen (1993).

4.2.1. El contraste de Tam y Reinsel

En la extensión estacional del contraste de Tanaka (1990), Tam y Reinsel (1997) consideraron el modelo $MA(1)_s$ estacional

$$z_t = (1 - \Theta B^s)a_t, \quad t = s + 1, \dots, T,$$

que puede escribirse en forma matricial como

$$\mathbf{z} = \boldsymbol{\nabla}_{s}(\boldsymbol{\Theta})\mathbf{a},\tag{4.3}$$

en donde $\mathbf{z} = (z_{k+1} \ z_{k+2} \ \dots \ z_T)'$ es un vector columna de T - s observaciones estacionarias, $\mathbf{a} = (a_1 \ a_2 \ \dots \ a_T)'$ es un vector columna de T errores aleatorios y $\nabla_s(\Theta)$ es la submatriz formada por las últimas T - s filas de la matriz cuadrada de orden T, $\mathbf{D}_s(\Theta) = \mathbf{D}(\Theta) \otimes \mathbf{I}_s$, que tiene unos en la diagonal principal, $-\Theta$ en la *s*-ésima superdiagonal y ceros en el resto y que también se puede denotar como $\mathbf{D}_s(\Theta) = \mathbf{I}_T - \Theta \mathbf{B}_T^s$. Bajo el supuesto normalidad **a** ~ $N(\mathbf{0}, \sigma_a^2 \mathbf{I}_T)$ se tiene que **z** ~ $N(\mathbf{0}, \sigma_a^2 \mathbf{\Omega}_s(\Theta))$, siendo

$$\mathbf{\Omega}_{s}(\Theta) = \mathbf{\nabla}_{s}(\Theta)\mathbf{\nabla}_{s}(\Theta)' = (1+\Theta^{2})\mathbf{I}_{T-s} - \Theta(\mathbf{B}_{T-s}^{s} + \mathbf{B}_{T-s}'^{s}) = \mathbf{\Omega}(\theta) \otimes \mathbf{I}_{s},$$

en donde $\Omega(\theta)$ es la matriz tridiagonal definida en (3.33). Así pues, el contraste LBIU para $H_0: \Theta = 0$ frente a $H_1 = \Theta < 1$ y que es la versión estacional de (3.35) es

$$\mathcal{T}_{s,n} = \frac{\mathbf{z}' \mathbf{\Omega}_s(1)^{-2} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_s(1)^{-1} \mathbf{z}}.$$
(4.4)

La distribución de exacta de $\mathcal{T}_{s,n}$ es

$$\mathcal{T}_{s,n} = \frac{\sum_{k=1}^{n-1} [\lambda_k(1)^2 / \lambda_k(\Theta)] \xi_k}{\sum_{k=1}^m [\lambda_k(1) / \lambda_k(\Theta)] \xi_k},\tag{4.5}$$

en donde $\xi_k = \mathbf{a}'(\mathbf{q}_k \mathbf{q}'_k \otimes \mathbf{I}_s)\mathbf{a}/\sigma_a^2 \sim \operatorname{iid}\chi_s^2$ y los autovalores $\lambda_k(\Theta)$ de $\mathbf{\Omega}_s(\Theta)^{-1} = \mathbf{\Omega}(\Theta)^{-1} \otimes \mathbf{I}_s$ son los de la matriz $\mathbf{\Omega}(\Theta)^{-1}$ con multiplicidad *s*. De aquí, cuando $n \to \infty$, la distribución límite de $\mathcal{T}_{s,n}/n$ bajo la secuencia de alternativas locales $H_1 : \Theta = 1 - c/n$ viene dada por

$$T_{s/n}/n \sim \frac{1}{s} \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4} \right) \xi_k.$$
 (4.6)

La demostración es idéntica a la descrita para el caso regular y simplemente requiere un mero cambio notacional consistente en remplazar las matrices *regulares* por las *estacionales*.

Tam y Reinsel (1997) expresaron también la distribución límite en términos de un funcional de movimientos brownianos

$$\mathcal{T}_{s,n}/n \to \frac{1}{s} \sum_{j=1}^{s} \int_{0}^{1} \left[V_{j}(r) + cV_{j}^{*}(r) \right]^{2} dr,$$
(4.7)

en donde $V_j(r) = W_j(r) - rW_j(1)$ es un puente browniano y $V_j^*(r) = \int_0^1 W_j(s)ds - r \int_0^1 W_j(s)ds$ es un movimiento browniano integrado, siendo $W_1(r)$, ..., $W_s(r)$ procesos Brownianos estándar mutuamente independientes.

También Tam y Reinsel (1997) extendieron el contraste LBIU de no invertibilidad esta-

cional al modelo MA(1)_s impuro

$$z_t = (1 - \Theta B^s)b_t, \quad t = s + 1, \dots, T,$$

$$\phi(B)b_t = \theta(B)a_t,$$
(4.8)

que expresaron como un modelo ARMA(p,q+s)

$$\phi(B)z_t = \theta(B)(1 - \Theta B^s)a_t, \quad t = s + 1, \dots, T.$$
(4.9)

Sin embargo, a diferencia de Saikkonen y Luukkonen (1993) siguieron una aproximación similar a la de Leybourne y McCabe (1994) consistente en el filtrado de la serie estacionaria para reducir el proceso ARMA(p,q+s) a un modelo $MA(1)_s$. La serie filtrada la obtienen de la representación matricial del proceso usada en la estimación de máxima verosimilitud exacta

$$\mathbf{\Phi}_m \mathbf{z}_m = \mathbf{\Theta}_m (\mathbf{\nabla}_{n-1,n}(\mathbf{\Theta}) \otimes \mathbf{I}_s) \mathbf{a}_T + \mathbf{F}_m \mathbf{z}_p^{(0)} + \mathbf{T}_m \mathbf{b}_q^{(0)},$$

en donde m = T - s, $\mathbf{\Phi}_m = \mathbf{I}_m - \phi_1 \mathbf{B}_m - \dots - \phi_p \mathbf{B}_m^p$ y $\mathbf{\Theta} = \mathbf{I} - \theta_1 \mathbf{B} - \dots - \theta_q \mathbf{B}^q$ son matrices banda de orden m, $\mathbf{z}_0 = (w_{s-p+1}, \dots, w_s)'$, $\mathbf{\Phi}_0 = [\phi_{p+i-j}]$ es una matriz $(T - s) \times p$ con $\phi_{p+i-j} = 0$ para p + i - j > p, $\mathbf{b}_0 = (a_{s-q+1} - a_{-q+1}, \dots, a_s - \Theta a_0)'$, $\mathbf{\Theta}_0 = [-\theta_{q+i-j}]$ es una matriz $(T - s) \times q$ con $\theta_{1+i-j} = 0$ para q + i - j > q. A partir de aquí proponen dos series filtradas. Por un lado, $\mathbf{z}_{*,n} = \mathbf{\Theta}^{-1}\mathbf{\Phi}\mathbf{w}$, que ignora los valores premuestrales; y por otro, $\mathbf{z}_{**,n} = \mathbf{\Theta}^{-1}\mathbf{\Phi}\mathbf{w} - \mathbf{\Theta}^{-1}\mathbf{\Phi}_0\mathbf{w}_0 - \mathbf{\Theta}^{-1}\mathbf{\Theta}_0\mathbf{b}_0$ que tiene en cuenta los valores premuestrales que habría que obtener usando el método de retrovisión. En cualquier caso, el estadístico ajustado $\mathcal{T}_{s,n}^*/n$ o $\mathcal{T}_{s,n}^{**}/n$ se obtiene de (4.4) reemplazando \mathbf{z} por \mathbf{z}_* o \mathbf{z}_{**} , respectivamente. La distribución límite de $\mathcal{T}_{s,n}^*/n$ bajo la secuencia de alternativas $H_1 : \mathbf{\Theta} = 1 - c/n$ es (4.7).

4.2.2. La aproximación de Saikkonen-Luukkonen

El modelo $IMA(1,1)_s$ estacional

$$\nabla_s y_t = (1 - \Theta B^s) a_t, \quad t = 1, \dots, T, \tag{4.10}$$

puede escribirse en forma matricial como

$$[\mathbf{D}_n \otimes \mathbf{I}_s]\mathbf{y}_T = [\mathbf{d}_n \otimes \mathbf{I}_s][\mathbf{y}_s^{(0)} - \Theta \mathbf{a}_s^{(0)}] + [\mathbf{D}_n(\Theta) \otimes \mathbf{I}_s]\mathbf{a}_T,$$
(4.11)

en donde $[\mathbf{y}_s^{(0)} - \Theta \mathbf{a}_s^{(0)}] = [y_{1-s} - \Theta a_{1-s}, \dots, y_0 - \Theta a_0]'$ es un vector $s \times 1$ de valores premuestrales que se suponen no estocásticos, y el resto de símbolos se han definido previamente. Premultiplicando (4.11) por la inversa $[\mathbf{C}_n \otimes \mathbf{I}_s]$ de $[\mathbf{D}_n \otimes \mathbf{I}_s]$, se obtiene

$$\mathbf{y}_T = [\mathbf{i}_n \otimes \mathbf{I}_s] [\mathbf{y}_s^{(0)} - \Theta \mathbf{a}_s^{(0)}] + [\mathbf{C}_n \mathbf{D}_n(\Theta) \otimes \mathbf{I}_s] \mathbf{a}_T,$$
(4.12)

que es un caso especial del modelo de regresión lineal generalizado (2.1) con matriz de diseño $\mathbf{X}_{T,s} = [\mathbf{i}_n \otimes \mathbf{I}_s]$ formada por *s* variables ficticias estacionales, vector de coeficientes de regresión $\boldsymbol{\beta}_s = [\mathbf{y}_s^{(0)} - \Theta \mathbf{a}_s^{(0)}]$ y término de error $\mathbf{e}_T = [\mathbf{C}_n \mathbf{D}_n(\Theta) \otimes \mathbf{I}_s] \mathbf{a}_T \sim N[\mathbf{0}, \sigma_a^2 \Omega_T(\Theta)] \operatorname{con} \Omega_T(\Theta) = [\mathbf{C}_n \mathbf{D}_n(\Theta) \mathbf{D}_n(\Theta)' \mathbf{C}'_n \otimes \mathbf{I}_s].$

El estadístico LBIU para el contraste de H_0 : $\Theta = 1$ frente a la alternativa bilateral $H_1 : \Theta \neq 1$ viene dado por

$$SL_{s,n} = \frac{\hat{\mathbf{e}}_T' [\mathbf{C}_n \mathbf{B}_n \mathbf{B}_n' \mathbf{C}_n' \otimes \mathbf{I}_k] \hat{\mathbf{e}}_T}{\hat{\mathbf{e}}_T' \hat{\mathbf{e}}_T}, \qquad (4.13)$$

donde **ê** son los residuos de mínimos cuadrados ordinarios de **y** sobre s variables ficticias estacionales.

La extensión estacional del estadístico LBIU modificado de Saikkonen y Luukkonen (1993) se obtiene escribiendo (4.8) en forma matricial

$$\mathbf{y}_T = (\mathbf{i}_n \otimes \mathbf{I}_s) \mathbf{y}_s^{(0)} + [\mathbf{C}_n \mathbf{D}_n(\Theta) \otimes \mathbf{I}_s] \mathbf{b}_T,$$

en donde $\mathbf{b}_T = (b_1, b_2, \dots, b_T)' \sim N(\mathbf{0}, \mathbf{\Omega}_{b,T})$. Definiendo $\mathbf{\Omega}_{b,T} = \sigma_a^2 \mathbf{\Omega}_{b,T}$ y $\mathbf{e}_T = [\mathbf{C}_n \mathbf{D}_n(\Theta) \otimes \mathbf{I}_s] \mathbf{b}_T$, se tiene que $\mathbf{e}_T \sim N(\mathbf{0}, \sigma_a^2 \mathbf{\Omega}_T(\Theta))$ con

$$\mathbf{\Omega}_{T}(\Theta) = [\mathbf{C}_{n}\mathbf{D}_{n}(\Theta) \otimes \mathbf{I}_{s}]\mathbf{\Omega}_{b,T}[\mathbf{D}_{n}(\Theta)'\mathbf{C}_{n}' \otimes \mathbf{I}_{s}]$$

у

$$\frac{\partial \mathbf{\Omega}_T(\Theta)^2}{\partial \Theta^2}\Big|_{\Theta=1} = 2\left(\mathbf{C}_n \mathbf{B}_n \otimes \mathbf{I}_s\right) \mathbf{\Omega}_{b,T} \left(\mathbf{B}'_n \mathbf{C}'_n \otimes \mathbf{I}_s\right).$$

Por el teorema 3, el estadístico LBIU para el contraste de H_0 : $\Theta = 1$ frente a H_1 : $\Theta \neq 1$ en (4.8) viene dado por

$$SL_{s,n}^{*} = \frac{\hat{\mathbf{e}}_{*,T}^{\prime} \mathbf{C}_{*,T} \mathbf{B}_{T}^{s} \mathbf{\Omega}_{b,T} \mathbf{B}_{T}^{\prime s} \mathbf{C}_{*,T}^{\prime} \hat{\mathbf{e}}_{*,T}}{\hat{\mathbf{e}}_{*,T}^{\prime} \hat{\mathbf{e}}_{*,T}}$$
(4.14)

en donde $\hat{\mathbf{e}}_{*,T} = \mathbf{M}_{*,T}\mathbf{y}_{*,T}, \mathbf{y}_{*,T} = \mathbf{\Omega}_{b,T}^{-1/2}\mathbf{y}_{T}, \mathbf{M}_{*,T} = \mathbf{I}_{T} - \mathbf{X}_{*,T,s}(\mathbf{X}'_{*,T,s}\mathbf{X}_{*,T,s})^{-1}\mathbf{X}_{*,T,s}',$ $\mathbf{X}_{*,T,s} = \mathbf{\Omega}_{b,T}^{-1/2}(\mathbf{i}_{n} \otimes \mathbf{I}_{s}) \text{ y } \mathbf{C}'_{*,T} = \mathbf{\Omega}_{b,T}^{-1/2}(\mathbf{C}_{n} \otimes \mathbf{I}_{s}).$ La expresion del estadístico en función de sus autovalores es la siguiente, en la que se considera $\nabla_{\#,m,T}(\Theta) = \mathbf{\Omega}_{b,T-s}^{-1/2} \nabla_{m,T}(\Theta) \mathbf{\Omega}_{b,T}^{1/2},$

$$SL_{s,n}^{*} = \frac{\sum_{k=1}^{n-1} [\lambda_{k,T}^{\#}(\Theta) / \lambda_{k,T}^{\#}(1)^{2}] \xi_{k}}{\sum_{k=1}^{m} [\lambda_{k,T}^{\#}(\Theta) / \lambda_{k,T}^{\#}(1)] \xi_{k}},$$
(4.15)

y donde $\lambda_{k,T}^{\#}(\Theta)$ son los recíprocos de los autovalores de la matriz $\nabla_{\#,m,T}(\Theta) \nabla_{\#,m,T}(\Theta)'$ y $\xi_k \sim iid\chi_s$, por lo que la distribución límite es entonces

$$SL_{s,n}^*/n \sim \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4}\right) \xi_k.$$
 (4.16)

4.2.3. La aproximación de Nyblom y Mäkeläinen

Se considera aquí la versión estacional del modelo de nivel local (3.1) definida por la ecuación

$$y_t = \mu_t + u_t,$$

 $\mu_t = \mu_{t-s} + v_t, \quad t = 1, \dots, T,$
(4.17)

en donde la serie estacional observada y_t se expresa como la suma de un paseo aleatorio estacional μ_t y un proceso de ruido blanco u_t . Se mantienen los supuestos sobre los errores u_t y v_t , así como el de condiciones iniciales μ_{-s+1} , μ_{-s+2} , ..., μ_0 no-estocásticas. Obviamente, cuando s = 1 se cumple que T = n y el modelo (4.17) se reduce a (3.1).

La forma matricial de (4.17), o versión estacional de (3.2), es

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{u},$$

$$\mathbf{D}_{s}\boldsymbol{\mu} = (\mathbf{i}_{(1)} \otimes \mathbf{I}_{s})\boldsymbol{\mu}_{0} + \mathbf{v},$$
(4.18)

en donde $\mathbf{y} = (y_1 \ y_2 \ \dots \ y_T)', \ \boldsymbol{\mu} = (\mu_1 \ \mu_2 \ \dots \ \mu_T)', \ \mathbf{u} = (u_1 \ u_2 \ \dots \ u_T)' \ \mathbf{y} \ \mathbf{v} =$

 $(v_1 \ v_2 \ \dots \ v_T)'$ son ahora vectores columna de orden T, $\mu_0 = (\mu_{-s+1} \ \mu_{-s+1} \ \dots \ \mu_0)'$ es el vector $s \times 1$ de condiciones iniciales, \mathbf{D}_s es una matriz cuadrada bidiagonal de orden T con todos sus elementos iguales a 1 en la diagonal principal y a -1 en la *s*ésima subdiagonal, y \otimes denota el producto de Kronecker. Expresando $\mathbf{D}_s = \mathbf{D} \otimes \mathbf{I}_s$ y denotando su inversa por $\mathbf{C}_s = \mathbf{D}_s^{-1}$, se cumple que $\mathbf{C}_s = (\mathbf{C} \otimes \mathbf{I}_s)$, en donde \mathbf{D} y \mathbf{C} son las matrices cuadradas orden n definidas en (3.3) y (3.4), respectivamente.

Por el teorema 2, el estadístico LBI para este problema de decisión es

$$\mathcal{NM}_{s,n} = \frac{\hat{\mathbf{e}}'\mathbf{C}_s\mathbf{C}'_s\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}},\tag{4.19}$$

en donde $\hat{\mathbf{e}}$ es el vector de residuos minimocuadráticos en la regresión de y_t sobre el conjunto completo de variables ficticias estacionales, residuos que se corresponden con los datos centrados de la serie respecto de las medias especificas en cada estación. Para expresar el numerador y denominador de $\mathcal{NM}_{s,n}$ en forma sumatoria, conviene indexar las observaciones en términos del periodo temporal $\tau = 1, 2, ..., n$ y la estación j = 1, 2, ..., s. Denotando por $y_{\tau j}$ la observación de la serie en la estación j del período τ y por $\bar{y}_j = \sum_{\tau=1}^n y_{\tau j}/n$, se tiene que $\mathbf{C}'_s \hat{\mathbf{e}} = [\sum_{\tau=t}^n (y_{\tau j} - \bar{y}_j)]$, $\mathbf{C}_s \hat{\mathbf{e}} =$ $[\sum_{\tau=1}^n (y_{\tau j} - \bar{y}_j)]$ y $\hat{\mathbf{e}}' \mathbf{C}_s \mathbf{C}'_s \hat{\mathbf{e}} = \hat{\mathbf{e}}' \mathbf{C}'_s \mathbf{C}_s \hat{\mathbf{e}}$, de manera que $\mathcal{NM}_{s,n}$ puede escribirse de dos formas equivalentes

$$\mathcal{NM}_{s,n} = \frac{\sum_{\tau=1}^{n} \left[\sum_{t=\tau}^{n} (y_{tj} - \bar{y}_j) \right]^2}{\sum_{\tau=1}^{n} (y_{\tau j} - \bar{y}_j)^2} = \frac{\sum_{\tau=1}^{n} \left[\sum_{t=1}^{\tau} (y_{tj} - \bar{y}_j) \right]^2}{\sum_{\tau=1}^{n} (y_{\tau j} - \bar{y}_j)^2}.$$
(4.20)

La distribución exacta de $\mathcal{NM}_{s,n}$ puede obtenerse siguiendo un desarrollo similar al descrito para \mathcal{NM}_n , que se basa en la relación aquí observada entre la matriz de proyección $\mathbf{M}_s = \mathbf{I}_T - (\mathbf{i} \otimes \mathbf{I}_s)[(\mathbf{i} \otimes \mathbf{I}_s)]^{-1}(\mathbf{i} \otimes \mathbf{I}_s)' = \mathbf{M}_1 \otimes \mathbf{I}_s$ en la regresión estacional y la matriz de diferencias estacionales

$$\mathbf{M}_{s} = \boldsymbol{\nabla}_{s}^{\prime} (\boldsymbol{\nabla}_{s} \boldsymbol{\nabla}_{s}^{\prime})^{-1} \boldsymbol{\nabla}_{s} = [\boldsymbol{\nabla}^{\prime} (\boldsymbol{\nabla} \boldsymbol{\nabla}^{\prime})^{-1} \boldsymbol{\nabla}] \otimes \mathbf{I}_{s}.$$
(4.21)

Además, se advierte que

$$\nabla_s \mathbf{C}_s \mathbf{C}'_s \nabla'_s = \nabla \mathbf{C} \mathbf{C}' \nabla' \otimes \mathbf{I}_s = \mathbf{I}_{n-1} \otimes \mathbf{I}_s = \mathbf{I}_{T-s}.$$

Estas identidades matriciales permiten expresar (4.19) como

$$\mathcal{NM}_{s,n} = \frac{\mathbf{y}' \mathbf{\nabla}'_s (\mathbf{\nabla}_s \mathbf{\nabla}'_s)^{-2} \mathbf{\nabla}_s \mathbf{y}}{\mathbf{y}' \mathbf{\nabla}'_s (\mathbf{\nabla}_s \mathbf{\nabla}'_s)^{-1} \mathbf{\nabla}_s \mathbf{y}'},\tag{4.22}$$

que es la versión estacional de (3.14). Teniendo en cuenta que λ_k y \mathbf{p}_k son los autovectores y autovalores de $(\nabla_s \nabla'_s)^{-1}$ y definiendo $\xi_k = \epsilon' (\mathbf{p}_k \mathbf{p}'_k \otimes \mathbf{I}_s)' \epsilon / [\sigma_u^2 (1 + \rho \lambda_k)] \sim$ $iid\chi_s^2$, se obtiene

$$\mathcal{NM}_{s,n} = \frac{\sum_{k=1}^{n-1} \lambda_k (1 + \rho \lambda_k) \xi_k}{\sum_{k=1}^{n-1} (1 + \rho \lambda_k) \xi_k},$$
(4.23)

en donde $\mathcal{NM}_{s,n}$ aparece expresado como una ratio de dos combinaciones lineales de variables χ_s^2 mutuamente independientes.

Cuando $n \to \infty$, la distribución límite de $\mathcal{NM}_{s,n}/n$ bajo la secuencia de alternativas $H_1: \rho = c^2/n^2$ es

$$\mathcal{NM}_{s,n}/n \sim \frac{1}{s} \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4} \right) \xi_k,$$
 (4.24)

en donde el divisor *s* aparece porque el denominador $\sum_{k=1}^{n} \xi_k / n$ converge en probabilidad a *s*.

El modelo de media local estacional se puede extender a un modelo generalizado empleando la aproximación de Leybourne y McCabe (1994)

$$\phi(B)y_t = \mu_t + u_t,$$

$$\mu_t = \mu_{t-s} + v_t, \quad t = 1, \dots, T,$$
(4.25)

siendo $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ un polinomio AR de orden *p* estacionario.

El procedimiento de Leybourne y McCabe (1994) consiste modificar el estadístico de contraste (4.19) considerando la serie filtrada $y_t^* = \phi(B)y_t$ como un proceso de nivel local estacional. Teniendo en cuenta que la estimación por máxima verosimilitud exacta del modelo sin restringir produce estimaciones $\tilde{\phi}_i$ (i = 1, ..., p) consistentes de los parámetros autorregresivos, se utiliza su forma reducida equivalente al proceso media local estacional y que es un modelo ARIMA(p, s, s) tal que

$$\phi(B)\nabla_s y_t = (1 - \Theta B^s)a_t. \tag{4.26}$$

El procedimiento de contraste es el mismo que el expuesto en el Capítulo 3; en primer lugar se estima por máxima verosimilitud el modelo utilizando la forma reducida, después se construye la serie filtrada centrada y finalmente se calcula el estadístico que se obtiene sustituyendo y_t por la serie filtrada centrada \bar{y}_t^* en (4.20)

$$\mathcal{NM}_{s,n} = \frac{\sum_{\tau=1}^{n} \left(\sum_{t=\tau}^{n} \bar{\tilde{y}}_{tj}^{*} \right)^{2}}{\sum_{\tau=1}^{n} (\bar{\tilde{y}}_{\tau j}^{*})^{2}}.$$
(4.27)

4.2.4. Valores críticos y potencia del contraste

En este apartado se evalúan la potencia del estadístico, la del contraste óptimo puntual bajo distintas alternativas locales $\Theta_1 = 1 - c_1/n$, así como la potencia envolvente.

En el cálculo de los valores críticos se utiliza el resultado (4.5) y para un conjunto de alternativas locales, $\Theta = 1 - c/n$, tenemos que

$$P(\mathcal{T}_{s,n} > x) = P\left[\sum_{k=1}^{n-1} \lambda_k(\Theta) \left(\frac{1}{n\lambda_k(1)^2} - \frac{x}{\lambda_k(1)}\right) \xi_s^2 < 0\right].$$
(4.28)

Bajo H_0 : $\Theta = 1$ la expresión (4.28) se reduce a

$$P(\mathcal{T}_{s,n} > x) = P\left[\sum_{k=1}^{n-1} \left(x - \frac{1}{n\lambda_k(1)}\right)\xi_s^2 > 0\right],$$

que se puede evaluar empleando métodos numéricos para cualquier valor crítico *x* con procedimientos como el desarrollado en Imhof (1961).

El Cuadro 4.1 recoge los valores críticos x del contraste $\mathcal{T}_{s,n}$ para distintos tamaños de muestra, $n = \{11, 21, 31, 101, 301, \infty\}$, para distintos valores de s, tal que $s = \{1, 4, 12\}$ y para $c_1 = \{3, 7, 14\}$.

Para evaluar la distribución del estadístico de contraste $\mathcal{T}_{s,n}(\Theta_1)$ considerando $\Theta_1 = 1 - c_1/n$ es útil la siguiente expresión

$$P(\mathcal{T}_{s,n}(\Theta_1) > x) = P\left[\sum_{k=1}^{n-1} \lambda_s(\Theta) \left(\frac{1}{n\lambda_k(\Theta_1)} - \frac{x}{\lambda_k(1)}\right) \xi_s^2 < 0\right],$$
(4.29)

donde $\xi_s^2 \sim \chi_s^2$. El Cuadro 4.2 recoge los valores críticos *x* para un nivel de significación α =0.05, del contraste $R(\Theta_1) = n[1 - \mathcal{T}_{s,n}(\Theta_1)]$ para distintos tamaños de muestra, n =

				n				
	α	11	21	31	101	201	301	∞
	0.5	0.154	0.136	0.130	0.122	0.121	0.120	0.119
	0.3	0.228	0.206	0.199	0.189	0.187	0.186	0.184
s=1	0.2	0.291	0.266	0.258	0.246	0.244	0.243	0.241
	0.1	0.401	0.375	0.366	0.353	0.350	0.349	0.346
	0.05	0.504	0.485	0.477	0.466	0.464	0.463	0.460
	0.01	0.707	0.730	0.735	0.742	0.743	0.743	0.738
	0.5	0.186	0.169	0.163	0.155	0.153	0.153	0.151
	0.3	0.229	0.210	0.204	0.195	0.193	0.193	0.191
s=4	0.2	0.259	0.240	0.233	0.224	0.222	0.221	0.219
	0.1	0.305	0.286	0.279	0.270	0.268	0.267	0.265
	0.05	0.348	0.328	0.323	0.313	0.311	0.311	0.308
	0.01	0.434	0.421	0.416	0.410	0.407	0.407	0.404
	0.5	0.195	0.178	0.173	0.165	0.163	0.163	0.161
	0.3	0.220	0.202	0.197	0.188	0.187	0.186	0.185
s=12	0.2	0.236	0.219	0.212	0.204	0.202	0.202	0.200
	0.1	0.261	0.242	0.236	0.228	0.226	0.225	0.224
	0.05	0.282	0.264	0.258	0.249	0.247	0.246	0.245
	0.01	0.325	0.308	0.302	0.293	0.291	0.291	0.290

Cuadro 4.1: Valores críticos para $\mathcal{T}_{s,n}$, tal que $P(\mathcal{T}_{s,n} > x) = \alpha$.

 $\{11, 21, \dots, 101\}$, para distintos valores de *s*, tal que $s = \{1, 4, 12\}$ y para $c = \{3, 7, 14\}$.

dro	4.2: Va	lores cr	ticos j	para R	$(\Theta_1) =$	= n[1 -	$-\mathcal{T}_{s,n}(0)$	$[\Theta_1)], \Theta_2$	$D_1 = 1$	$-c_{1}/n$	ı (α=0
						n					
	c ₁	11	21	31	41	51	61	71	81	91	101
	$c_1 = 3$	-0.2045	-0.3796	-0.4360	-0.4626	-0.4784	-0.4889	-0.4961	-0.5017	-0.5061	-0.5093
s=1	$c_1 = 7$	-0.2850	-0.6641	-0.7700	-0.8191	-0,8476	-0.8661	-0.8791	-0.8887	-0.8961	-0.9018
	$c_1 = 14$	-2.4890	-2.9443	-2.9443	-2.9358	-2.9274	-2.9214	-2.9170	-2.9132	-2.9107	-2.9083
	$c_1 = 3$	-1.2395	-1.2718	-1.2797	-1.2831	-1.2850	-1.2862	-1.2869	-1.2875	-1.2881	-1.2884
s=4	$c_1 = 7$	-2.9483	-2.7426	-2.6759	-2.6434	-2.6241	-2.6115	-2.6025	-2.5958	-2.5905	-2.5864
	$c_1 = 14$	-10.2991	-7.0251	-6.3371	-6.0425	-5.8792	-5.7756	-5.7037	-5.6512	-5.6109	-5.5790

-1.6218 -3.3467

-7.0768

-1.6191 -3.3203

-6.9287

-1.6171

-3.3015

-6.8269

-1.6156

-3.2877

-6.7523

-1.6145

-3.2770

-6.6956

-1.6136

-3.2684

-6.6509

-1.6258 -3.3873

-7.3117

-1.6325 -3.4573

-7.7406

-1.6455

-3.6060

-8.7732

 $c_1 = 3$ $c_1 = 7$

s=12

-1.6817 -4.1351

-14.5441

C

La potencia del estadístico LBUI y el POI se pueden calcular utilizando (4.28) y (4.29). Para obtener la potencia envolvente consideramos en (4.29) que $\Theta = \Theta_1$, o bien que $c = c_1$, por lo que la potencia envolvente será la mayor potencia del contraste para un determinado punto. Los Cuadros 4.3 y 4.4 muestran las potencias de los contrastes LBUI, POI y envolvente para $\Theta = \{ 1, 0.99, 0.98, \dots, 0.90, 0.85, 0.8, 0.75 \}, n = 101$ y $\alpha = 0.05$. Por otra parte la potencia del contraste POI se calcula para tres alternativas locales $c = \{3, 7, 14\}.$
					0			
					0			
		1	0.99	0.98	0.97	0.96	0.95	0.94
s=1	Potencia envolvente	0.0500	0.0620	0.1000	0.1626	0.2414	0.3267	0.4110
	$POI(\Theta_1) c=3$	0.5000	0.0619	0.0999	0.1626	0.2409	0.3232	0.4023
	$POI(\Theta_1)$ c=7	0.0500	0.0607	0.0958	0.1567	0.2364	0.3239	0.4103
	$POI(\Theta_1)$ c=14	0.0500	0.0583	0.0864	0.1384	0.2116	0.2973	0.3863
	TR	0.0500	0.0620	0.0999	0.0617	0.2372	0.3155	0.3896
s=4	Potencia envolvente	0.0500	0.0723	0.1558	0.3130	0.5055	0.6792	0.8079
	$POI(\Theta_1) c=3$	0.5000	0.0721	0.1555	0.3130	0.5041	0.6732	0.7964
	$POI(\Theta_1)$ c=7	0.0500	0.0697	0.1460	0.2981	0.4940	0.6744	0.8070
	$POI(\Theta_1)$ c=14	0.0500	0.0651	0.1243	0.2514	0.4353	0.6246	0.7747
	TR	0.0500	0.0723	0.1556	0.3102	0.4956	0.6590	0.7800
s=12	Potencia envolvente	0.0500	0.0877	0.2595	0.5766	0.8410	0.9588	0.9919
	$POI(\Theta_1)c = 3$	0.5000	0.0878	0.2590	0.5766	0.8391	0.9562	0.9882
	$POI(\Theta_1)c = 7$	0.0500	0.0836	0.2415	0.5495	0.8302	0.9556	0.9918
	$POI(\Theta_1)c = 14$	0.0500	0.0771	0.1940	0.4628	0.7650	0.9303	0.9864
	TR	0.0500	0.0881	0.2592	0.5713	0.8309	0.9497	0.9873

Cuadro 4.3: Tamaño empírico, potencia para el estadístico $\mathcal{T}_{s,n}$, óptimo puntual $POI(\Theta_1)$ y envolvente para $\alpha = 0.05$ (n=101).

Cuadro 4.4: Tamaño empírico, potencia para el estadístico $T_{s,n}$, óptimo puntual $POI(\Theta_1)$ y envolvente para $\alpha = 0.05$ (n=101).

					Θ			
		0.93	0.92	0.91	0.90	0.85	0.80	0.75
s=1	Potencia envolvente	0.4907	0.5635	0.6286	0.6860	0.8734	0.9534	0.9837
	$POI(\Theta_1) c=3$	0.4746	0.5392	0.5661	0.6459	0.8137	0.8979	0.9416
	$POI(\Theta_1)$ c=7	0.4907	0.5628	0.6262	0.6811	0.8573	0.9350	0.9695
	$POI(\Theta_1)$ c=14	0.4718	0.5504	0.6203	0.6812	0.8736	0.9514	0.9812
	TR	0.4566	0.5161	0.5686	0.6148	0.7752	0.8626	0.9120
s=4	Potencia envolvente	0.8916	0.9412	0.9701	0.9851	0.9997	0.9998	0.9993
	$POI(\Theta_1) c=3$	0.8781	0.9279	0.9573	0.9757	0.9983	0.9999	0.9993
	$POI(\Theta_1)$ c=7	0.8916	0.9417	0.9693	0.9838	0.9988	0.9996	0.9999
	$POI(\Theta_1)$ c=14	0.8752	0.0347	0.9670	0.9831	0.9997	0.9989	0.9998
	TR	0.8610	0.9125	0.9457	0.9660	0.9963	0.9995	0.9999
s=12	Potencia envolvente	0.9987	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
	$POI(\Theta_1) c=3$	0.9980	0.9987	0.9958	0.9999	0.9997	0.9999	0.9999
	$POI(\Theta_1)$ c=7	0.9976	0.9942	0.9998	0.9988	0.9970	0.9988	0.9999
	$POI(\Theta_1) c=14$	0.9947	0.9997	0.9991	0.9959	1.0000	0.9952	0.9993
	TR	0.9970	0.9993	0.9999	1.0000	1.0000	1.0000	1.0000

4.3. Contraste de raíz MA unitaria estacional en un proceso ARIMA general

En las secciones previas se han desarrollado contrastes de raíces unitarias en modelos estacionales, tanto en modelos estructurales como en modelos ARIMA, siendo un aspecto relevante su extensión a modelos generalizados. En esta tesis se propone el uso del estadístico basado en los residuos exactos por varios motivos; tiene una expresión muy conveniente en términos computacionales, se puede aplicar en presencia de una raíz unitaria extra y además tiene un funcionamiento razonable. Su extensión al caso estacional implica un desarrollo muy similar al detallado en el capítulo anterior, en el que únicamente se modifican las matrices de acumulación del numerador del estadísti-co.

El punto de partida es el trabajo de Tam y Reinsel (1997) y su contraste para detectar una raíz unitaria estacional, así como el desarrollo propuesto en esta tesis para extender a un modelos MA(1)_s los contrastes que aquí se han desarrollado sobre los artículos de Saikkonen y Luukkonen (1993) y Nyblom y Mäkeläinen (1983). Los estadísticos de contraste básicos para $H_0: \Theta = 0$ frente a $H_0: \Theta \neq 1$ son los siguientes

$$\mathcal{NM}_{s,n} = \frac{\mathbf{y'}\mathbf{M}_s\mathbf{C}_s\mathbf{C}_s'\mathbf{M}_s\mathbf{y}}{\mathbf{y'}\mathbf{M}_s\mathbf{y}}, \quad \mathcal{T}_{s,n} = \frac{\mathbf{z'}(\nabla_s\nabla_s')^{-2}\mathbf{z}}{\mathbf{z'}(\nabla_s\nabla_s')^{-1}\mathbf{z}} \quad y \quad \mathcal{SL}_{s,n} = \frac{\mathbf{y'}\mathbf{M}_s\mathbf{C}_s\mathbf{B}_s\mathbf{B}_s'\mathbf{C}_s'\mathbf{M}_s\mathbf{y}}{\mathbf{y'}\mathbf{M}_s\mathbf{y}}.$$

También se ha presentado la extensión del estadístico de Tam y Reinsel (1997) que utiliza un filtrado de la serie empleando los residuos de un modelo ARMA(p,q) obtenidos por el método de retrovisión, y a la vez se proponen las extensiones de dos contrastes que se han desarrollado únicamente en el caso regular y que son la corrección paramétrica de Saikkonen y Luukkonen (1993) y la de Leybourne y McCabe (1994). En esta sección se amplía al caso estacional el procedimiento derivado en el apartado 3.5. para el contraste de raíz unitaria en un proceso ARIMA general, y en presencia de raíces unitarias estacionales, a un contrate para una raíz unitaria estacional.

4.3.1. El contraste para el nivel local estacional generalizado

El modelo de Nyblom y Mäkeläinen (1983) generalizado en (4.17) se puede expresar como

$$y_t = \mu_t + u_t, \quad \phi(B)u_t = \theta(B)w_{1t},$$

 $\mu_t = \mu_{t-s} + v_t, \quad \phi(B)v_t = \theta(B)w_{2t}, \quad t = 1, \dots, T,$
(4.30)

donde los supuestos son los mismos que para el modelo definido en (3.62), inluida la condición de que las raíces de $\phi(B)$ estén fuera del círculo unitari. Por el contrario se permite la existencia de raíces unitarias en el polinomio $\theta(B)$. Esta cuestión es importante puesto que la ausencia de otras raíces unitaria en el polinomio media móvil extra en $\theta(B)$ es un requisito obligatorio para poder aplicar contrastes como el Tam y Reinsel (1997).

Además, se supone que ambos polinomios no tienen raíces comunes que indicarían la presencia de factores redundantes. Los s valores premuestrales μ_0 son no estocásticos, $w_{1t} \sim \text{iid}N(0, \sigma_1^2), w_{2t} \sim \text{iid}N(0, \sigma_2^2)$ y $E(w_{1t}w_{2,t-k}) = 0 \forall k$.

La forma reducida equivalente a (4.30) es un modelo ARIMA(p, s, q + s)

$$\phi(B)\nabla_s y_t = \theta(B)(1 - \Theta B^s)a_t. \tag{4.31}$$

y análogamente a (3.64), se puede expresar como

$$abla_{s}y_{t} = v_{t} + u_{t} - u_{t-s}, \quad t = s+1, \ldots, T,$$

donde

104

$$\phi(B)\nabla_s y_t = \theta(B)[w_{2t} + (1 - B^s)w_{1t}],$$

siendo $w_{2t} + (1 - B^s)w_{1t}$ un proceso MA(1)_s con parámetro Θ y también se define el parámetro ρ , tal que $\rho = \sigma_2^2 / \sigma_1^2$.

Como en el desarrollo del capítulo anterior el modelo de regresión generalizado es

$$\mathbf{y} = (\mathbf{i} \otimes \mathbf{I}_s)\mu_0 + \mathbf{e}, \quad \mathbf{e} = \mathbf{C}_s \mathbf{v} + \mathbf{u}, \tag{4.32}$$

en donde ahora $\mathbf{u} \sim N[\mathbf{0}, \sigma_1^2 \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})], \mathbf{v} \sim N[\mathbf{0}, \sigma_2^2 \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})]$ y la matriz $\mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})$ depende del vector de parámetros $\boldsymbol{\varphi} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$. Por tanto, la distribución muestral de y viene dada por

$$\mathbf{y} \sim N\left(\mathbf{i} \otimes \mathbf{I}_{s}, \, \sigma_{1}^{2} \boldsymbol{\Omega}_{\mathbf{y},T}(\rho, \boldsymbol{\varphi})\right) \quad \text{con} \quad \boldsymbol{\Omega}_{\mathbf{y},T}(\rho, \boldsymbol{\varphi}) = \rho \mathbf{C}_{s} \boldsymbol{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \mathbf{C}_{s}' + \boldsymbol{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}), \quad (4.33)$$

 \mathbf{C}_s y \mathbf{C}'_s son intercambiables de manera que $\mathbf{\Omega}_{\mathbf{y},T}(\rho, \boldsymbol{\varphi}) = \rho \mathbf{C}'_s \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \mathbf{C}_s + \mathbf{\Omega}_{\mathbf{y},T}(\boldsymbol{\varphi})$. Para simplificar la notación se escribe la matriz $\mathbf{\Omega}_{\mathbf{y},T}(\rho, \boldsymbol{\varphi})$ de orden $T \times T$ sólo en términos del parámetro de interés ρ , $\mathbf{\Omega}_{\mathbf{y},T}(\rho)$. Alternativamente, y siguiendo una aproximación similar a la de Tanaka (1990), se multiplica (4.32) por la matriz de diferencias $\nabla_{m,T}$, recordando que m = T - s, se obtiene la forma matricial

$$\mathbf{z} = \boldsymbol{\nabla}_{m,T} \mathbf{e}, \quad \boldsymbol{\nabla}_{m,T} \mathbf{e} = \boldsymbol{\nabla}_{m,T} \mathbf{C}_s \mathbf{v} + \boldsymbol{\nabla}_{m,T} \mathbf{u},$$

donde $\nabla_{m,T} \mathbf{C}_s = [\mathbf{0} | \mathbf{I}_{T-s}]$, **z** es la serie diferenciada, y su distribución muestral es

$$\mathbf{z} \sim N\left[\mathbf{0}, \ \sigma_1^2 \mathbf{\Omega}_{\mathbf{z},m}(\rho, \boldsymbol{\varphi})\right] \quad \text{con} \quad \mathbf{\Omega}_{\mathbf{z},m}(\rho, \boldsymbol{\varphi}) = \rho \mathbf{\Omega}_{\mathbf{u},m}(\boldsymbol{\varphi}) + \boldsymbol{\nabla}_{m,T} \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \boldsymbol{\nabla}_{m,T}', \quad (4.34)$$

o en términos de Θ y σ_a^2

$$\mathbf{z} \sim N\left[\mathbf{0}, \ \sigma_a^2 \mathbf{\Omega}_{\mathbf{z},m}(\theta, \boldsymbol{\varphi})\right] \quad \text{con} \quad \mathbf{\Omega}_{\mathbf{z},m}(\Theta, \boldsymbol{\varphi}) = \boldsymbol{\nabla}_{m,T}(\Theta) \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \boldsymbol{\nabla}_{m,T}(\theta)'.$$

Una vez que el problema de contraste está definido en términos de los parámetros ρ y Θ es más conveniente, por la sencillez de su desarrollo, la obtención del estadístico LBI que emplea la primera derivada de la matriz de covarianzas Ω_z en lugar de la segunda.

4.3.2. Estadísticos de contraste

El estadístico de contraste es LBI que se obtiene aplicando el teorema 2 y tiene la siguiente región crítica,

$$\mathcal{NM}_{s,n}^* = \frac{\hat{\mathbf{e}}_{*,T}' \mathbf{C}_{*,T} \mathbf{\Omega}_{u,T} \mathbf{C}_{*,T}' \hat{\mathbf{e}}_{*,T}}{\hat{\mathbf{e}}_{*,T}' \hat{\mathbf{e}}_{*,T}}.$$
(4.35)

en donde $\hat{\mathbf{e}}'_{*,T}$ y $\mathbf{C}'_{*,T}$ tienen la definición previa del estadístico $\mathcal{SL}^*_{s,n}$.

A continuación se propone la extensión estacional del contraste \mathcal{GM}_n^* , aplicando la metodología expuesta en Tanaka (1990) a (4.34). Utilizando el corolario (2.17) para la obtención del estadístico LBI se obtiene la expresión

$$\mathcal{GM}_{s,m}^* = \frac{1}{m} \frac{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z},m}(\rho_0)^{-1} \mathbf{\Omega}_{\mathbf{u},m} \mathbf{\Omega}_{\mathbf{z},m}(\rho_0)^{-1} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z},m}(\rho_0)^{-1} \mathbf{z}}.$$
(4.36)

La equivalencia con \mathcal{NM}_n^* se puede observar teniendo en cuenta que la matriz de diferencias transformada es $\nabla_{*,m,T} = \nabla_{m,T} \Omega_{\mathbf{y},T}^{1/2}$, la matriz de proyección $\mathbf{M}_{*,T}$ puede escribirse como $\mathbf{M}_{*,T} = \nabla'_{*,m,T} (\nabla_{*,m,T} \nabla'_{*,m,T})^{-1} \nabla_{*,m,T}$ porque $\nabla_{*,m,T} \mathbf{I}_{*,T} = \nabla_{*,m,T} \Omega_{\mathbf{y},T}^{1/2} \Omega_{b,T}^{-1/2} \mathbf{X}_{*,T,s} =$ $\mathbf{0}_{T,s}$. Usando estas matrices, (4.35) puede escribirse como

$$\mathcal{GM}_{s,n}^{*} = \frac{\mathbf{y}_{*,T}^{\prime} \mathbf{\nabla}_{*,m,T}^{\prime} (\mathbf{\nabla}_{*,m,T} \mathbf{\nabla}_{*,m,T}^{\prime})^{-1} \mathbf{\nabla}_{*,m,T} \mathbf{\Omega}_{u,T} \mathbf{\nabla}_{*,m,T}^{\prime} (\mathbf{\nabla}_{*,m,T} \mathbf{\nabla}_{*,m,T}^{\prime})^{-1} \mathbf{y}_{*,T}}{\mathbf{y}_{*,T}^{\prime} \mathbf{\nabla}_{*,m,T}^{\prime} (\mathbf{\nabla}_{*,m,T} \mathbf{\nabla}_{*,m,T}^{\prime})^{-1} \mathbf{\nabla}_{*,m,T}^{\prime} \mathbf{y}_{*,T}},$$

en donde $\mathbf{y}_{*,T}$ es la transformación de MCG del modelo (4.32). Finalmente, la expresión del estadístico $\mathcal{GM}^*_{s,m}$ para la hipóteis de contraste $H_0: \Theta = 1$ frente $H_1: \Theta \neq 1$ tiene la forma

$$\mathcal{GM}_{s,m}^* = \frac{1}{m} \frac{\mathbf{z}_m' \mathbf{\Omega}_{z,m}(1)^{-1} \mathbf{\Omega}_z(0) \mathbf{\Omega}_{z,m}(1)^{-1} \mathbf{z}_m}{\mathbf{z}_m' \mathbf{\Omega}_{z,m}(1)^{-1} \mathbf{z}_m}.$$
(4.37)

4.3.3. Distribución exacta

Para obtener la distribución exacta de los estadísticos $\mathcal{GM}_{s,m}^*$ y $\mathcal{SL}_{s,m}^*$, se siguen el mismo desarrollo que en el apartado 3.5.3. Se considera la transformación de $\nabla_{m,T}$ tal que $\nabla_{\#,m,T} = \Omega_{\mathbf{u},m}^{-1/2}(\boldsymbol{\varphi}) \nabla_{m,T} \Omega_{\mathbf{u},T}^{1/2}(\boldsymbol{\varphi}) = \Omega_{\mathbf{u},m}^{-1/2}(\boldsymbol{\varphi}) \nabla_{*,m,T}$ que permite transformar \mathbf{z}_m , tal que $\mathbf{z}_{*,m} = \Omega_{\mathbf{u},m}^{-1/2}(\boldsymbol{\varphi}) \mathbf{z}_m$.La distribución de $\mathbf{z}_{*,m}$ es $\mathbf{z}_{*,m} \sim N\left(\mathbf{0}, \sigma_1^2(\boldsymbol{\rho}\mathbf{I}_m + \nabla_{\#,m,T}\nabla_{\#,m,T}')\right)$ y el estadístico se puede reescribir como

$$\mathcal{GM}_{s,m}^* = \frac{\boldsymbol{\epsilon}_m' (\boldsymbol{\nabla}_{\#,m,T} \boldsymbol{\nabla}_{\#,m,T}')^{-1} \boldsymbol{\epsilon}_m}{\boldsymbol{\epsilon}_m' \boldsymbol{\epsilon}_m}.$$
(4.38)

Teniendo en cuenta que $\Omega_{u,m}^{1/2}(\varphi)(\nabla \Omega_{u,T}(\varphi)\nabla')^{-1}\Omega_{u,m}^{1/2}(\varphi) = (\nabla_{\#,m,T}\nabla'_{\#,m,T})^{-1}$ y que la variable ϵ_m tiene la siguiente distribución

$$\boldsymbol{\epsilon}_{m} = (\boldsymbol{\nabla}_{\#,m,T} \boldsymbol{\nabla}_{\#,m,T}')^{-1/2} \mathbf{z}_{*,m}' \sim N\left(\mathbf{0}, \ \sigma_{1}^{2} (\mathbf{I}_{m} + \rho \boldsymbol{\nabla}_{\#,m,T} \boldsymbol{\nabla}_{\#,m,T}')\right).$$

Por tanto, el estadístico $\mathcal{GM}^*_{s,m}$ tiene la siguiente distribución

$$\mathcal{GM}_{s,n}^* = \frac{\sum_{k=1}^m \lambda_k (1+\rho\lambda_k) \xi_k}{\sum_{k=1}^m (1+\rho\lambda_k) \xi_k},$$
(4.39)

106

en donde los autovalores λ_k (k = 1, ..., T - s). Estos autovalores son los recíprocos de los autovalores de la matriz $\nabla_{\#,m,T} \nabla'_{\#,m,T}$. Su distribución límite es por tanto,

$$\mathcal{GM}_{s,n}^*/n \stackrel{a}{\sim} \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4}\right) \xi_k.$$
 (4.40)

donde $\xi_k \sim iid\chi_s$.

La expresion del estadístico $S\mathcal{L}_{s,n}^*$ en función de sus autovalores se obtiene siguiendo los mismos pasos que 3.5.3, sustituyendo la matrices *regulares* por las *estacionales*

$$SL_{s,n}^{*} = \frac{\sum_{k=1}^{m} [\lambda_{k,T}^{\#}(\Theta) / \lambda_{k,T}^{\#}(1)^{2}] \xi_{k}}{\sum_{k=1}^{m} [\lambda_{k,T}^{\#}(\Theta) / \lambda_{k,T}^{\#}(1)] \xi_{k}},$$
(4.41)

donde $\lambda_{k,T}^{\#}(\Theta)$ son los recíprocos de los autovalores de la matriz $\nabla_{\#,m,T}(\Theta)\nabla_{\#,m,T}(\Theta)'$ y $\xi_k \sim iid\chi_1$, por lo que la distribución límite es entonces

$$SL_{s,n}^*/n \sim \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4}\right) \xi_1.$$
 (4.42)

4.3.4. Estadísticos aproximados

Después de revisar disintos estadísticos para el contraste de una raíz unitaria estacional en un modelo generalizado, es interesante expresar dicho estadísticos de una forma homogénea, por lo que aquí se proponen expresiones de distintos estadísticos obtenidos en los apartados anteriores, basados en los residuos exactos y con expresiones de cálculo más sencillo. Para ello es conveniente expresar el modelo en términos del modelo (4.32) de MCG tal que

$$\mathbf{y}_{*,T} = (\mathbf{i}_{*,n} \otimes \mathbf{I}_s)\mu_0 + \mathbf{e}_{*,T}, \quad \mathbf{e}_{*,T} \sim N(0,\sigma_a^2 \mathbf{I}_T), \tag{4.43}$$

donde los residuos exactos se pueden reescribir como $\tilde{\mathbf{a}} = E(\mathbf{a}|\mathbf{z}') = E(\tilde{\mathbf{a}}\mathbf{z}')E(\mathbf{z}\mathbf{z}')\mathbf{z} = \mathbf{L}_{\mathbf{b}}^{s'} \nabla_{m,T}' (\nabla_{m,T} \mathbf{\Omega}_{b,T-s} \nabla_{m,T}')^{-1} \mathbf{z} = \mathbf{M}_{*,T} \mathbf{y}_{*,T} = \tilde{\mathbf{e}}_{*,T}$, que coinciden con los residuos de regresión generalizados del modelo y donde $\mathbf{z} = \mathbf{L}_{\mathbf{b}}^{s} \mathbf{a}$ es la serie estacionaria y recordando que $\mathbf{L}_{\mathbf{b}}$ es el factor de Cholesky de $\mathbf{\Omega}_{b,T-s}$. Así pues, ahora es posible reescribir de forma aproximada algunos de los estadísticos referidos anteriormente, en términos

de los residuos exactos. En primer lugar el estadístico $\mathcal{GM}^*_{s,n}$

$$\mathcal{GM}_{s,n}^* = \frac{1}{m} \frac{\tilde{\mathbf{a}}' \mathbf{C}_s \mathbf{C}'_s \tilde{\mathbf{a}}}{\tilde{\mathbf{a}}' \tilde{\mathbf{a}}}.$$
(4.44)

La corrección de Leybourne y McCabe (1994), que este capítulo se extiende a una raíz MA estacional, también se puede expresar como

$$\mathcal{LM}_{s,n}^* = \frac{1}{m} \frac{\tilde{\mathbf{a}}_0' \mathbf{M}_s \mathbf{C}_s \mathbf{C}_s' \mathbf{M}_s \tilde{\mathbf{a}}_0}{\tilde{\mathbf{a}}_0' \tilde{\mathbf{a}}_0}$$

donde $\tilde{\mathbf{a}}_0$ son los residuos condicionales del modelo (4.26), suponiendo, como ya se ha señalado anteriormente, que los valores premuestrales son nulos.

La extensión del estadístico de Tam y Reinsel (1997) detallado en la sección anterior consiste en la sustitución de la serie z por una serie filtrada z^* que proviene de un proceso ARMA(p,q), y donde los residuos del modelo se obtienen por el método de retrovisión. Dado que los residuos exactos $\tilde{a} = \nabla'_{*,m,T} (\nabla_{*,m,T} \nabla'_{*,m,T})^{-1} \nabla_{*,m,T} y_*$ son los mismos que los residuos del modelo de regresión generalizado generalizado, \tilde{e} , por los que el estadístico se puede reescribir como.

$$\mathcal{TR}_{s,n}^* = \frac{1}{m} \frac{\tilde{\mathbf{e}}' \mathbf{M}_s \mathbf{C}_s \mathbf{C}'_s \mathbf{M}_s \tilde{\mathbf{e}}}{\tilde{\mathbf{e}}'_0 \tilde{\mathbf{e}}_0}$$

4.4. Constraste de no invertibilidad para un MA(2)

Como ya se ha apuntado en la introducción, una de las críticas que recibe el contraste de Tam y Reinsel (1997) es que realiza de forma conjunta un contraste de no invertibilidad sobre s raíces unitarias al mismo tiempo.

Por tanto es de interés extender los contraste de raíces unitarias planteados en esta tésis a las frecuencias estacionales en el contexto ARIMA, y en especial del estadístico \mathcal{GM} . Algunas referencias a este tipo de contrastes se pueden encontrar en Taylor (2003), aunque el trabajo se centra en el marco de un modelo estructural.

El objetivo es contrastar distintas hipótesis nulas sobre los factores del polinomio $\theta_0(B)$

(Gallego y Treadway (1995))

$$(1 - \theta_0 B^s) = (1 - \theta^{1/s} B)(1 + \theta^{1/s} B + \theta^{2/s} B^2 + \dots + \theta^{(s-1)/s} B^{s-1})$$
$$= (1 - \theta^{1/s} B)(1 + \theta^{1/s} B) \prod_{k=1}^{[s/2]-1} (1 - 2c_k \theta^{1/s} B + \theta^{2/s} B^2),$$

donde pasamos de un solo parámetro a *s*. En concreto, $\theta_0(B)$ sobre el que se aplica el contraste de raíces unitarias que se formula en esta sección se refiere al siguiente modelo MA(2)

$$\theta_0(B) = (1 - 2c_k \theta^{1/2} B + \theta B^2).$$

En síntesis, esta formulación permite derivar un contraste para determinar la existencia de raíces unitarias en cada una de las frecuencias estacionales, y que pueden ser la causa del rechazo de la hipótesis nula en el modelo agregado $MA(1)_s$.

Por esto, en esta sección se desarrollan los contrastes para un ciclo estacional en un modelo generalizado, y destacando al igual que sucede en los apartados anteriores las ventajas que el estadístico \mathcal{GM} .

4.4.1. la extensión del contraste de Tam y Reinsel

El modelo MA(2), en la línea de los trabajos de Tanaka (1990) y Tam y Reinsel (1997), se puede formular como

$$(1 - 2c_k B + B^2)y_t \equiv z_t = (1 - 2c_k \Theta^{1/2} B + \Theta B^2)a_t, \quad t = 3, \dots T,$$
(4.45)

donde el polinomio $(1 - 2c_k\Theta^{1/2}B + \Theta B^2)$ denota un modelo MA de orden 2 que proviene de la factorización de $(1 - \Theta B^s)$ para la frecuencia *k*-ésima y con periodo estacional s. El objeto de este contraste es determinar si H_0 : $\Theta = 1$, y supone que el modelo contiene dos raíces unitarias en la frecuencia *k*-ésima, frente a la alternativa, $H_1 : \Theta \neq 1$, en que es una media móvil invertible de orden 2.La notación empleada es la misma que en las secciones anteriores y se considera que z_t es la serie estacionaria y $c_k = \cos(2\pi k/s)$. La representación matricial para el modelo (4.45) es

$$\mathbf{z} = \boldsymbol{\nabla}_2(\Theta)$$

donde la serie estacionaria $\mathbf{z} = (z_3 \ z_4 \ \dots \ z_T)'$ es un vector columna de T - 2 observaciones estacionarias, $\mathbf{a} = (a_1 \ a_2 \ \dots \ a_T)'$ es un vector columna de T errores aleatorios y $\nabla_2(\Theta)$ es una submatriz que está compuesta por las últimas T - 2 filas de la matriz cuadrada de orden T y $\mathbf{D}_2(\Theta) = (\mathbf{I} - 2c_k\Theta^{1/2}\mathbf{L} + \Theta\mathbf{L}^2)$ que tiene unos en la diagonal principal, $-2c_k\Theta$ en la primera superdional, Θ en la segunda superdiagonal y ceros en el resto. Bajo el supuesto normalidad de los errores, $\mathbf{a} \sim N(\mathbf{0}, \sigma_a^2\mathbf{I}_T)$, se tiene que la serie diferenciada tiene una distribución normal de media cero y matriz de varianzas y covarianzas

$$\mathbf{\Omega}_2(\Theta) = \mathbf{\nabla}_2(\Theta) \mathbf{\nabla}_2(\Theta)',$$

donde $\mathbf{z} \sim N(\mathbf{0}, \sigma_a^2 \mathbf{\Omega}_2(\Theta)).$

El estadístico de contraste para H_0 : $\Theta = 1$ frente a H_0 : $\Theta \neq 1$ es LBUI y su expresión proviene del Teorema 3 establecido en el capítulo 2

$$\mathcal{T}_{2,n} \equiv \frac{\mathbf{z}'(\boldsymbol{\nabla}_2 \boldsymbol{\nabla}_2')^{-1} (2\mathbf{I} - c_k (\mathbf{L} + \mathbf{L}')) (\boldsymbol{\nabla}_2 \boldsymbol{\nabla}_2')^{-1} \mathbf{z}}{\mathbf{z}'(\boldsymbol{\nabla}_2 \boldsymbol{\nabla}_2')^{-1} \mathbf{z}}.$$
(4.46)

4.4.2. La aproximación de Saikkonen y Luukkonen

También se puede obtener el estadístico con una aproximación similar a la derivada en Saikkonen y Luukkonen (1993). El modelo MA(2) definido en

$$y_t = \mu_{0,t} + u_t, \qquad t = 1,2$$

$$(1 - 2c_k B + B^2)y_t = (1 - 2c_k \Theta^{1/2} B + \Theta B^2)u_t, \quad t = 3, \dots T,$$
(4.47)

se puede expresar como,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}_2^{-1}\mathbf{D}_2(\boldsymbol{\Theta})\mathbf{u},$$

donde $\mathbf{y} = (y_1, \dots, y_T)'$ es un vector de orden $T \times 1$ de observaciones, $\mathbf{u} \sim N(\mathbf{0}, \sigma_u^2 \mathbf{I}_T)$, $\mathbf{D}_2(\Theta) = \mathbf{I} - 2c_k \Theta^{1/2} \mathbf{L} + \Theta \mathbf{L}^2$, $\mathbf{X} = \mathbf{D}_2^{-1}[\mathbf{i}_k, \mathbf{i}_{k+1}] = [\mathbf{c}_k, \mathbf{s}_k]$ es una matriz de orden $T \times 2$ donde \mathbf{i}_j es un vector $T \times 1$ de ceros salvo un 1 en la posición *j*-ésima, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ es un vector 2×1 de parámetros, $\mathbf{u} = (u_1, \dots, u_T)'$ es un vector $T \times 1$ de perturbaciones aleatorias y la serie estacionaria es \mathbf{z} donde que $\mathbf{D}_2 \mathbf{y} = \mathbf{z}$, y $\mathbf{D}_2 = \mathbf{D}_2(\Theta)|_{\Theta=1}$.

La distribución de **z** es **z** ~ $N(0, \sigma_u^2 \Omega_2(\Theta))$. La matriz de varianzas y covarianzas del

proceso es $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma_u^2 \mathbf{D}_2^{-1} \mathbf{\Omega}_2(\Theta) \mathbf{D}_2^{-1'})$ donde $\mathbf{\Omega}_2(\Theta) = (\mathbf{I} - 2c_k \Theta^{1/2} \mathbf{L} + \Theta \mathbf{L}^2)(\mathbf{I} - 2c_k \Theta^{1/2} \mathbf{L} + \Theta \mathbf{L}^2)'$.

El estadístico de contraste es LBIU y se obtiene a partir de la segunda derivada de la función de log-verosimilitud $L(\Theta, \beta, \sigma_u^2)$ correspondiente al modelo (4.47) evaluada bajo la hipótesis nula $H_0 : \Theta = 1$. Parece claro que la primera derivada evaluada bajo la hipótesis nula es igual a cero. Utiliando el Teorema 3. La expresión del estadistico para el modelo MA(2) es la siguiente

$$\mathcal{SL}_2 = \frac{\hat{\mathbf{u}}' \mathbf{D}_2^{-1} (2\mathbf{I} - c_k (\mathbf{L} + \mathbf{L}')) \mathbf{D}_2'^{-1} \hat{\mathbf{u}}}{\hat{\mathbf{u}}' \hat{\mathbf{u}}}, \qquad (4.48)$$

donde $\hat{\mathbf{u}}$ son los residuos del modelo (4.47).

La distribución límite del estadístico SL_2 se puede expresar en términos de procesos brownianos tal que

$$\mathcal{SL}_2 \xrightarrow{d} \sum_{k=1}^2 \int_0^1 V_k^2(r) dr \equiv CvM(2),$$

donde $V_k(r) = W(r) - rW(1)$ es un puente Browniano estándar y CvM(2) denota una distribución Cramér-von Mises con 2 grados de libertad.

Para el desarrollo de la distribución límite es conveniente utilizar la equivalencia entre el estadístico (4.48) y el contraste obtenido en un modelo estructural para s = 2, dado que las expresiones matriciales de los numeradores de ambos estadísticos son iguales;

$$2\mathbf{I} - ck(\mathbf{L} + \mathbf{L}') = \boldsymbol{\nabla}_2(\mathbf{C}_k\mathbf{C}'_k + \mathbf{S}_k\mathbf{S}_k)\boldsymbol{\nabla}'_2,$$

cuya prueba se puede encontrar en los elementos de la matriz $2\mathbf{I} - ck(\mathbf{L} + \mathbf{L}') = [\delta_{ij}]$ son

$$\delta_{i,j} = \begin{cases} c_{ki}^2 + c_{k(i+2)}^2 + s_{ki}^2 + s_{k(i+2)}^2 = 2, \text{ para } i = j \\ -c_{ki}c_{k(i+1)} - s_{ki}s_{k(i+1)} = -c_{k1}, \text{ para } i = j+1 \\ -c_{kj}c_{k(j+1)} - s_{kj}s_{k(j+1)} = -c_{k1}, \text{ para } j = i+1. \end{cases}$$

Teniendo en cuenta la equivalencia entre los estadísticos de contraste en ambas repre-

sentaciones, el residuo para el modelo bajo H_0 se puede expresar como

$$\hat{u}_t = c_{kt}(\alpha_k - \hat{\alpha}_k) + s_{kt}(\beta - \hat{\beta}_k) + u_t$$

donde,

112

$$\hat{\alpha}_{k} = \frac{\sum_{t=1}^{T-2} s_{kt}^{2} \sum_{t=1}^{T} c_{kt} y_{t} - \sum_{t=1}^{T} c_{kt} s_{kt} \sum_{t=1}^{T} s_{kt} y_{t}}{\sum_{t=1}^{T} c_{kt}^{2} \sum_{t=1}^{T} s_{kt}^{2} - \left(\sum_{t=1}^{T} c_{kt} s_{kt}\right)} = \frac{\sum_{t=1}^{T} c_{kt} \alpha_{k} + \sum_{t=1}^{T} c_{kt} u_{t}}{T/2}$$

у

$$\hat{\beta}_{k} = \frac{\sum_{t=1}^{T-2} c_{kt}^{2} \sum_{t=1}^{T} s_{kt} y_{t} - \sum_{t=1}^{T} c_{kt} s_{kt} \sum_{t=1}^{T} c_{kt} y_{t}}{\sum_{t=1}^{T} c_{kt}^{2} \sum_{t=1}^{T} s_{kt}^{2} - \left(\sum_{t=1}^{T} c_{kt} s_{kt}\right)} = \frac{\sum_{t=1}^{T} s_{kt} \beta_{k} + \sum_{t=1}^{T} s_{kt} u_{t}}{T/2},$$

por lo que el residuo se puede reescribir como

$$\hat{u}_t = u_t + (T/2)^{-1/2} c_{kt} \sum_{i=1}^T c_{kt} u_i + (T/2)^{-1/2} s_{kt} \sum_{i=1}^T s_{kt} u_i.$$

La extensión del estadístico a un modelo general, en la línea de lo establecido por Saikkonen y Luukkonen (1993), requiere reescribir el modelo para el MA(2)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}_2^{-1}\mathbf{D}_2(\boldsymbol{\Theta})\mathbf{b}_T,$$

donde \mathbf{b}_T es un proceso ARMA(p,q) cuya definición es la misma que en las secciones previas en donde $\mathbf{b}_T = (b_1, b_2, \dots, b_T)' \sim N(\mathbf{0}, \mathbf{\Omega}_{\mathbf{b},T})$, con matriz de varianzas y covarianzas $\mathbf{\Omega}_{\mathbf{b},T} = [\mathbf{\Omega}_b(i-j)]$ y $\gamma_b(k) = E(b_t b_{t-k})$. En función del término deerror **a**, $\mathbf{\Omega}_{b,T} = \sigma_a^2 \mathbf{\Omega}_{b,T}$ y estableciendo $\mathbf{e}_T = \mathbf{D}_2^{-1} \mathbf{D}_2(\Theta) \mathbf{b}_T$, se tiene que $\mathbf{e}_T \sim$ $N(\mathbf{0}, \sigma_a^2 \mathbf{D}_2' \mathbf{\Omega}_T(\Theta) \mathbf{D}_2)$ con

$$\mathbf{\Omega}_T(\Theta) = (\mathbf{I} - 2c_k \Theta^{1/2} \mathbf{L} + \Theta \mathbf{L}^2) \mathbf{\Omega}_{b,T} (\mathbf{I} - 2c_k \Theta^{1/2} \mathbf{L} + \Theta \mathbf{L}^2)'$$

El estadístico de contraste corregido se obtiene a partir del Teorema 3 es LBIU expresar como

$$\mathcal{SL}_{2,T}^{*} = \frac{\tilde{\mathbf{e}}_{*}^{\prime} \mathbf{D}_{2*}^{-1} \left(\left(\frac{1}{2} c_{k} \mathbf{L} \right) \mathbf{\Omega}_{b} \mathbf{D}_{2}^{\prime} + \mathbf{D}_{2} \mathbf{\Omega}_{b} \left(\frac{1}{2} c_{k} \mathbf{L} \right)^{\prime} + 2 \left(-c_{k} \mathbf{L} + \mathbf{L}^{2} \right) \mathbf{\Omega}_{b} \left(-c_{k} \mathbf{L} + \mathbf{L}^{2} \right)^{\prime} \right) \mathbf{D}_{2*}^{-1} \tilde{\mathbf{e}}_{*}}{\tilde{\mathbf{e}}_{*}^{\prime} \tilde{\mathbf{e}}_{*}},$$

donde se emplean las distintas variables transformadas $\mathbf{D}_{2*}^{-1} = \mathbf{\Omega}_b^{-1/2} \mathbf{D}_2^{-1}$ y $\tilde{\mathbf{e}}_* = \mathbf{\Omega}_b^{-1/2} \tilde{\mathbf{e}}$, y se considera que $\tilde{\mathbf{e}}$ son los residuos de Mínimos Cuadrados Generalizados. Finalmente, apuntar que en la formulación del estadístico y su distribución se han empleado algunos resultados que se pueden encontrar en el apéndice.

4.4.3. Contraste para de estabilidad paramétrica para una frecuencia

La representación trigonométrica de la estacionalidad, equivalente a la forma reducida en (4.47) para la *k*-ésima frecuencia se define como

$$y_t = \alpha_{kt}c_{kt} + \beta_{kt}s_{kt} + u_t,$$

$$\alpha_{kt} = \alpha_{kt-1} + \omega_{kt}$$

$$\beta_{kt} = \beta_{kt-1} + \omega_{kt}^*, \quad t = 1, \dots, T,$$

$$(4.49)$$

donde el efecto estacional en el periodo t está representado por $\alpha_{kt}c_{kt} + \beta_{kt}s_{kt}$, los parámetros estacionales α_{kt} y β_{kt} evolucionan en el tiempo como paseos aleatorios, $c_{kt} = \cos(2\pi kt/s)$ y $s_{kt} = \sin(2\pi kt/s)$ y donde las perturbaciones $\omega_{kt} \sim N(0, \sigma_{\omega_k}^2)$ y $\omega_{kt}^* \sim N(0, \sigma_{\omega_k}^2)$ están incorreladas. Su representación matricial es

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{C}_k \boldsymbol{\omega}_k + \mathbf{S}_k \boldsymbol{\omega}_k^* + \mathbf{u}, \tag{4.50}$$

donde $\mathbf{y} = (y_1, \dots, y_T)'$ es un vector de orden $T \times 1$ de variables observadas, \mathbf{X} es una matriz de orden $T \times 2$ tal que $\mathbf{X} = [\mathbf{c}_k, \mathbf{s}_k]$ donde $\mathbf{c}_k = (c_{k1}, \dots, c_{kt})'$, $\mathbf{s}_k = (s_{k1}, \dots, s_{kt})'$ son vectores, de orden $T \times 1$, el vector de valores iniciales $\boldsymbol{\beta}_k = (\alpha_k, \beta_k)'$ es de orden 2×1 y \mathbf{C}_k y \mathbf{S}_k denotan matrices de orden $T \times T$, tal que $\mathbf{C}_k = \mathbf{c}_k \odot \mathbf{C}$, $\mathbf{S}_k = \mathbf{s}_k \odot \mathbf{C}$, donde $\mathbf{C} = [c_{ij}]$ es una matriz triangular inferior de orden $T \times T$ con $c_{i,j} = 1$ para $i \leq j$, siendo ceros el resto de elementos. Finalmente, los vectores de perturbaciones aleatorias $\mathbf{u} = (u_1, \dots, u_T)'$, $\boldsymbol{\omega}_k = (\omega_{k1}, \dots, \omega_{kT})'$ y $\boldsymbol{\omega}_k^* = (\omega_{k1}^*, \dots, \omega_{kT}^*)'$ son de dimensión $T \times 1$. Además, considerando valores iniciales deterministas, se pueden expresar los parámetros $\alpha_{k,t}$ y β_{kt} como tal que $\alpha_{kt} = \alpha_k + \sum_{i=1}^t \omega_{ki}$ y $\beta_{kt} = \beta_k + \sum_{i=1}^t \omega_{ki}^*$. El parámetro de interés ρ = es ahora $\rho = \sigma_{\omega_k}^2/\sigma_{\epsilon}^2 = \sigma_{\omega_k}^2/\sigma_{\epsilon}^2 = 0$,

El estadístico para contrastar dos raíces unitarias en la frecuencia *k*-ésima, para la hipótesis nula H_0 : $\rho = 0$ frente a H_1 : $\rho < 1$ se obtiene directamente de la aplica-

ción del Teorema 2 al modelo (4.49) y empleando el siguiente resultado sobre la matriz de varianzas y covarianzas

$$\left. rac{d \mathbf{\Omega}(
ho)}{d
ho}
ight|_{
ho=0} = \mathbf{C}_k \mathbf{C}_k' + \mathbf{S}_k \mathbf{S}_k',$$

nos lleva a

$$\mathcal{NM}_{2,n} = \frac{\hat{\mathbf{e}}' \left(\mathbf{C}_k \mathbf{C}'_k + \mathbf{S}_k \mathbf{S}'_k \right) \hat{\mathbf{e}}}{\hat{\mathbf{e}}' \hat{\mathbf{e}}},\tag{4.51}$$

que se puede expresar en términos de los residuos de mínimos cuadrados ordinarios, $\hat{\mathbf{e}}$ de la regresión de \mathbf{y} sobre la matriz $\mathbf{X} = [\mathbf{c}_k, \mathbf{s}_k]$.

4.4.4. El contraste para las frecuencias estacionales generalizado

Se puede generalizar el modelo (4.49) como

$$y_{t} = \alpha_{kt}c_{kt} + \beta_{kt}s_{kt} + u_{t}, \quad \phi(B)u_{t} = \theta(B)w_{k1,t}$$

$$\alpha_{kt} = \alpha_{kt-1} + \omega_{k,t}, \qquad \phi(B)\omega_{kt} = \theta(B)w_{k2,t} \qquad (4.52)$$

$$\beta_{kt} = \beta_{kt-1} + \omega_{kt}^{*}, \qquad \phi(B)\omega_{kt}^{*} = \theta(B)w_{k2,t}^{*}, \quad t = 1, ..., T,$$

donde los errores $w_{k1,t}$, $w_{k2,t}$ y $w_{k2,t}^*$ son procesos ARMA(p,q) independientes, en los que se supone que el polinomio $\phi(B)$ no contiene raíces unitarias, que en cambio sí se permiten en el polinomio $\theta(B)$ que puede contener raíces unitarias extra, y además no se produce cancelación de operadores entre ambos polinomios. Al igual que en las secciones precedentes, los dos valores premuestrales α_{k0} y β_{k0} son deterministas. Los términos de error se pueden definir como $w_{1k,t} \sim iidN(0, \sigma_{k1}^2)$, $w_{2k,t} \sim iidN(0, \sigma_{k2}^2)$, no existiendo correlación entre ellos.

El modelo equivalente a (4.52) es un modelo ARIMA(p, 2, q + 2)

$$\phi(B)(1 - 2c_k B + B^2)y_t = \theta(B)(1 - 2c_k \Theta^{1/2} B + \Theta B^2)a_t, \quad t = 3, \dots, T,$$
(4.53)

y análogamente, la representación estacionaria de (4.52) se puede formular de la si-

guiente forma

$$(1 - 2c_k B + B^2)y_t = (1 - c_k B)\omega_{kt} + s_k B\omega_{kt}^* + (1 - 2c_k B + B^2)\epsilon_t, \quad t = 3, \dots, T.$$
(4.54)

En este contexto, se pueden plantear dos contrastes equivalentes empleando dos representaciones distintas, por una parte el contraste de no invertibilidad en la representación ARIMA para las hipotésis H_0 : $\Theta = 1$ frente H_1 : $\theta \neq 1$ en (4.53) y el contraste de estacionalidad determinista en la frecuencia k para H_0 : $\rho = 0$ frente H_1 : $\rho > 0$ en (4.52); donde ahora, el parámetro ρ tiene la siguiente definición $\rho = \sigma_{k1}^2 / \sigma_{k2}^2 = \sigma_{k1}^2 / \sigma_{k2}^{*2}$. La representación matricial del modelo de regresión generalizado para (4.52) es

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} = \mathbf{C}_k \boldsymbol{\omega}_k + \mathbf{S}_k \boldsymbol{\omega}_k^* + \mathbf{u}, \tag{4.55}$$

donde $\mathbf{u} \sim N[\mathbf{0}, \sigma_{k1}^2 \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})], \, \boldsymbol{\omega}_k \sim N[\mathbf{0}, \sigma_{k2}^2 \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})] \, \mathbf{y} \, \boldsymbol{\omega}_k^* \sim N[\mathbf{0}, \sigma_{k2}^{*2} \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})], \, \mathbf{y} \, \mathbf{la}$ matriz $\mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})$ depende del vector de parámetros $\boldsymbol{\varphi} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$. La distribución muestral de \mathbf{y} viene dada por

$$\mathbf{y} \sim N\left(\mathbf{X}\boldsymbol{\beta}, \ \sigma_{k1}^{2}\boldsymbol{\Omega}_{y,T}(\boldsymbol{\rho},\boldsymbol{\varphi})\right), \quad \boldsymbol{\Omega}_{\mathbf{y},T}(\boldsymbol{\rho},\boldsymbol{\varphi}) = \rho\left(\mathbf{C}_{k}\boldsymbol{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})\mathbf{C}_{k}' + \mathbf{S}_{k}\boldsymbol{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})\mathbf{S}_{k}'\right) + \boldsymbol{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}),$$
(4.56)

que se puede reescribir dado que las matrices las matrices \mathbf{C}_s y \mathbf{S}_s son intercambiables por sus traspuestas, por lo que $\mathbf{\Omega}_{y,T}(\rho, \boldsymbol{\varphi}) = \rho \left(\mathbf{C}'_k \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \mathbf{C}_k + \mathbf{S}'_k \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \mathbf{S}_k \right) +$ $\mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi})$. Al igual que en las secciones anteriores, y con objeto de simplicar la notación de $\mathbf{\Omega}_{\mathbf{y},T}(\rho, \boldsymbol{\varphi})$, se omite la referencia a $\boldsymbol{\varphi}$, tal que $\mathbf{\Omega}_{y,T}(\rho)$. También se establece una nueva definición para el número de observaciones de la serie diferenciada que ahora se denota m = T - 2 para ajustarse al modelo MA(2).

De nuevo, y para poder obtener el estadístico siguiendo la aproximación de Tanaka (1990) es preciso diferenciar (4.55) utilizando la matriz de diferencias rectangular ∇_2 , de orden $(T - 2) \times T$ previamente definida

$$\mathbf{z} = \nabla_2 \mathbf{e}, \quad \nabla_2 \mathbf{e} = \nabla_2 \mathbf{C}_k \boldsymbol{\omega}_k + \nabla_2 \mathbf{S}_k \boldsymbol{\omega}_k^* + \nabla_2 \mathbf{u},$$

donde z es la serie diferenciada. Para calcular la distribución muestral de z es con-

veniente tener en cuenta los siguientes resultados matriciales, $\nabla_2 \mathbf{C}_k = [\mathbf{0}|\mathbf{I}_{T-2}] - c_k[\mathbf{0}|\mathbf{B}_{T-2}] = \nabla_2^c \text{ y } \nabla_2 \mathbf{S}_k = s_k[\mathbf{0}|\mathbf{I}_{T-2}] = \nabla_2^s$, entonces $\mathbf{z} \sim N \left[\mathbf{0}, \sigma_{k1}^2 \mathbf{\Omega}_{z,m}(\rho, \boldsymbol{\varphi})\right]$ donde

$$\boldsymbol{\Omega}_{\mathbf{z},m}(\rho,\boldsymbol{\varphi}) = \rho \left(\boldsymbol{\nabla}_{2}^{c} \boldsymbol{\Omega}_{\mathbf{u},m}(\boldsymbol{\varphi}) \boldsymbol{\nabla}_{2}^{c'} + \boldsymbol{\nabla}_{2}^{s} \boldsymbol{\Omega}_{\mathbf{u},m}(\boldsymbol{\varphi}) \boldsymbol{\nabla}_{2}^{s'} \right) + \boldsymbol{\nabla}_{2} \boldsymbol{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \boldsymbol{\nabla}_{2}^{c'}, \qquad (4.57)$$

que se puede simplicar más teniendo en cuenta que $\nabla_2^c \Omega_{\mathbf{u},m}(\boldsymbol{\varphi}) \nabla_2^{c'} = (1+c_k)^2 \Omega_{\mathbf{u},m}(\boldsymbol{\varphi}) - c_k \mathbf{L}_{T-2} \Omega_{\mathbf{u},m}(\boldsymbol{\varphi}) - c_k \Omega_{\mathbf{u},m}(\boldsymbol{\varphi}) \mathbf{L}'_{T-2}$ y que $\nabla_2^s \Omega_{\mathbf{u},m}(\boldsymbol{\varphi}) \nabla_2^{s'} = s_k^2 \Omega_{\mathbf{u},m}(\boldsymbol{\varphi}).$

Finalmente el modelo en términos de Θ y σ_a^2 es

$$\mathbf{z} \sim N\left[\mathbf{0}, \sigma_a^2 \mathbf{\Omega}_{\mathbf{z},m}(\theta, \boldsymbol{\varphi})\right] \quad \text{con} \quad \mathbf{\Omega}_{\mathbf{z},m}(\Theta, \boldsymbol{\varphi}) = \mathbf{\nabla}_2(\Theta) \mathbf{\Omega}_{\mathbf{u},T}(\boldsymbol{\varphi}) \mathbf{\nabla}_2(\theta)'.$$

4.4.5. Estadísticos de contraste

En este apartado, y en coherencia con las secciones anteriores, se formulan los estadísticos de contraste en términos de tres tipos de residuos. Por una parte tenemos la extensión de Leybourne y McCabe (1994), cuyos residuos condicionados provienen del filtrado de la serie temporal empleando un polinomio AR y que considera unos valores iniciales nulos. La de Tam y Reinsel (1997), que como ya hemos visto en las secciones previas, utiliza los residuos de un modelo ARMA calculados utilizando el método de retrovisión. Y el estadístico que utiliza los residuos exactos, que como ya se ha apuntado, tiene una expresión más sencilla que el estadístico exacto de Saikkonen y Luukkonen (1993), y permite la existencia de otras raíces unitarias MA adicionales a la que se considera bajo la hipótesis de contraste.

Teniendo en cuenta que los fundamentos para la obtención de estos contraste se detallan ampliamente en la sección 3, la formulación de los mismos únicamente requiere de su aplicación en el marco, ahora, de un MA(2). En cualquier caso, se sugiere de nuevo la aplicación del contraste LBI, por la simplicidad de su aplicación.

En concreto, los estadísticos $\mathcal{GM}_{2,n}^*$ y $\mathcal{GM}_{2,n-1}^*$ se obtiene aplicando de forma directa el Teorema 2 al modelo MA(2) generalizado, tanto en niveles como diferenciado, y que supone únicamente sustituir las matrices de varianzas y covarianzas en los estadísticos (3.71) y (3.72).

A continuación, el estadístico basado en los residuos exactos, en la que se consideran

todos los supuestos de la sección 3.5. para el modelo generalizado, tiene la siguientes expresión

$$\mathcal{GM}_2 = rac{1}{T-2} rac{ ilde{\mathbf{a}}' \left(\mathbf{C}_k \mathbf{C}'_k + \mathbf{S}_k \mathbf{S}'_k
ight) ilde{\mathbf{a}}}{ ilde{\mathbf{a}}' ilde{\mathbf{a}}},$$

donde **ã** denota el vector de los residuos exactos. Como puede observarse, su expresión es similar a la de los estadísticos para una raíz unitaria regular y estacional, la principal diferencia entre los tres estadísticos obtenidos está en las matrices centrales del numerador, que actúan como matrices de acumulación de los residuos.

De forma equivalente, se puede obtenemos el estadístico de contraste generalizado

$$\mathcal{NM}_{2,n}^{*} = \frac{\mathbf{\hat{e}}_{*,T}^{\prime} \left(\mathbf{C}_{*,k,T} \mathbf{\Omega}_{\mathbf{u},T} \mathbf{C}_{*,k,T}^{\prime} + \mathbf{S}_{*,k,T} \mathbf{\Omega}_{\mathbf{u},T} \mathbf{S}_{*,k,T}^{\prime} \right) \mathbf{\hat{e}}_{*,T}}{\mathbf{\hat{e}}_{*,T}^{\prime} \mathbf{\hat{e}}_{*,T}}, \qquad (4.58)$$

en donde se considera que $\mathbf{\hat{e}}_{*,T} = \mathbf{M}_{*,T}\mathbf{y}_{*,T}, \mathbf{y}_{*,T} = \mathbf{\Omega}_{b,T}^{-1/2}\mathbf{y}_{T}, \mathbf{X}_{*,T,s} = \mathbf{\Omega}_{b,T}^{-1/2}[\mathbf{c}_{k}, \mathbf{s}_{k}],$ $\mathbf{M}_{*,T} = \mathbf{I}_{T} - \mathbf{X}_{*,T,s}(\mathbf{X}'_{*,T,s}\mathbf{X}_{*,T,s})^{-1}\mathbf{X}_{*,T,s}', \mathbf{C}_{*,k,T} = \mathbf{\Omega}_{b,T}^{-1/2}\mathbf{C}_{k} \mathbf{y} \mathbf{S}_{*,k,T} = \mathbf{\Omega}_{b,T}^{-1/2}\mathbf{S}_{k},$

y que es equivalente al estadístico

$$\mathcal{GM}_{s,n}^{*} = \frac{\mathbf{y}_{*,T}^{\prime} \mathbf{\nabla}_{*,2}^{\prime} (\mathbf{\nabla}_{*,2} \mathbf{\nabla}_{*,2}^{\prime})^{-1} \mathbf{\nabla}_{*,2} \left(\mathbf{C}_{k} \mathbf{\Omega}_{\mathbf{u},T} \mathbf{C}_{k}^{\prime} + \mathbf{S}_{k} \mathbf{\Omega}_{\mathbf{u},T} \mathbf{S}_{k}^{\prime} \right) \mathbf{\nabla}_{*,2}^{\prime} (\mathbf{\nabla}_{*,2} \mathbf{\nabla}_{*,2}^{\prime})^{-1} \mathbf{y}_{*,T}}{\mathbf{y}_{*,T}^{\prime} \mathbf{\nabla}_{*,2}^{\prime} (\mathbf{\nabla}_{*,2} \mathbf{\nabla}_{*,2}^{\prime})^{-1} \mathbf{\nabla}_{*,2}^{\prime} \mathbf{y}_{*,T}},$$

en las que se emplea la transformación de la matriz de diferencias rectangular $\nabla_{*,2} = \nabla_2 \Omega_{y,T}^{1/2}$ y la utilísima relación matricial $\mathbf{M}_{*,T} = \nabla'_{*,2} (\nabla_{*,2} \nabla'_{*,2})^{-1} \nabla_{*,2}$.

Otra extensión paramétrica para recoger autocorrelación ya desarrallada a lo largo de esta tesis es la de Leybourne y McCabe (1994), que sugiere filtrar la serie considerando que es generada por un proceso AR(p), donde ahora $y_{*,t} = (\phi(B)/\theta(B))y_t$,

$$\mathcal{LM}_{2,n} = \frac{\mathbf{y}_{*}' \mathbf{M} \left(\mathbf{C}_{k,T} \mathbf{C}_{k,T}' + \mathbf{S}_{k,T} \mathbf{S}_{k,T}' \right) \mathbf{M} \mathbf{y}_{*}}{\mathbf{y}_{*}' \mathbf{M} \mathbf{y}_{*}}, \qquad (4.59)$$

que también se puede formular en téminos del modelo ARIMA, tal que

$$\mathcal{LM}_{2,n} = \frac{\mathbf{z}'_{*}(\boldsymbol{\nabla}_{2}\boldsymbol{\nabla}'_{2})^{-1}(2\mathbf{I} - c_{k}(\mathbf{B} + \mathbf{B}'))(\boldsymbol{\nabla}_{2}\boldsymbol{\nabla}'_{2})^{-1}\mathbf{z}_{*}}{\mathbf{z}'_{*}(\boldsymbol{\nabla}_{2}\boldsymbol{\nabla}'_{2})^{-1}\mathbf{z}_{*}},$$

donde $\mathbf{z}_* = \boldsymbol{\nabla}_2 \mathbf{y}_*$.

Finalmente, la última corrección propuesta es la de Tam y Reinsel (1997), aplicada al modelo MA(2). El modelo aplicado es

$$\begin{split} \phi(B) \nabla_2 y_t = \theta(B) z_{*,t}, \\ z_{*,t} = (1 - 2c_k \Theta^{1/2} B + \Theta B), \end{split}$$

donde $z_{*,t}$ sigue siendo el proceso MA(2) de referencia. De forma equivalente al estadístico se puede aplicar la corrección de Tam y Reinsel (1997), que al igual que la corrección paramétrica de Leybourne y McCabe (1994) filtra la serie utilizando los residuos obtenidos por el método de retrovisión $\tilde{\mathbf{e}}_{*}$

$$\mathcal{T}_{2,n} = rac{ ilde{\mathbf{e}}'_* \left(\mathbf{C}_k \mathbf{C}'_k + \mathbf{S}_k \mathbf{S}'_k
ight) ilde{\mathbf{e}}_*}{ ilde{\mathbf{e}}'_* \mathbf{M} ilde{\mathbf{e}}_*}.$$

4.4.6. Distribución exacta

Para obtener la distribución exacta es preciso realizar algunas transformaciones. Teniendo en cuenta que $\mathbf{z}_* = \mathbf{\Omega}_{u,n-1}^{-1/2}(\boldsymbol{\varphi})\mathbf{z}$, se tiene que $\mathbf{z}_* \sim N\left(\mathbf{0}, \sigma_1^2(\boldsymbol{\rho}\mathbf{A} + \nabla_{\#}\nabla'_{\#})\right)$, $\mathbf{A} = 2\mathbf{I}_{T-2} - c_k(\mathbf{L}_{T-2} + \mathbf{L}'_{T-2})$ en donde $\nabla_{\#} = \mathbf{\Omega}_{u,n-1}^{-1/2}(\boldsymbol{\varphi})\nabla\mathbf{\Omega}_{u,n}^{1/2}(\boldsymbol{\varphi})$.

Además, notando que $\Omega_{u,n-1}^{1/2}(\varphi)(\nabla_2\Omega_{u,n}(\varphi)\nabla'_2)^{-1}\Omega_{u,n-1}^{1/2}(\varphi) = (\nabla_{\#}\nabla'_{\#})^{-1}$, puede escribirse como

$$\mathcal{GM}_{2,n}^* = \frac{\epsilon' (\nabla_{2,\#} \nabla'_{2,\#})^{-1} \epsilon}{\epsilon' \epsilon}, \qquad (4.60)$$

dado que

$$\boldsymbol{\epsilon} = (\boldsymbol{\nabla}_{\#} \boldsymbol{\nabla}'_{\#})^{-1/2} \mathbf{z}'_{*} \sim N \left(\mathbf{0}, \ \sigma_{1}^{2} (\mathbf{I}_{n-1} + \rho \boldsymbol{\nabla}_{2,\#} \boldsymbol{\nabla}'_{2,\#}) \right),$$

en donde $\nabla_{\#} \mathbf{A} \nabla'_{\#} = \nabla_{2,\#} \nabla'_{2,\#}$. Si se consideran λ_k los autovalores de $\nabla_{2,\#} \nabla'_{2,\#}$

$$\mathcal{GM}_{2,n}^{*} = \frac{\sum_{k=1}^{n-1} \lambda_{k} (1+\rho\lambda_{k}) \xi_{k}}{\sum_{k=1}^{n-1} (1+\rho\lambda_{k}) \xi_{k}},$$
(4.61)

en donde λ_k (k = 1, ..., n - 2) son ahora los recíprocos de los autovalores de $\nabla_{2,\#} \nabla'_{2,\#}$. De forma equivalente, se puec obtener la distribución límite bajo la secuencia de alternativas locales $H_1: \rho/n$ de $\mathcal{GM}_{2,n}^*$

$$\mathcal{GM}_{s,n}^*/n \stackrel{a}{\sim} \sum_{k=1}^{\infty} \left(\frac{1}{k^2 \pi^2} + \frac{c^2}{k^4 \pi^4} \right) \xi_k.$$
 (4.62)

donde $\xi_k \sim iid\chi_2$.

4.4.7. Evaluación de los contrastes

Finalmente se considera oportuno hacer una evaluación de los contrastes para comprobar si la presencia de autocorrelación altera su distribución muestral bajo la hipótesis nula en muestras finitas. Se considera que el valor del parámetro bajo la hipótesis nula es conocido.

En primer lugar se tabulan mediante procedimientos de integración numérica las distribuciones muestrales de los estadísticos $SL_{s,n}$ y $GM_{s,n}$ suponiendo que el modelo teórico es un MA(s+1) no invertible

$$z_t = (1 - B^s)(1 - \theta B)a_t,$$

donde el primer factor MA, $1 - B^s$, recoge la hipótesis nula de no invertibilidad y se considera que s = 4, 8. El segundo factor MA se corresponde con el parámetro de control del experimento y recoge la correlación. Cuando $\theta = 0$, las distribuciones muestrales de los estadísticos $S\mathcal{L}_{s,n}$ y $\mathcal{GM}_{s,n}$ deben coincidir con la tabulada en Tam y Reinsel (1997) replicada en el Cuadro 4.1. Además, el modelo MA(s+1) también permite evaluar la correlación serial AR si tenemos en cuenta que puede aproximarse por un AR de orden finito

$$(1+\theta B+\theta^2 B^2+\cdots+\theta^p B^p)z_t=(1-B^s)a_t.$$

Antes de formular los estadísticos se considera necesario introducir la siguiente notación para el proceso estacionario z_t del modelo MA(s+1), donde Ω_z es la matriz de covarianzas, \mathbf{L}_z es el factor de Cholesky de Ω_z , y Ψ_z y Π_z son matrices triangulares con los pesos ψ_j y π_j del MA(s+1). De forma análoga se pueden definir las correspondientes matrices para el proceso auxiliar $u_t = (1 - \theta B^s) w_t$. De esta forma los estadísticos de no invertibilidad pueden expresarse en términos de un vector $\boldsymbol{\epsilon} \sim N(0, \sigma^2 \mathbf{I})$:

1. Saikkonen-Luukkonen

$$\mathcal{SL}_{s}^{*} = \frac{1}{T-s} \frac{\tilde{\mathbf{e}}_{s}^{\prime} \mathbf{L}_{\mathbf{u},T}^{-1} \mathbf{C}_{s}^{\prime} \mathbf{\Omega}_{\mathbf{u},T} \mathbf{C}_{s} \mathbf{L}_{\mathbf{u},T}^{\prime - 1} \tilde{\mathbf{e}}_{*}}{\tilde{\mathbf{e}}_{*}^{\prime} \tilde{\mathbf{e}}_{*}} = \frac{1}{T-s} \frac{\boldsymbol{\epsilon}^{\prime} \mathbf{L}_{\mathbf{u},T-s}^{\prime} \mathbf{\Omega}_{\mathbf{z},T-s}^{-1} \mathbf{L}_{\mathbf{u},T-s} \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}^{\prime} \boldsymbol{\epsilon}}.$$

2. Gallego-Mazas

$$\mathcal{GM}_{s}^{*} = \frac{1}{T-s} \frac{\tilde{\mathbf{a}}' \mathbf{M}_{s} \mathbf{C}_{s}' \mathbf{C}_{s} \mathbf{M}_{s} \tilde{\mathbf{a}}}{\tilde{\mathbf{a}}' \mathbf{M}_{s} \tilde{\mathbf{a}}} = \frac{1}{T-s} \frac{\boldsymbol{\epsilon}' \mathbf{\Psi}_{z} \mathbf{L}_{u}^{-1} \mathbf{M}_{s} \mathbf{C}_{s}' \mathbf{C}_{s} \mathbf{M}_{s} \mathbf{\Psi}_{z} \mathbf{L}_{u}' \boldsymbol{\epsilon}}{\boldsymbol{\epsilon}' \mathbf{\Psi}_{z} \mathbf{L}_{u}^{-1} \mathbf{M}_{s} \mathbf{\Psi}_{z} \mathbf{L}_{u}' \boldsymbol{\epsilon}}.$$

El Cuadro 4.5 muestra los valores críticos de los estadísticos SL_s^* y GM_s^* al nivel de significación del 5% para s = 4 y 12 y n = 10, 20, 30, tal que T = ns. Se consideran también distintos valores para el parámetro de control $\theta = -1, -0.75, ..., 1$.

Cuando $\theta = 0$ los valores criticos de SL_s^* y GM_s^* estadísticos para s = 4 y s = 12 coinciden puesto que no se aplica ninguna corrección. El valor crítico para ambos estdísticos, $\theta = 0$ y s = 4 es 0.313, y para s = 12 es 0.249. De forma general, los valores críticos de ambos estadíticos coinciden, o son muy próximos, aunque hay se pueden hacer algunas consideraciones. Cuando s = 4 los valores críticos de los estadísticos coinciden cuando $\theta < 0.5$, para valores mayores del parámetro de control θ se observan pequeñas diferencias que se acentúan para los tamaños de muestra más pequeños. Esta posible diferencia entre los valores críticos para s = 4 y s = 12 se debe a que el número de observaciones de la muestra empleada en la serie mensual es mayor. Para s = 12 coinciden todos los valores críticos de SL_s^* y GM_s^* para todos los tamaños muestrales, a excepción del valor crítico para n = 10 y $\theta = 1$ en que difieren en el tercer decimal. A medida que θ se aproxima uno, el comportamiento de los estadísticos SL_s^* y GM_s^* es el mismo, observándose un ligero desplazamiento hacia la izquierda de la distribución muestral. Este desplazamiento de la distribución recuerda los resultados de valores críticos para una doble raíz unitaria del Capítulo 3, aunque en ese caso el desplazamiento de la distribución es mucho más evidente. Esto quiere decir que la corrección paramétrica en los estadísticos SL_s^* y GM_s^* tiene un comportamiento razonable, incluso en presencia de una raíz unitaria extra.

De esta forma, los principales resultados del Cuadro 4.5 son los siguientes:

1. Los valores críticos obtenidos en la tabulación de la distribución muestral de los

estadísticos de no invertibilidad SL_s^* y GM_s^* proporcionan resultados similares.

- 2. Se observa una mayor coincidencia entre los resultados de ambos estadísticos para s = 12, que se debe a un mayor tamaño muestral.
- 3. Para valores del parámetro de control θ cercanos a uno se observa un ligerísimo desplazamiento de la distribución muestral hacia la izquierda.

		θ									
	п	-1	-0.75	-0.5	-0.25	0	0.25	0.5	0.75	1	
s = 4	10	0.316	0.339	0.346	0.348	0.348	0.348	0.347	0.339	0.316	
	20	0.292	0.323	0.328	0.329	0.329	0.329	0.328	0.323	0.292	
	30	0.284	0.318	0.322	0.323	0.323	0.323	0.322	0.318	0.284	
	100	0.273	0.312	0.313	0.313	0.313	0.313	0.313	0.312	0.273	
s = 12	10	0.274	0.282	0.282	0.282	0.282	0.282	0.282	0.282	0.274	
	20	0.254	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.254	
	30	0.247	0.258	0.258	0.258	0.258	0.258	0.258	0.258	0.247	
	100	0.238	0.249	0.249	0.249	0.249	0.249	0.249	0.249	0.238	
(b) GM_{π}											
		0									
		1	0.75	0.5	0.25	0	0.25	0.5	0.75	1	
- 1	$\frac{n}{10}$	-1	-0.75	-0.5	-0.25	0 2 4 9	0.25	0.5	0.75	<u>I</u>	
s = 4	10	0.318	0.347	0.348	0.348	0.348	0.348	0.348	0.347	0.318	
	20	0.293	0.328	0.329	0.329	0.329	0.329	0.329	0.328	0.293	
	30	0.285	0.322	0.323	0.323	0.323	0.323	0.323	0.322	0.285	
	100	0.273	0.313	0.313	0.313	0.313	0.313	0.313	0.313	0.273	
. 10	10	0 272	0 202	0 202	0 202	0 202	0 202	0 202	0 202	0 272	
s = 12	10	0.272	0.282	0.282	0.282	0.282	0.282	0.282	0.282	0.272	
	20	0.253	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.253	
	3U 100	0.24/	0.258	0.238	0.238	0.238	0.238	0.238	0.238	0.24/	
	100	0.238	0.249	0.249	0.249	0.249	0.249	0.249	0.249	0.238	

Cuadro 4.5: Valores críticos al 5 % bajo $z_t = (1 - B^s)(1 - \theta B)a_t$ (a) SL_n

4.5. Contraste para múltiples raíces unitarias en un modelo general

El siguiente paso una vez han sido derivados distintos contraste de raíces unitarias en el contexto de un modelo generalizado es plantear un procedimiento unificado con el que se puedan realizar contrastes que satisfagan diversas hipótesis.

El elemento clave para establecer este procedimiento unificado es el estadístico \mathcal{GM} basado en los residuos exactos, y que tiene varias ventajas con respecto a otros contraste examinados. Por una parte, como ya se ha demostrado en la comparación con otros estadísticos, tiene una expresión de cómputo sencillo; por otra a parte, permite que en el modelo sobre el que se aplica el contraste exista una raíz unitaria extra, y su funcionamiento, comparado con el estadístico exacto es considerablemente mejor que otros estadísicos examinados. Por esto, es de interés formular un procedimiento que sirva para contrastar varias hipótesis de contraste en un modelo general.

La idea general fundamental es que en un modelo ARIMA

$$\phi(B)w_t = \theta(B)a_t, \tag{4.63}$$

donde el polinomio media móvil $\theta(B) = \theta_0(B)\theta_*(B)$, que se puede dividir entre los factores que son objeto de contraste $\theta_0(B)$ y los restantes $\theta_*(B)$ que forman parte de la estructura del modelo generalizado, se pueda realizar un contraste para varias hipótesis nulas teniendo en cuenta que $\theta_*(B)$ puede tener raíces adicionales. En particular los polinomios $\theta_0(B)$ empleados en este tesis son

$$\theta_0(B) = \begin{cases} (1 - \theta B) \\ (1 - \Theta B^s) \\ (1 - 2c_k \Theta^{1/2} B + \Theta B^2). \end{cases}$$
(4.64)

Por esto, el contraste \mathcal{GM} es una opción conveniente, puesto que se puede aplicar en

cualquier contexto siendo las expresiones de estos estadísticos la siguientes:

$$\mathcal{GM}_{1} = \frac{\tilde{\mathbf{a}}'\left(\mathbf{C}\mathbf{C}'\right)\tilde{\mathbf{a}}}{\tilde{\mathbf{a}}'\tilde{\mathbf{a}}}, \quad \mathcal{GM}_{2} = \frac{\tilde{\mathbf{a}}'\left(\mathbf{C}_{k}\mathbf{C}'_{k} + \mathbf{S}_{k}\mathbf{S}'_{k}\right)\tilde{\mathbf{a}}}{\tilde{\mathbf{a}}'\tilde{\mathbf{a}}}, \quad y \quad \mathcal{GM}_{s} = \frac{\tilde{\mathbf{a}}'\left(\mathbf{C}_{s}\mathbf{C}'_{s}\right)\tilde{\mathbf{a}}}{\tilde{\mathbf{a}}'\tilde{\mathbf{a}}}$$

Observando estos estadísticos basados en los residuos exactos se aprecia que su forma es muy similar, la única diferencia radica en las matrices centrales del numerador que actúan como matrices de acumulación de los residuos. Dependiendo de la hipótesis nula que se considere, el único cambio que se aprecia en la fórmula de este estadístico está en la forma en que éstos se acumulan en el numerador. Por tanto, podríamos formular el contraste \mathcal{GM}_i , para que, con carácter general se pueda aplicar para diversas hipótesis de contraste, y tiene la siguiente forma

$$\mathcal{GM}_i = rac{1}{m} rac{ ilde{\mathbf{a}}' \mathbf{M}_i \mathbf{A}_i \mathbf{M}_i ilde{\mathbf{a}}}{ ilde{\mathbf{a}}' ilde{\mathbf{a}}},$$

donde el contador *i* denota el número de raíces unitarias bajo la hipótesis nula, \mathbf{A}_i es la matriz de acumulación que se correponde son esa hipótesis nula y $\mathbf{\tilde{a}}$ son los residuos exactos. En los casos evaluados en este trabajo, \mathbf{M}_i es la matriz \mathbf{M} correspondiente a cada caso y la matriz \mathbf{A} tiene las siguientes definiciones

$$\mathbf{A}_{i} = \begin{cases} \mathbf{C}\mathbf{C}', & i = 1, \\ \mathbf{C}_{k}\mathbf{C}_{k} + \mathbf{S}_{k}\mathbf{S}'_{k} & i = 2, \\ \sum_{k=1}^{[s/2]-1} \left(\mathbf{C}_{k}\mathbf{C}_{k} + \mathbf{S}_{k}\mathbf{S}'_{k}\right) & i = s - 1, \\ \mathbf{C}_{s}\mathbf{C}'_{s} & i = s, \end{cases}$$
(4.65)

donde $\sum_{k=1}^{[s/2]-1} (\mathbf{C}_k \mathbf{C}_k + \mathbf{S}_k \mathbf{S}'_k)$ denota la expresión de la matriz \mathbf{A}_i cuando se desea contrastar s - 1 frecuencias.

En cualquier caso, y teniendo en cuenta la siguiente descomposición del polinomio MA

$$(1 - \theta_0 B^s) = (1 - \Theta^{1/s} B)(1 + \Theta^{1/s} B) \prod_{k=1}^{[s/2]-1} (1 - 2c_k \Theta^{1/s} B + \Theta^{2/s} B^2),$$
(4.66)

el estadístico \mathcal{GM}_i propuesto se puede aplicar a cualquier hipótesis de contraste de raíces unitarias en (4.66).

Para concluir este capítulo es oportuno resaltar algunos de los corolarios que se pueden extraer. Por una parte se ha derivado un estadístico de contraste de raíces unitarias que puede ser aplicado para cualquier hipótesis de contraste únicamente cambiando la forma de acumulación de los residuos y su cálculo computacional es relativamente sencillo. Por otra parte, el estadístico se puede aplicar en un modelo general en el que pueden existir otras raíces unitarias, en último lugar su funcionamiento en razonablamente bueno.

Capítulo 5

Conclusiones

En esta tesis doctoral se ha propuesto un procedimiento general para contrastar la presencia de múltiples raíces MA unitarias en cualquier modelo ARIMA, que puede también incluir otras raíces MA unitarias además de las que son de interés. La consideración de estas raíces adicionales resulta conveniente como alternativa a la inclusión de términos deterministas (tendencia lineal), y es también útil en el marco de modelos estacionales múltiplicativos. De este modo, la representación usada permite formular de manera común una gran diversidad de contrastes de no invertibilidad o estabilidad paramétrica.

El procedimiento de contraste es similar al propuesto por Saikkonen y Luukkonen (1993) para el caso de una raíz MA regular en el sentido de que se tiene en cuenta la estructura ARMA adicional, pero presenta dos principales diferencias: la estrategia seguida para obtener el estadístico de contraste y las expresiones proporcionadas para el mismo. En primer lugar, el estadístico de contraste se obtiene en el marco de un modelo estructural equivalente, que simplifica su obtención porque solo requiere calcular la primera derivada de la función de verosimilitud. En segundo lugar, las expresiones propuestas son convenientes para calcular el estadístico en términos de los residuos, así como para evaluar su distribución muestral tanto bajo H_0 como H_1 .

Las propuestas más relevantes, de forma breve, se pueden agrupar en tres bloques.

En primer lugar se derivan de forma homogénea los contrastes de raíces unitarias más relevantes como son los propuestos por Nyblom y Mäkeläinen (1983), Tanaka (1990) y Saikkonen y Luukkonen (1993) en el caso más sencillo, esto es, para la hipótesis nula de una raíz unitaria regular en un modelo MA(1). Constituye una aportación interesante el marco metodológico común desarrollado para la obtención de estos contrastes, en el que se emplean nuevas relaciones matriciales que permiten apreciar la equivalencia entre ellos y facilitan la obtención de su distribución.

En segundo lugar, otra cuestión de interés es la extensión de los contrastes de Nyblom y Mäkeläinen (1983) y Saikkonen y Luukkonen (1993) para la hipótesis nula de una o varias raíces unitarias estacionales. El marco metodológico del Capítulo 3 juega un papel muy importante en el desarrollo de estos estadísticos puesto que permite obtener extensiones novedosas para los contrastes de raíces unitarias en modelos ARIMA, equivalentes a los desarrollados en Canova y Hansen (1995), Harvey y Busetti (2003) o Caner (1998), y empleando de nuevo relaciones matriciales que facilitan su comparación.

Y en tercer lugar se propone un nuevo estadístico basado en los residuos exactos que se puede aplicar en el contraste de hipótesis para una o varias raíces unitarias en un modelo generalizado. El estadístico exacto es el propuesto por Saikkonen y Luukkonen (1993), pero tiene un cómputo complejo y solo se ha desarrollado en el caso más simple. Uno de los objetivos era, por una parte proponer un estadístico aproximado cuyo funcionamiento fuera similar al estadístico exacto de Saikkonen y Luukkonen (1993), incluso en presencia de raíces unitarias adicionales pero con un desarrollo más simple. Para simplificar la obtención del estadístico, el modelo general que se emplea es un modelo estructural donde los errores son un proceso ARMA

$$y_t = \mu_t + u_t, \quad \phi(B)u_t = \theta(B)w_{1t},$$

 $\mu_t = \mu_{t-k} + v_t, \quad \phi(B)v_t = \theta(B)w_{2t}, \quad t = 1, ..., T,$

donde *k* denota el número de frecuencias a contrastar. Aquí juega un papel clave el desarrollo de la matriz de diferencias rectangular ∇_k que diferencia el modelo para hacerlo estacionario

$$\mathbf{z} = \boldsymbol{\nabla}_k \mathbf{e}, \quad \boldsymbol{\nabla}_k \mathbf{e} = \boldsymbol{\nabla}_k \mathbf{C}_k \mathbf{v} + \boldsymbol{\nabla}_k \mathbf{u}.$$

Esta representación tienen una ventaja sobre la que emplean Saikkonen y Luukkonen (1993), y por eso simplifica de forma considerable la obtención del estadístico, y es que

el contraste para el modelo diferenciado es LBI y el contraste para el estadístico exacto es LBIU, y necesita de la segunda derivada de la matriz de varianzas y convarianzas del proceso. El estadístico de contraste se obtiene de forma directa utilizando el Teorema 2 del Capítulo 2 y tiene, de forma general, la siguiente expresión

$$\mathcal{GM}_k = \frac{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(\rho_0)^{-1} \mathbf{\Omega}_{\mathbf{u}} \mathbf{\Omega}_{\mathbf{z}}(\rho_0)^{-1} \mathbf{z}}{\mathbf{z}' \mathbf{\Omega}_{\mathbf{z}}(\rho_0)^{-1} \mathbf{z}}$$

que se puede expresar en términos de ∇_k y la matriz de varianzas y covarianzas de **u**, que recoge la estructura ARMA, y para su cálculo se utilizan los residos exactos. Por tanto, el estadístico propuesto supone algunas ventajas que hacen recomendable su uso. Por una parte tiene una formulación sencilla en comparación al estadístico exacto de Saikkonen y Luukkonen (1993). En segundo lugar se puede aplicar para contrastar distintas hipótesis nulas, tanto regulares como estacionales, y en el contexto de un modelo general. De hecho, la diferencia en la formulación del estadístico para una u otra de las hipótesis de raíces unitarias regulares y estacionales definidas en este tesis, se centra únicamente en las matrices de acumulación de los residuos que se encuentran en el numerador de la forma cuadrática. Otra de las ventajas de este estadístico es que salva algunas de las limitaciones de los contrastes de raíces unitarias existentes puesto que se puede aplicar en presencia de otra raiz MA extra. Además, este estadístico tiene un funcionamiento razonable en muestras finitas si se compara con otros estadísticos que también tienen correcciones para recoger autocorrelación en presencia de raíces unitarias extra.

Para comprobar su funcionamiento se han presentado varios experimentos con el fin de evaluar la distribución muestral de este estadístico en muestras finitas. El primero de ellos considera que el parámetro bajo la hipótesis nula es conocido, y por tanto no hace falta estimarlo; y en el segundo, el parámetro θ_0 se estima por el método de maxima verosimilitud. Los resultados en ambos casos son muy similares y constituyen una evidencia en favor del uso del estadístico basado en los residuos exactos. De hecho, en términos comparativos con otros estadísticos que proponen correcciones para recoger autocorrelación, la distribución muestral del estadístico basado en los residuos exactos es la que más se aproxima a la del estadístico exacto. A partir de aquí, las futuras líneas de investigación que abre esta tesis consisten fundamentalmente en la extensión de los estadísticos obtenidos, y en especial del estadístico basado en los residuos exactos a modelos más complejos, como pueden ser los modelos que incorporan variables de intervencion o multivariantes. APÉNDICE A

Algoritmos para el estadístico de Tam y Reinsel (1997)

Sea el proceso ARIMA(p, 1, q + 1)

$$\phi(B)\nabla y_t = \theta(B)a_t, \quad t = 1, \dots, n, \tag{A.1}$$

para el que se quiere contrastar la hipótesis nula de que el polinomio MA(q + 1) tiene una raíz real unitaria, esto es, $H_0: \theta(B) = 0$ frente a $H_1: \theta(B) \neq 0$ para B = 1. Se presenta ahora un algoritmo para calcular la forma alternativa (??) del estadístico de Tam y Reinsel (1997)

$$\mathcal{TR}_n = \frac{1}{n-1} \frac{\sum_{t=1}^n \left(\sum_{\tau=1}^t \left(\tilde{a}_{\tau} - \tilde{\bar{a}} \right)^2 \right)}{\sum_{t=1}^n \left(\tilde{a}_t - - \tilde{\bar{a}} \right)^2},$$

en donde \tilde{a}_t son los denominados residuos exactos del modelo (A.1).

Siguiendo a Ljung y Box (1979), el modelo (A.1) puede escribirse en forma matricial como

$$\Phi \mathbf{z} = \mathbf{\Theta} \mathbf{a} + \mathbf{G} \mathbf{F} \mathbf{u}$$

en donde $\mathbf{z} = (z_2 \dots z_n)', z_t = \nabla y_t (t = 2, \dots, n), \mathbf{a} = (a_1 \dots a_n)' \mathbf{y} \mathbf{u} = (\mathbf{z}'_0, \mathbf{a}'_0)$ es el vector de valores premuestrales con $\mathbf{z}'_0 = (z_{1-p}, \dots, z_{-1}, z_0)$ and $\mathbf{a}'_0 = (a_{-q}, \dots, a_{-1}, a_0)$. Además, $\mathbf{\Phi}$ es una matriz triangular inferior de orden $(n - 1) \times (n - 1)$ cuya *j*-ésima subdiagonal contiene el parámetro ϕ_j en todas sus entradas ($\phi_0 = 1$ y $\phi_j = 0$ para j > p). Análoga estructura tiene la matriz $\mathbf{\Theta}$, pero con los q + 1 coeficientes MA. Si r = máx(p, q + 1), entonces la matriz $\mathbf{G} = [\mathbf{I}_r | \mathbf{0}]$ tiene orden $n \times r$, en donde la barra vertical denota concatenación vertical. Finalmente, la matriz $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2]$ es la concatenación horizontal de las matrices de orden $r \times p$ y $r \times (q + 1)$, respectivamente,

$$\mathbf{F}_{1} = \begin{pmatrix} \phi_{p} & \phi_{r-1} & \dots & \phi_{1} \\ 0 & \phi_{p} & \dots & \phi_{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{p} \end{pmatrix} \quad \mathbf{y} \quad \begin{pmatrix} \theta_{q+1} & \theta_{r-1} & \dots & \theta_{1} \\ 0 & \theta_{q+1} & \dots & \theta_{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_{q+1} \end{pmatrix},$$

a las que habría añadir filas nulas si p < r o q + 1 < r.

Los residuos exactos de A.1 se definen como

$$E(\mathbf{a}|\mathbf{z}) = E(\mathbf{a}\mathbf{z}')E(\mathbf{z}\mathbf{z}')^{-1}\mathbf{z},$$

en donde $E(\mathbf{a}\mathbf{z}') = \mathbf{\Theta}' \mathbf{\Phi}'^{-1}$, $E(\mathbf{z}\mathbf{z}') = \sigma_a^2 \mathbf{\Omega}_z \mathbf{y}$

$$\mathbf{\Omega}_{z} = \mathbf{\Phi}^{-1} \mathbf{\Theta} [\mathbf{I}_{n-1} + \mathbf{\Theta}^{-1} \mathbf{G} \mathbf{F} \mathbf{\Omega}_{u} \mathbf{F}' \mathbf{G}' \mathbf{\Theta}'^{-1}] \mathbf{\Theta}' \mathbf{\Phi}'^{-1}.$$

Usando una de las expresiones de la inversa de una suma de matrices dadas por Henderson y Searle (1981), se tiene que

$$\boldsymbol{\Omega}_{z}^{-1} = \boldsymbol{\Phi}' \boldsymbol{\Theta}'^{-1} \{ \mathbf{I}_{n-1} - \boldsymbol{\Theta}^{-1} \mathbf{G} [\mathbf{I}_{r} + (\mathbf{F} \boldsymbol{\Omega}_{u} \mathbf{F}') \mathbf{G}' \boldsymbol{\Theta}'^{-1} \boldsymbol{\Theta}^{-1} \mathbf{G}]^{-1} (\mathbf{F} \boldsymbol{\Omega}_{u} \mathbf{F}') \mathbf{G}' \boldsymbol{\Theta}'^{-1} \} \boldsymbol{\Theta}^{-1} \boldsymbol{\Phi},$$

expresión que se prefiere aquí frente a la comúnmente usada

$$\boldsymbol{\Omega}_{z}^{-1} = \boldsymbol{\Phi}^{\prime}\boldsymbol{\Theta}^{\prime-1}\{\mathbf{I}_{n-1} - \boldsymbol{\Theta}^{-1}\mathbf{G}[(\mathbf{F}\boldsymbol{\Omega}_{u}\mathbf{F}^{\prime})^{-1} + \mathbf{G}^{\prime}\boldsymbol{\Theta}^{\prime-1}\boldsymbol{\Theta}^{-1}\mathbf{G}]^{-1}\mathbf{G}^{\prime}\boldsymbol{\Theta}^{\prime-1}\}\boldsymbol{\Theta}^{-1}\boldsymbol{\Phi}$$

que requiere invertir la matriz $\mathbf{F} \mathbf{\Omega}_{u} \mathbf{F}'$ y que puede ser singular en modelos sobreparametrizados con $\phi_{p} = 0$ o $\theta_{q+1} = 0$. Gallego (2009) sugiere calcular esta matriz a partir de la relación

$$\mathbf{F} \mathbf{\Omega}_{u} \mathbf{F} = \mathbf{\Phi}_{r} \mathbf{\Omega}_{z,r} \mathbf{\Phi}_{r}' + \mathbf{\Theta}_{r} \mathbf{\Theta}_{r}',$$

que simplifica considerablemente el cálculo de Ω_u y en donde el subíndice *r* indica de la dimensión de las matrices arriba definidas.

1. Se calcula los residuos denominados condicionados $\hat{a}_0 = \Theta^{-1} \Phi$ usando la ecuación en diferencias

$$\hat{a}_{0,t} = z_t - \phi_1 z_{t-1} - \dots - \phi_p \hat{z}_{t-p} + \theta_1 \hat{a}_{0,t-1} + \dots + \theta_{q+1} \hat{a}_{0,t-q+1}, \quad t = 1, \dots, n,$$

en donde los valores premuestrales se fijan en cero.

2. Se contruye la matriz $\mathbf{X} = \mathbf{\Theta}^{-1}\mathbf{G}$ de orden $n \times r$ usando la siguiente recurrencia para los elementos de la primera columna

$$X_{t,1} = \theta_1 X_{t-1,1} + \dots + \theta_{q+1} X_{t-q-1,1} + g_{t,1}$$

en donde los valores premuestrales son cero y $\delta_{t,1}$ es 1 para t = 1; y 0, en otro caso. Las r - 1 restantes columnas cumplen la relación $X_{t,j} = X_{t-j,1}$.

3. Finalmente, el vector de residuos exactos se calcula como

$$\tilde{\mathbf{a}} = \mathbf{\hat{a}}_0 - \mathbf{X} [\mathbf{I}_r + (\mathbf{F} \mathbf{\Omega}_u \mathbf{F}') \mathbf{X}' \mathbf{X}]^{-1} (\mathbf{F} \mathbf{\Omega}_u \mathbf{F}') \mathbf{X} \mathbf{\hat{a}}_0.$$

APÉNDICE B

Desarrollos del modelo MA(2) en la aproximación de Saikkonen y Luukkonen

B.0.1. Estadístico de contraste

El estadístico de contrasta rechaza la hipótesis nula formulada, $H_0: \Theta = 1$ para valores grandes de

$$\frac{\partial^2 L(\Theta, \boldsymbol{\beta}, \sigma^2)}{\partial \Theta^2} \bigg|_{\Theta = 1, \boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}, \sigma^2 = \tilde{\sigma}^2} > d,$$

donde *d* es una constante arbitraria. Directamente, la expresión de la segunda derivada es

$$\begin{split} \frac{\partial^2 L(\Theta, \boldsymbol{\beta}, \sigma_u^2)}{\partial \Theta^2} \Big|_{\Theta=\Theta_0, \boldsymbol{\beta}=\tilde{\boldsymbol{\beta}}, \sigma_u^2=\tilde{\sigma}_u^2} &= -\frac{1}{2} tr\left(\mathbf{D}_2^{-1} \left(\frac{d\mathbf{\Omega}_2(\Theta)}{d\Theta}\right|_{\Theta=1} (\mathbf{D}_2\mathbf{D}_2')^{-1} \frac{d\mathbf{\Omega}_2(\Theta)}{d\Theta}\Big|_{\Theta=1}\right) \\ &+ \mathbf{D}_2^{-1} \frac{d^2 \mathbf{\Omega}_2(\Theta)}{d\Theta^2}\Big|_{\Theta=1}\right) \mathbf{D}_2'^{-1} \mathbf{M} \mathbf{y} \Big) \\ &+ \frac{T-2}{2} \left(\frac{\mathbf{y}' \mathbf{M} \mathbf{D}_2' \left((\frac{1}{2}c_k \mathbf{B}) \mathbf{D}_2' + \mathbf{D}_2(\frac{1}{2}c_k \mathbf{B})'\right) \mathbf{D}_2 \mathbf{M} \mathbf{y}}{\mathbf{y}' \mathbf{M} \mathbf{y}} \right. \\ &+ \frac{\mathbf{y}' \mathbf{M} \mathbf{D}_2' (2(-c_k \mathbf{B} + \mathbf{B}^2)(-c_k \mathbf{B} + \mathbf{B}^2)') \mathbf{D}_2 \mathbf{M} \mathbf{y}}{\mathbf{y}' \mathbf{M} \mathbf{y}} \Big). \end{split}$$

Además, se consideran los siguentes resultados sobre la matriz de covarianzas $\Omega(\Theta)$:

$$\begin{aligned} \frac{d\mathbf{\Omega}_{2}(\Theta)}{d\Theta}\Big|_{\Theta=1} &= \left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)\mathbf{D}_{2}' + \mathbf{D}_{2}\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)'\\ \frac{d^{2}\mathbf{\Omega}_{2}(\Theta)}{d\Theta^{2}}\Big|_{\Theta=1} &= \left(\frac{1}{2}c_{k}\mathbf{B}\right)\mathbf{D}_{2}' + \mathbf{D}_{2}\left(\frac{1}{2}c_{k}\mathbf{B}\right)' + 2\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)'\\ \frac{d\mathbf{\Omega}_{2}^{-1}(\Theta)}{d\Theta}\Big|_{\Theta=1} &= \left(\mathbf{D}_{2}\mathbf{D}_{2}'\right)^{-1}\left(\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)\mathbf{D}_{2}' + \mathbf{D}_{2}\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)'\right)\left(\mathbf{D}_{2}\mathbf{D}_{2}'\right)^{-1}\\ \frac{d^{2}\mathbf{\Omega}_{2}^{-2}(\Theta)}{d\Theta^{2}}\Big|_{\Theta=1} &= \left(\mathbf{D}_{2}\mathbf{D}_{2}'\right)^{-1}\left(\left(\frac{1}{2}c_{k}\mathbf{B}\right)\mathbf{D}_{2}' + \mathbf{D}_{2}\left(\frac{1}{2}c_{k}\mathbf{B}\right)'\right)\\ &+ 2\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)\left(-c_{k}\mathbf{B} + \mathbf{B}^{2}\right)'\right)\left(\mathbf{D}_{2}\mathbf{D}_{2}'\right)^{-1}\end{aligned}$$

B.0.2. Distribución del estadístico

El denominador $\hat{\mathbf{u}}'\hat{\mathbf{u}}$, se puede expresar como

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \sum_{t=1}^{T} \left(u_t + (T/2)^{-1/2} c_{kt} \sum_{i=1}^{T} c_{ki} u_i + (T/2) s_{kt} \sum_{i=1}^{T} s_{ki} u_i \right)^2$$
$$= \sum_{t=1}^{T} u_t^2 + \sum_{t=1}^{T} (T/2)^{-2} c_{kt}^2 \left(\sum_{i=1}^{T} c_{ki} u_i \right)^2 + \sum_{t=1}^{T} (T/2)^{-2} s_{kt}^2 \left(\sum_{i=1}^{T} s_{ki} u_i \right)^2$$
$$+ 2 \sum_{t=1}^{T} u_t (T/2)^{-1} c_{kt} \left(\sum_{i=1}^{T} c_{ki} u_i \right) + 2 \sum_{t=1}^{T} u_t (T/2)^{-1} s_{kt} \left(\sum_{i=1}^{T} s_{ki} u_i \right)^2$$

Los tres primeros términos,

$$T^{-1}u_t^2 \to \sigma_u^2$$

$$T^{-1}\sum_{t=1}^T (T/2)^{-2} c_{kt}^2 \left(\sum_{i=1}^T c_{ki}u_i\right)^2 \to 0$$

$$T^{-1}\sum_{t=1}^T (T/2)^{-2} s_{kt}^2 \left(\sum_{i=1}^T s_{ki}u_i\right)^2 \to 0$$

y los dos últimos términos son cero ya que $\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} c_{it} c_{jt} = 0$ y $\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} s_{it} s_{jt} = 0$ para $i \neq j$. El numerador del estadístico es $\hat{\mathbf{u}}' \mathbf{C}_k \mathbf{C}_k' \hat{\mathbf{u}}$, donde la suma parcial,

$$\sum_{i=1}^{T} c_{kt} \left(u_t + (T/2)^{-1/2} c_{kt} \sum_{i=1}^{T} c_{kt} u_t + (T/2)^{-1/2} s_{kt} \sum_{i=1}^{T} s_{kt} u_t \right)$$
$$= \sum_{t=1}^{T} c_{kt} u_t + (T/2)^{-1} \sum_{t=1}^{T} c_{kt}^2 \left(\sum_{i=1}^{T} c_{ki} u_i \right)$$

donde

$$T^{-1/2}\sum_{t=1}^{[Tr]} \left(c_{kt}u_t - c_{kt}^2 \left(\frac{\sum_{i=1}^T c_{ki}u_i}{T/2} \right) \right)$$

Análogamente, para la componente $\hat{\mathbf{u}}' \mathbf{S}_k \mathbf{S}'_k \hat{\mathbf{u}}$, donde la suma parcial,

$$\sum_{i=1}^{T} s_{kt} \left(u_t + (T/2)^{-1/2} c_{kt} \sum_{i=1}^{T} c_{kt} u_t + (T/2)^{-1/2} s_{kt} \sum_{i=1}^{T} s_{kt} u_t \right)$$
$$= \sum_{t=1}^{T} s_{kt} u_t + (T/2)^{-1} \sum_{t=1}^{T} s_{kt}^2 \left(\sum_{i=1}^{T} s_{ki} u_i \right)$$

donde

$$T^{-1/2} \sum_{t=1}^{[Tr]} \left(s_{kt} u_t - s_{kt}^2 \left(\frac{\sum_{i=1}^T s_{ki} u_i}{T/2} \right) \right)$$

por tanto para el numerador tenemos que $T^{-1/2} \sum_{t=1}^{[Tr]} (c_{kt}\hat{u}_t + s_{kt}\hat{u}_t) \rightarrow \sigma_u (V_{k,1}(r) + V_{k,2}(r)),$ donde $V_{k,i}(r) = W_{k,i}(r) - rW_{k,i}(1)$ es un puente Browniano estándar.

Bibliografía

- T. W. Anderson. Statistical analysis of time series. John Wiley and Sons, New York, 1992.
- T. W. Anderson y D. A. Darling. Asymptotic theory of certain "goodness of fitçriteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2):193–212, Jun. 1952.
- C.F. Ansley. An algorithm for the exact likelihood of a mixed autoregressive-moving average model. *Biometrika*, 66(1):59–65, 1979.
- C.F. Ansley y P. Newbold. Finite sample properties of estimators for autoregressive moving average models. *Journal of Econometrics*, 70:159–183, 1983.
- S. Beveridge y C.R. Nelson. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to a new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics*, 7:151–174, 1981.
- Harald Bohman. From characteristic function to distribution function via fourier analysis. *BIT Numerical Mathematics*, 12:279–283, 1972.
- G. E. P. Box, G. M. Jenkins, y G. C. Reinsel. *Time series analysis, forecasting and control*.Wiley, New Jersey, 4th edition, 2008.
- G. E. P. Box, G. M. Jenkins, y G. C. Reinsel. *Time series analysis: forecasting and control*.Wiley, New Jersey, 5th edition, 2015.
- F. Busetti y A. Harvey. Testing for the presence of a random walk in series with structural breaks. *JTSA*, 22(2):127–150, March 2001.

- F. Busetti y A.M.R. Taylor. Seasonality tests. *Journal of Econometrics*, 21(1):21–53, November 2003.
- M. Caner. A locally optimal seasonal unit-root test. *Journal of Business & Economic Statistics*, 16(3):349–356, Jul. 1998.
- F. Canova y B.E. Hansen. Are seasonal patterns constant over time? a test for seasonal stability. *JBES*, 13(3):237–252, July 1995.
- In Choi. *Almost All About Unit Roots: Foundations, Developments, and Applications*. Cambridge University Press, Cambridge, 2015.
- D. Coniffe y J.E. Spencer. When moments of ratios are ratios of moments. *The Statistician*, 50(2):161–168, 2001.
- J. D. Cryer y J. Ledolter. Small-sample properties of the maximum likelihood estimator in the first- order moving average model. *Biometrika*, 68(3):691–694, Dec. 1981.
- R. B. Davies. Numerical inversion of a characteristic function. *Biometrika*, 60(2):415–417, 1973.
- R. B. Davies. Algorithm as 155: The distribution of a linear combination of χ^2 random variables. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 29(3):323–333, 1980.
- W. Dent. Computation of the exact likelihood function of an arima process. *Journal of Statistical Computation and Simulation*, 5:193–206, 1977.
- D. A. Dickey, D. P. Hasza, y W. A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 79: 355–367, 1984.
- David A. Dickey y Wayne A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366): 427–431, jun 1979.
- R. W. Farebrother. The distribution of a quadratic form in normal variables. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 39:294–309, 1990.

- T. S. Ferguson. *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York, 1967.
- J. L. Gallego. The exact likilihood function of a vector autoregressive moving average process. *Statistics and Probability Letters*, 79:711–714, 2009.
- J.L. Gallego y C. Díaz. Testing for deterministic components in vector seasonal time series. *Open Journal of Statistics*, 1(3):145–150, 2011.
- J.L. Gallego y A.B. Treadway. The general family of seasonal stochastic procresses. Departamento de Economía, Universidad de Cantabria, 1995.
- R.C. Geary. A general expression for the moments of certain symmetrical functions of normal samples. *Biometrika*, 25(1/2):184–186, may 1933.
- J. Gil-Pelaez. Note on the inversion theorem. *Biometrika*, 38(3/4):481–482, 1951.
- James D. Hamilton. Time series analysis. Princeton University Press, Princeton, NJ, 1994.
- Andrew C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge, 1989.
- Andrew C. Harvey y Fabio Busetti. Seasonality tests. *Journal of Business & Economic Statistics*, 21(3):420–436, 2003.
- H. V. Henderson y S. R. Searle. On deriving the inverse of a sum of matrices. *SIAM Review*, 23(1):53–60, Jan. 1981.
- S. Hylleberg, R.F. Engle, C.W.J. Granger, y B.S. Yoo. Seasonal integration and cointegration. *JE*, 44:215–238, 1990.
- J.P. Imhof. Computing the distribution of cuadratic forms in normal variables. *Biometrika*, 48:419–426, 1961.
- K. R. Kadiyala. Testing for the independence of regression disturbences. *Econometrica*, 38(1):97–117, 1970.
- T. Kariya. A robustness property of the test for serial correlation. *The Annals of Statistics*, 5(6):1212–1220, 1977.
- T. Kariya. Locally robust test for serial correlation in least squares regression. *The Annals of Statistics*, 8(5):1065–1070, 1980.
- M.L. King. Robust tests for spherical symmetry and their application to least squares regression. *The Annals of Statistics*, 8(6):1265–1271, 1980.
- M.L. King y G.H. Hillier. Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society. Series B*, 47(1): 98–102, 1985.
- M.L. King y P.X. Wu. Locally optimal one-sided tests for multiparameter hypothesis. *Econometric Reviews*, 16:131–156, 1997.
- J. Koerts y A.P.J. Abrahamse. *On the theory and application of the general linear model*. Rotterdam University Press, Rotterdam, 1971.
- D. Kwiatkowski, P.C.B. Phillips, P. Schmidth, y Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics*, 54:159–178, 1992.
- E.L. Lehman. Testing statistical hypothesis. Wiley-Interscience, New York, 1959.
- E.L. Lehman. Testing statistical hypothesis. Wiley-Interscience, New York, 1986.
- S J Leybourne y B P M McCabe. A consistent test for a unit root. *Journal of Business & Economic Statistics*, 12(2):157–66, 1994.
- G.M. Ljung y G.E.P. Box. The likelihood function of stationary autoregressive-moving average models. *Biometrika*, 66(2):265–70, 1979.
- Z. Lu y M.L. King. Improving the numerical technique for computing the accumulated distribution of a quadratic form in normal variables. *Econometric Reviews*, 21(2):149– 65, 2002.
- I.B. MacNeill. Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Annals of Statistics*, 6:422–433, 1978.

- J. Nyblom. Testing for deterministic linear trend in time series (in theory and methods). *Journal of the American Statistical Association*, 81(394):545–549, Jun. 1986.
- J. Nyblom y A. C. Harvey. Testing against smooth stochastic trends. *Journal of Applied Econometrics*, 16(3):415–429, 2001.
- J. Nyblom y T. Mäkeläinen. Comparisons of test for the presence of random walk coefficientes in a simple linear model. *Journal of the American Statistical Association*, 78 (384):856–864, 1983.
- Kerry Patterson. *Unit Root Tests in Time Series Volume 1: Key Concepts and Problems*. Palgrave Macmillan, London, 2011.
- P. Perron. Testing for a unit root in a time series with a changing mean. *Journal of Business & Economic Statistics*, 8(2):153–162, Apr. 1990.
- P.C.B. Phillips y P. Perron. Testing for a unit root in time series regression. *Biometrika*, 75:335–346, 1988.
- B. M. Pötscher. Noninvertibility and pseudo-maximum likelihood estimation of misspecified arma models. *Econometric Theory*, 7(4):435–449, 1991.
- S. E. Said y D. A. Dickey. Hypothesis testing in arima(p, 1, q) models. *Journal of the American Statistical Association*, 80:369–374, 1985.
- P. Saikkonen y R. Luukkonen. Testing for a moving average unit root in autoregressive integrated moving average models. *JASA*, 88(422):596–601, 1993.
- J. D. Sargan y Alok Bhargava. Maximum likelihood estimation of regression models with first order moving average errors when the root lies on the unit circle. *Econometrica*, 51(3):799–820, 1983.
- W. Tam y G.C. Reinsel. Tests for seasonal moving average unit root in arima models. *JASA*, 92(438):725–738, 1997.
- W. Tam y G.C. Reinsel. Seasonal moving-average unit root tests in the presence of a linear trend. *Journal of Time Series Analysis*, 19(5):609–625, September 1998.

- K. Tanaka. Testing for a moving average unit root. *Econometríc Theory*, pages 433–44, 1990.
- K. Tanaka. *Time Series Analysis: nonstationary and noninvertible distribution theory*. Wiley, New York, 1996.
- A. M. R. Taylor. Locally optimal tests against unit roots in seasonal time series processes. *Journal of Time Series Analysis*, 24(5):591–612, 2003.