

043
101

UNIVERSIDAD DE SEVILLA

UNIVERSIDAD DE SEVILLA
FACULTAD DE MATEMATICAS
BIBLIOTECA

FACULTAD DE MATEMATICAS

DEPARTAMENTO DE ESTADISTICA MATEMATICA

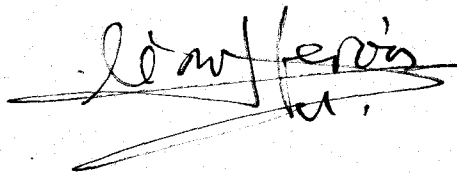
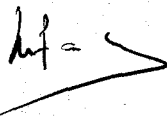
NUEVOS TESTS PARA CONTRASTAR LA BONDAD DE MODELOS ARMA
Y MARMA DE SERIES TEMPORALES.

Visado en Sevilla

Octubre de 1984.

El Director Prof.Dr. José Maria
Caridad y Ocerín.El Ponente Prof.Dr. Rafael
Infante Macias.

Memoria que presenta Cesar Hervás
Martinez para optar al grado de
Doctor en Ciencias Matemáticas.
Sevilla. Octubre de 1984.



NUEVOS TESTS PARA CONTRASTAR LA BONDAD DE MODELOS

ARMA Y MARCA DE SERIES TEMPORALES

28-11-84

2469

UNIVERSIDAD DE SEVILLA

FACULTAD DE MATEMATICAS

NUEVOS TESTS PARA CONTRASTAR LA BONDAD DE MODELOS ARMA
Y MARMA DE SERIES TEMPORALES.

CESAR HERVAS MARTINEZ

Memoria para optar al grado de Doctor
en Ciencias Matemáticas realizada
bajo la dirección de D. José María
Caridad y Ocerín en el Departamento
de Estadística de la Universidad de
Córdoba.

C O N T E N I D O

Pag.

CAPITULO I.- "TESTS PARA DIAGNOSTICAR MODELOS ARMA: TEST DEL ESTADISTICO PORTMANTEAU".

1.0.- Sumario	3
1.1.- Introducción.....	4
1.2.- Tests basados en el estadístico "portmanteau".....	5
1.3.- Lemas y Teorema acerca de la distribución asintótica de \hat{r}	8
1.4.- Tests estadísticos basados en \hat{r} : $Q(\hat{r})$ y $\tilde{Q}(\hat{r})$	17
1.5.- Test para un m no necesariamente de orden superior a \sqrt{n}	20
1.6.- Conclusiones	21

CAPITULO II.- "METODO DE LOS SCORES EN EL CONTRASTE DE HIPOTESIS : APLICACION DEL TEST DE LOS MULTIPLICADORES DE LAGRANGE PARA CONTRASTAR LA BONDAD DE MODELOS ARMA"

2.0.- Sumario	25
2.1.- Introducción	26
2.2.- Propositiones y lemas previos para hallar la distribución asintótica del test de los multiplicadores de Lagrange	29
2.3.- Estimaciones restringidas	39
2.4.- Test de los multiplicadores de Lagrange para matrices de información singulares	46
2.5.- Aplicación del Test de los multiplicadores de Lagrange para contrastar la bondad de modelos ARMA	57
2.6.- Equivalencia asintótica del test del estadístico portmanteau y del "score" test en un caso particular	75

CAPITULO III.- "SCORE TEST Y ESTADISTICO PORTMANTEAU PARA DIAGNOSTICAR
MODELOS MARMA".

3.0.- Sumario	84
3.1.- Score test para diagnosticar modelos MARMA.....	85
3.2.- Estadístico Portmanteau para modelos MARMA	103
3.3.- Cálculo práctico del estadístico del score test	113
3.4.- APENDICE	117
BIBLIOGRAFIA.....	123

AGRADECIMIENTOS.-

Deseo expresar mi más profundo agradecimiento al Profesor Dr. D. José Maria Caridad y Ocerín por la constante ayuda e interés puesto en la realización de esta monografía, que hago extensivo a mis compañeros de Departamento.

Igualmente quiero agradecer al profesor Dr. D. Rafael Infante Macias las facilidades y comprensión prestados para la presentación de esta memoria, y en general a todo el Departamento de Estadística de la Facultad de Ciencias Matemáticas de la Universidad de Sevilla.

Por último, agradezco a la Srta. María José Hidalgo la mecanografía de estas páginas.

Córdoba, Septiembre de 1984.

A mi madre

PROLOGO.-

La presente memoria consta de 3 capítulos en donde se realizan estudios de los problemas de contraste de hipótesis acerca de la bondad de modelos ARMA y MARMA de series temporales.

Desde los estudios precursores de una nueva metodología acerca del análisis de las series temporales, BOX y JENKINS en 1970, hasta nuestros días, han aparecido una gran diversidad de trabajos, publicaciones y software de ordenador acerca de los problemas de Identificación, Estimación, Contraste y Predicción de tales series con publicaciones específicas como Journal of Time Series Analysis ó Journal of Forecasting.

Nuestro trabajo se ha centrado en el problema del contraste del modelo, donde dada la complejidad del mismo los métodos habituales no son siempre fáciles de aplicar, adaptándose métodos estadísticos generales y desarrollando otros originales.

En el capítulo primero se estudian los estadísticos del tipo portmanteau, aplicables al chequeo del modelo, pero sin un modelo alternativo de contraste, dándose sus distribuciones asintóticas así como una valoración de las mismas.

Continuamos la memoria con un segundo capítulo en el que una vez abordada la teoría del contraste de hipótesis para estimadores restringidos, se ha estructurado para poderla aplicar a modelos AR y ARMA frente a diferentes familias de modelos alternativos, resolviéndose los problemas de singularidad en la matriz de covarianzas de los parámetros del modelo, usando nuevos estadísticos.

Continúa el capítulo con el estudio asintótico

de las distribuciones de los estadísticos usados en estos contrastes, para terminar con la equivalencia asintótica de ambos tests (portmanteau y score) en un caso concreto de familia de modelos alternativos.

En el tercer capítulo se generaliza el test tratado en el capítulo anterior, para contrastar modelos múltiples de series temporales (MARMA) usando en el desarrollo de las demostraciones técnicas de álgebra tensorial.

Sigue el capítulo con un estudio del estadístico definido en este caso, en el que se hace hincapié no solo en su distribución asintótica sino en su cálculo práctico con vistas a su inclusión en el paquete de programas W.M.T.S. de la Universidad de Wisconsin.

Acabamos con la equivalencia asintótica de los tests portmanteau y score y con un Apéndice donde se propone una transformación propia que permite simplificar las demostraciones de los distintos resultados.

C A P I T U L O I

TESTS PARA DIAGNOSTICAR MODELOS ARMA : TEST DEL
ESTADISTICO PORTMANTEAU.

"TESTS PARA DIAGNOSTICAR MODELOS ARMA: TEST DEL ESTADISTICO PORTMANTEAU"

1.0.-	Sumario	3
1.1.-	Introducción	4
1.1.1.-	Primer Método : Sobreparametrización	4
1.1.2.-	Segundo Método: Contraste usando los residuos..	5
1.2.-	Tests basados en el estadístico "Portmanteau".....	5
1.2.1.-	Definiciones de las funciones $\tilde{Q}(\vec{r})$ y $Q(\vec{r})$. Resultados asintóticos de las mismas.....	6
1.2.2.-	Cálculo de las medias y varianzas de $\tilde{Q}(\vec{r})$ y $Q(\vec{r})$ para valores finitos de n	6
1.3.-	Lemas y Teorema acerca de la distribución asintótica de \hat{r}	8
1.3.1.-	Lema 1ª	8
1.3.2.-	Lema 2ª	9
1.3.3.-	Lema 3ª	12
1.3.4.-	Teorema	14
1.3.5.-	Consecuencia del Teorema.....	16
1.4.-	Tests estadísticos basados en \hat{r} , $Q(\hat{r})$ y $\tilde{Q}(\hat{r})$	17
1.4.1.-	Cálculo de las medias de los estadísticos $Q(\hat{r})$ y $\tilde{Q}(\hat{r})$	17
1.4.2.-	Distribuciones asintóticas de $Q(\hat{r})$ y $\tilde{Q}(\hat{r})$	19
1.5.-	Test para un m no necesariamente de orden superior a \sqrt{n} ..	20
1.6.-	Conclusiones	21

1.0.- SUMARIO

En este primer capítulo, realizamos una exposición de los aspectos mas importantes del chequeo de un modelo ARMA aplicables en el análisis de una serie temporal en el dominio del tiempo, según los desarrollos de BOX-JENKINS (1976) y HANNAN (1970).

Empezamos con una introducción, donde ademas de fijar el modelo, se citan los dos métodos que fundamentalmente se usan para contrastar la bondad de dicho modelo, ciñéndonos a continuación al método que usa el estadístico portmanteau ó modificaciones del mismo, estadístico que es función del vector $\hat{\vec{r}}$ de autocorrelaciones muestrales del modelo, donde los parámetros del mismo se han sustituido por sus estimadores eficientes.

Despues de citar la distribución asintótica del vector de autocorrelaciones muestrales \vec{r} ; ANDERSON (1942), ANDERSON y WALKER (1964) y HANNAN (1970), se definen las funciones de \vec{r} , $Q(\vec{r})$ y $\tilde{Q}(\vec{r})$, así como sus medias y varianzas. A continuación se estudia $\hat{\vec{r}}$ y se desarrolla su distribución asintótica, para a partir de ella, obtener las de $Q(\hat{\vec{r}})$ y $\tilde{Q}(\hat{\vec{r}})$, en dicho desarrollo se combinan resultados obtenidos por BOX-PIERCE (1970), LJUNG-BOX (1978), IAN MAC LEOD (1977a y 1978), HANNAN (1970) y SILVEY (1959). Se termina el capítulo con la definición del estadístico S, basado también en $\hat{\vec{r}}$, pero donde m no necesariamente debe ser de orden superior a \sqrt{n} , concluyendo con una valoración de los tres estimadores.

TESTS PARA DIAGNOSTICAR MODELOS "ARMA" : TEST DEL ESTADISTICO
PORTMANTEAU

1.1. INTRODUCCION.

Sea el modelo ARMA para una serie observada : $\{x_t\} = \{x_1, \dots, x_n\}$

$$(1.1) \quad \sum_{j=0}^p \phi_j x_{t-j} = \sum_{j=0}^q \theta_j a_{t-j} \quad \text{con} \quad \phi_0 = \theta_0 = 1$$

La sucesión a_t es una sucesión de variables aleatorias independientes e idénticamente distribuidas $N(0, \sigma^2)$.

Si se introducen las funciones generatrices

$$\phi(z) = \sum_{i=0}^p \phi_i z^i \quad \text{y} \quad \theta(z) = \sum_{i=0}^q \theta_i z^i$$

entonces se define la región \mathcal{R} de \mathbb{C}^{p+q} donde $\phi(z) \neq 0$ y $\theta(z) \neq 0$ con $|\bar{z}| \leq 1$ y donde $\phi(z)$ y $\theta(z)$ no tienen ceros comunes, estas condiciones son suficientes para que el modelo sea estacionario e invertible (Hannan (1970) , Box-Jenkins (1976)).

Nuestro objetivo no solo es determinar la bondad del modelo, sino sugerir el camino en el que el modelo debe ser modificado.

Dos métodos son los que fundamentalmente se usan para contrastar si el modelo esta correctamente especificado.

1.1.1. Primer Método: Sobreparametrización.

La sobreparametrización consiste en elegir

p y q de forma tal que el modelo (1.1) se obtenga anulando ciertos parámetros del modelo más general, al estadístico que se construye se le denomina "score" pues está basado en el vector de las derivadas, respecto a los parámetros del modelo, del logaritmo de la función de verosimilitud.

1.1.2.- Segundo Método: Contraste usando los residuos.

El estadístico utilizado se construye a partir de las autocorrelaciones muestrales de los residuos $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_n)$, esto es

$$\hat{r}_k = \frac{\sum_{t=1}^{n-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^n \hat{a}_t^2}$$

donde estos se calculan usando el modelo ARMA estimado

$$\hat{\phi}(B) x_t = \hat{\theta}(B) \hat{a}_t,$$

siendo B el operador de retardo, esto es, $Bx_t = x_{t-1}$

A este estadístico se le llama "portmanteau".

1.2.- TEST BASADO EN EL ESTADISTICO "PORTMANTEAU".

Si el modelo (1.1) es el adecuado, entonces las v.a. a_t estarán incorreladas, $E(a_t a_{t-k}) = 0 \quad \forall k \neq 0$ y sus m primeras autocorrelaciones muestrales $\vec{r} = (r_1, r_2, \dots, r_m)'$,

siendo

$$r_k = \frac{\sum_{t=1}^{n-k} a_t a_{t+k}}{\sum_{t=1}^n a_t^2}$$

se distribuyen asintóticamente según una distribución normal con vector de medias cero y matriz de varianzas-covarianzas con elementos de la forma siguiente:

$$(1.2) \quad V(r_k) = \frac{n-k}{n(n+2)} \quad y \quad \text{Cov}(r_k, r_l) = 0 \quad \text{si } k \neq l$$

(Anderson (1942), Anderson-Walker (1964) y Hannan (1970))

1.2.1.- Definiciones de las funciones $\tilde{Q}(\vec{r})$ y $Q(\vec{r})$ y resultados asintóticos de las mismas.

De los resultados anteriores se deduce que la distribución asintótica de la variable aleatoria

$$\tilde{Q}(\vec{r}) = n(n+2) \sum_{k=1}^m (n-k)^{-1} r_k^2 \quad \text{es una } \chi_m^2$$

Análogamente, si suponemos que $V(r_k)$ es del mismo orden que $1/n$, entonces la distribución del otro estadístico.

$$Q(\vec{r}) = n \sum_{k=1}^m r_k^2 \quad \text{es asintóticamente } \chi_m^2$$

(Box-Pierce 1970).

1.2.2.- Cálculo de las medias y varianzas de $Q(\vec{r})$ y $\tilde{Q}(\vec{r})$ para valores finitos de n .

Cálculo de las medias.

En este caso se demuestra que $E(\tilde{Q}(\vec{r})) = m$

mientras que

$$(1.2.2.1) \quad E(Q(\vec{r})) = \frac{m+n}{n+2} \left(1 - \frac{m+1}{2+n}\right)$$

valor que es inferior al que toma la esperanza asintótica de la variable, luego el estadístico $Q(\vec{r})$ tiene una distribución, para valores de n pequeños, que discrepa de la χ_m^2 .

Cálculo de las varianzas de $Q(\vec{r})$ y $\tilde{Q}(\vec{r})$.

El cálculo de las varianzas de $\tilde{Q}(\vec{r})$ y $Q(\vec{r})$ se deduce haciendo uso de la identidad

$$E(r_k^i r_l^j) = \frac{E\left\{ \left(\sum a_t a_{t+k}\right)^i \left(\sum a_t a_{t+l}\right)^j \right\}}{E\left(\sum a_t^2\right)^{i+j}}$$

que se sigue de la independencia de las variables aleatorias

$$r_k \quad \text{y} \quad \sum_{t=1}^n a_t^2 \quad (\text{Anderson 1971})$$

Suponemos además que $V(a_i)=1$, sin perder generalidad por ello, y entonces

$$\sum_{i=1}^n a_i^2 \in \chi_n^2$$

luego

$$E\left(\sum a_i^2\right)^{i+j} = n(n+2) \cdot \dots \cdot (n+2i+2j-2)$$

Usando la fórmula de Leibnitz de la potencia n-ésima de un polinomio, se demuestra por el método de inducción transfinita que si $k < 1/2 n$

$$V(r_k^2) = E(r_k^4) - V^2(r_k) = \frac{6(3n-5k)+3(n-k)^2}{n(n+2)(n+4)(n+6)} - \frac{(n-k)^2}{n^2(n+2)^2}$$

y que

$$\text{Cov}(r_k^2, r_l^2) = \frac{(n-k)(n-l)+4(n-l)+8(n-k-l)}{n(n+2)(n+4)(n+6)} - \frac{(n-k)(n-l)}{n^2(n+2)^2}$$

Teniendo en cuenta los resultados anteriores, ya es fácil ver que

$$V\{Q(\vec{r})\} \approx 2m \left(1 + \frac{m-10}{n}\right)$$

donde se ha supuesto que $n \gg m$ y se ignoran los términos de orden superior a $1/n$.

Análogamente, con idénticos supuestos, se prueba que

$$V\{\tilde{Q}(\bar{r})\} \approx 2m \left(1 + \frac{2m-5}{n}\right)$$

1.3.-DISTRIBUCION ASINTOTICA DE \hat{r} :

Sea $\hat{r} = (\hat{r}_1, \dots, \hat{r}_m)'$ el vector de las autocorrelaciones muestrales del modelo (1.1) calculadas sustituyendo el vector de parámetros del modelo $\bar{\beta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$ por un estimador asintóticamente eficiente del mismo $\hat{\beta} = (\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)'$ por ejemplo, el obtenido por mínimos cuadrados en (Box-Jenkins (1976))

Para obtener la distribución asintótica de \hat{r} , lo haremos a partir de relaciones lineales de \hat{r}_k en función de \bar{r}_k , donde aplicamos el resultado obtenido en (1.2) y desarrollando el trabajo de I. MacLeod (1978) se siguen los siguientes lemas:

1.3.1.-Lema 1^a

Para cualquier m fijo.

$$r_k = Q_k + o(1/n) \quad (1 \leq k \leq m)$$

siendo

$$Q_k = n^{-1} \sum_{t=1}^{n-k} a_t a_{t-k} \quad \text{y} \quad Q_0 = n^{-1} \sum_{t=1}^n a_t^2$$

Demostración.-

Desarrollando en serie de Taylor la expresión $r_k = \frac{Q_k}{Q_0}$ como función de (Q_k, Q_0) , en un entorno de $(0, 1)$. ya que $E(Q_k) = 0$ y $E(Q_0) = 1$

se tiene que

$$g(Q_k, Q_0) = g(0,1) + \frac{\int g}{\int Q_k} \Big|_{(0,1)} (Q_k - 0) + \frac{\int g}{\int Q_0} \Big|_{(0,1)} (Q_0 - 1) + \\ + \frac{1}{2!} \frac{\int^2 g}{\int Q_k^2} \Big|_{(0,1)} (Q_k - 0)^2 + \frac{1}{2!} \frac{\int^2 g}{\int Q_0^2} \Big|_{(0,1)} (Q_0 - 1)^2 + \frac{1}{2!} \frac{\int^2 g}{\int Q_0 \int Q_k} \Big|_{(0,1)} (Q_k - 0)(Q_0 - 1) + \dots$$

luego

$$r_k = \frac{Q_k}{Q_0} = g(Q_k, Q_0) = 0 + \frac{1}{Q_0} \Big|_{(0,1)} Q_k - \frac{Q_k}{Q_0^2} \Big|_{(0,1)} (Q_0 - 1) + \frac{1}{2!} 0 + \\ + \frac{1}{2!} \frac{2 Q_k}{Q_0^3} \Big|_{(0,1)} (Q_0 - 1)^2 + \frac{1}{2!} \left[-\frac{1}{Q_0^2} \right]_{(0,1)} (Q_k - 0)(Q_0 - 1) + \dots$$

así

$$r_k \equiv Q_k - \frac{1}{2!} Q_k (Q_0 - 1) + \dots$$

como queriamos demostrar

1.3.2-LEMA 2º.

Sea $\hat{\beta}$ el estimador mínimo cuadrático de $\vec{\beta}$ obtenido al minimizar, la suma de residuos al cuadrado $S(\hat{\beta}) = \sum_{t=1}^n \dot{a}_t^2$ siendo $\vec{\beta}$ cualquier solución admisible del espacio paramétrico Θ

Entonces

$$\hat{\beta} - \vec{\beta} = \begin{bmatrix} \delta_{vv} & \delta_{uv} \\ \delta_{uv} & \delta_{uu} \end{bmatrix}^{-1} \bar{s} + o(1/n)$$

donde el vector \bar{s} se define en la forma

$$s_i = \begin{cases} -n^{-1} \sum_{t=1}^n a_t v_{t-i} & \text{para } 1 \leq i \leq p \\ -n^{-1} \sum_{t=1}^n a_t u_{t+p-i} & \text{para } p+1 \leq i \leq p+q \end{cases}$$

Demostración.-

(Hannan 1970) demuestra que $\hat{\beta}$ se distribuye, asintóticamente, según una $N(\vec{\beta}, n^{-1} \Lambda^{-1})$

donde

$$\Lambda = \begin{bmatrix} \gamma_{vv}(i-j) & \gamma_{vu}(i-j) \\ \gamma_{uv}(i-j) & \gamma_{uu}(i-j) \end{bmatrix} \begin{matrix} p \\ q \end{matrix}$$

Siendo γ_{vv} , γ_{vu} y γ_{uu} las matrices de varianzas-covarianzas de las variables u y v definidas en la forma

$$(1.3.2.1) \quad \phi(B)v_t = -a_t \quad \text{y} \quad \theta(B)u_t = a_t$$

es decir

$$(1.3.2.2) \quad \gamma_{vv} = (\gamma_{vv}(k)) \quad \text{con} \quad \gamma_{vv}(k) = E(v_t v_{t+k}) = E(v_t v_{t-k})$$

$$\gamma_{vu} = (\gamma_{vu}(k)) \quad \text{y} \quad \gamma_{uv} = (\gamma_{uv}(k)) = (\gamma_{vu}(-k)) \quad \text{con}$$

$$\gamma_{vu}(k) = E(v_t u_{t+k})$$

por último

$$\gamma_{uu} = (\gamma_{uu}(k)) \quad \text{con} \quad \gamma_{uu}(k) = E(u_t u_{t+k}) = E(u_t u_{t-k})$$

Para cada $\vec{\beta}$, si reagrupamos el modelo (1.1), tenemos que la serie ruido blanco estimada \dot{a}_t se calcula de la forma

$$(1.3.2.3) \quad \dot{a}_t = -\sum_{j=1}^q \theta_j a_{t-j} + \sum_{j=0}^p \phi_j x_{t-j}$$

para $p+1 \leq t \leq n$

y

$$\dot{a}_t = 0 \quad \text{para} \quad t \leq p$$

Estamos en condiciones de calcular el vector $\frac{\int S}{\int \beta}$

$$\frac{\int S(\vec{\beta})}{\int \beta} = \frac{\int \sum \dot{a}_t^2}{\int \beta} = \frac{\int \dot{a}' \dot{a}}{\int \beta} = 2 \frac{\int \dot{a}}{\int \beta} \dot{a}$$

y si descomponemos $\frac{J S(\vec{\beta})}{J \vec{\beta}}$

en $\frac{J S(\vec{\beta})}{J \vec{\phi}}$ y $\frac{J S(\vec{\beta})}{J \vec{\theta}}$

donde $\vec{\beta}' = (\vec{\phi}' : \vec{\theta}')$ con $\vec{\phi} = (\phi_1, \dots, \phi_p)'$

y $\vec{\theta} = (\theta_1, \dots, \theta_q)'$

entonces

$$\frac{J S}{J \phi_i} = 2 \sum_{t=1}^n \dot{a}_t v_{t-i}$$

para $i = 1, \dots, p$

$$y \quad \frac{J S}{J \theta_i} = 2 \sum_{t=1}^n \dot{a}_t u_{t-i}$$

para $i = 1, \dots, q$

donde
$$v_{t-i} = \frac{J \dot{a}_t}{J \phi_i}$$

y
$$u_{t-i} = \frac{J \dot{a}_t}{J \theta_i}$$

Además por definición, $\hat{\vec{\beta}}$ minimiza $S(\vec{\beta})$, luego

$$\left. \frac{J S}{J \vec{\beta}} \right|_{\vec{\beta} = \hat{\vec{\beta}}} = \vec{0}$$

Ahora como $S(\vec{\beta})$ es una función de $\vec{\beta} \in \mathcal{Q} \subset \mathcal{R} \subset \mathbb{C}^{p+q}$ y como

\mathcal{R} es abierto (Deitstler. H; Dunsmuir. W y Hannan. E.J. 1978), si suponemos que existen las derivadas parciales de $S(\vec{\beta})$, y que son finitas, desarrollando en serie de Taylor, para $\vec{\beta} = \hat{\vec{\beta}}$, la expresión $J S / J \vec{\beta}$, se tiene

$$\frac{J S}{J \vec{\beta}} = \left. \frac{J S}{J \vec{\beta}} \right|_{\vec{\beta} = \hat{\vec{\beta}}} + \left. \frac{J^2 S}{J \vec{\beta} J \vec{\beta}} \right|_{\vec{\beta} = \hat{\vec{\beta}}} (\hat{\vec{\beta}} - \vec{\beta}) + o(\hat{\vec{\beta}} - \vec{\beta})$$

luego
 (1,3,2,4)
$$\hat{\vec{\beta}} - \vec{\beta} = \frac{-\frac{1}{2} \frac{\int S}{\int \vec{\beta}}}{\frac{1}{2} \frac{\int^2 S}{\int \vec{\beta} \int \vec{\beta}'}} + o(1/n)$$

ya que $(\hat{\vec{\beta}} - \vec{\beta})$ son infinitesimos de orden superior a $1/n$ y las derivadas parciales de $S(\vec{\beta})$ de cualquier orden, son de orden \sqrt{n} .

Ahora teniendo en cuenta el resultado, (Poskitt y Tremayne (1981))

en el que

$$\frac{1}{2} \frac{\int^2 S}{\int \vec{\beta} \int \vec{\beta}'} = n \Lambda + o(\sqrt{n})$$

sigue que sustituyendo este resultado en (1.3.2.4)

$$\hat{\vec{\beta}} - \vec{\beta} = \frac{-\frac{1}{2} \frac{\int S}{\int \vec{\beta}}}{n \Lambda} + o(1/n) = \Lambda^{-1} \left(-\frac{1}{2n} \frac{\int S}{\int \vec{\beta}} \right) + o(1/n) = \Lambda^{-1} \vec{s} + o(1/n)$$

siendo \vec{s} el vector definido en el enunciado del lema, como queriamos demostrar.

1.3.3.-LEMA 3º

La distribución conjunta de $\sqrt{n} (\hat{\vec{\beta}} - \vec{\beta}, \vec{r}) \xrightarrow{\bar{L}} N(0, C)$ siendo

C , la matriz de varianzas-covarianzas de la forma

$$C = \begin{bmatrix} \Lambda^{-1} & -\Lambda^{-1} X' \\ -X \Lambda^{-1} & I \end{bmatrix}$$

donde la matriz $X = (\underline{\phi}^* ; \underline{\theta}^*)$

siendo

$$\underline{\phi}^* = (\phi_{i-j}^*) \text{ y } \underline{\theta}^* = (\theta_{i-j}^*)$$

matrices de dimensiones $\infty \times p$ y $\infty \times q$, con coeficientes definidos de la

forma

$$\phi^{-1}(B) = \sum_{k=0}^{\infty} \phi_k^* B^k \quad \text{y} \quad \theta^{-1}(B) = \sum_{k=0}^{\infty} \theta_k^* B^k$$

donde

$$\phi_k^* = \theta_k^* = 0 \quad \text{si } k < 0$$

y calculados por la técnica recursiva desarrollada por (Box-Jenkins, 1976).

Demostración.-

En los lemas 1 y 2 se ha demostrado que tanto $\hat{\beta} - \vec{\beta}$ como \vec{r} , si despreciamos los términos de orden superior a $1/n$, son en realidad medias de una serie de martingalas y haciendo uso del Teorema Central del Límite para martingalas (Billingsley, 1961) se tiene que

$$\sqrt{n}(\hat{\beta} - \vec{\beta}, \vec{r}) \xrightarrow{L} N(\vec{0}, C)$$

Cálculo de la matriz C de varianzas-covarianzas.-

En primer lugar, como pasos previos, calcularemos la covarianza asíntótica de

$$\sqrt{n} s_i = -\sqrt{n} \frac{\sum_{t=1}^n a_t v_{t-i}}{n}$$

y de

$$\sqrt{n} Q_j = \sqrt{n} \frac{\sum_{t=1}^{n-j} a_t a_{t+j}}{n}$$

Para $1 \leq i \leq p$

$1 \leq j \leq m$

$$(1.3.3.1) \quad \lim_{n \rightarrow \infty} n E(s_i Q_j) = -\lim_{n \rightarrow \infty} n^{-1} E\left(\sum_t \sum_s a_t v_{t-i} a_s a_{s+j}\right) = \gamma_{va}(j-i) = \gamma_{av}(i-j) = \phi_{i-j}^* \quad (\text{Hannan, 1970})$$

donde

$$\gamma_{av} = \frac{\gamma_{av}(k)}{\phi(B)} \quad \text{y} \quad v_t = \frac{-a_t}{\phi(B)} = -\sum_{i=0}^{\infty} \phi_i^* B^i a_t$$

siendo

$$\gamma_{av}(k) = E(a_t v_{t-k})$$

Análogamente se calculan las covarianzas asintóticas de $\sqrt{n} s_i$ y de $n Q_j$, pero ahora para

$$\sqrt{n} s_k = -\sqrt{n} \frac{\sum_{t=1}^n a_t u_{t-p-k}}{n} \quad \begin{array}{l} \text{si } p+1 \leq k \leq p+q \\ \text{y } 1 \leq j \leq m \end{array}$$

$$(1.3.3.2) \quad \lim_{n \rightarrow \infty} n E(s_i, Q_j) = -\lim_{n \rightarrow \infty} n^{-1} E\left(\sum_t \sum_s a_t u_{t-i} a_s a_{s+j}\right) = \\ = \gamma_{ua}(j-i) = \gamma_{au}(i-j) = -\Theta_{i-j}^* \quad \begin{array}{l} \text{para } i = 1 \dots q \text{ con } i = k-p \\ j = 1 \dots m \end{array}$$

donde

$$\gamma_{au} = \frac{\gamma_{au}(k)}{\Theta(B)}$$

y

$$u_t = \frac{a_t}{\Theta(B)} = \sum_{i=0}^{\infty} \Theta_i^* B^i a_t$$

siendo

$$\gamma_{au}(k) = E(a_t u_{t-k})$$

De (1.3.3.1) y (1.3.3.2) se deduce la matriz de covarianzas asintótica, entre $\sqrt{n}(\hat{\beta} - \beta)$ y $\sqrt{n} \bar{r}$

$$\text{Cov}(\sqrt{n}(\hat{\beta} - \beta), \sqrt{n} \bar{r}) = E(\sqrt{n}(\hat{\beta} - \beta) \sqrt{n} \bar{r}) = n E((\hat{\beta} - \beta) \bar{r}) = \\ = n E(\Lambda^{-1} \bar{s} \bar{r}) = \Lambda^{-1} n E(\bar{s} \bar{r}) \xrightarrow{n \rightarrow \infty} \Lambda^{-1} (-X')$$

Luego tenemos probado el lema 3. Puesto que es fácil probar que

$$V(\sqrt{n}(\hat{\beta} - \beta)) \xrightarrow{n \rightarrow \infty} \Lambda^{-1} \quad \text{y} \quad V(\sqrt{n} \bar{r}) \xrightarrow{n \rightarrow \infty} I$$

1.3.4.-TEOREMA .-

La distribución de las autocorrelaciones muestrales \hat{r} del modelo (1.1) tiende asintóticamente a una normal de vector de medias $\bar{0}$ y de matriz de varianzas-covarianzas $n^{-1}(I_m - X \Lambda^{-1} X')$

Demostración.-

Sea $\dot{\vec{\beta}}$ cualquier solución de $\vec{\beta}$ perteneciente al espacio paramétrico abierto $\mathcal{R} \subset \mathbb{C}^{p+q}$, definimos entonces las m correlaciones residuales

$$\dot{r}_k = \frac{\sum_{t=1}^{n-k} \dot{a}_t \dot{a}_{t+k}}{\sum_{t=1}^n \dot{a}_t^2} \quad \text{para } 1 \leq k \leq m$$

donde \dot{a}_t está definida según (1.3.2.3)

Si calculamos el vector de derivadas de $\dot{\vec{r}} = (\dot{r}_1, \dots, \dot{r}_m)'$ respecto de $\vec{\phi}$ y de $\vec{\Theta}$ se tiene

$$\frac{\partial \dot{r}_i}{\partial \phi_j} = n^{-1} \sum_{t=1}^{n-i} a_t v_{t+i-j} + O(1/n) = -\phi_{i-j}^* + O(1/n) \quad \text{si } \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq p \end{matrix}$$

y

$$\frac{\partial \dot{r}_i}{\partial \Theta_j} = n^{-1} \sum_{t=1}^{n-i} a_t u_{t+i-j} + O(1/n) = \Theta_{i-j}^* + O(1/n) \quad \text{si } \begin{matrix} 1 \leq i \leq m \\ p+1 \leq j \leq p+q \end{matrix}$$

Donde se han tenido en cuenta los resultados obtenidos en (1.3.3.1) y (1.3.3.2) y el resultado del Lema 1

Ahora desarrollando en serie de Taylor $\dot{\vec{r}}$, función de $\vec{\beta}$ definida en \mathcal{R} , en un entorno de $\vec{\beta} = \hat{\vec{\beta}}$ y tomando $\hat{\vec{\beta}}$ como $\hat{\vec{\beta}}$ se sigue que

$$(1.3.4.1) \quad \hat{\vec{r}} = \vec{r} + X(\hat{\vec{\beta}} - \vec{\beta}) + O(1/n)$$

y aplicando el lema 3^a queda demostrado el teorema.

Se demuestra en (McLeod 1977a) que para m suficientemente grande el teorema anterior obtiene un resultado equivalente al obtenido por (Box-Pierce 1970) acerca de la distribución $\hat{\vec{r}}$.

1.3.5.- Consecuencia del Teorema.-

Basándonos en el Teorema anterior y en el supuesto de que m sea suficientemente grande, de un orden de magnitud superior a \sqrt{n} , entonces para $i > m$ suponemos que los parámetros ϕ_i^* y θ_i^* de la matriz X definida en (1.3.3), tienden a anularse, de forma tal que puedan ignorarse los efectos finales en la matriz del diseño X , así $X'X$ tiende a Λ , puesto que en el supuesto anterior dicha matriz X es de la forma:

$$X = (-\underline{\phi}^* : \underline{\theta}^*) = \begin{pmatrix} -1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ -\phi_1^* & -1 & \dots & 0 & \theta_1^* & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\phi_{m-1}^* & \dots & \dots & -\phi_{m-p}^* & \theta_{m-1}^* & \dots & \dots & \theta_{m-q}^* \end{pmatrix} \quad m$$

p q

ó también

$$X = \begin{pmatrix} 1 & 0 & \dots & 0 \\ A_1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{m-1} & A_{m-2} & \dots & A_{m-p+q} \end{pmatrix} \quad m \times (p+q)$$

donde

$$A(z) = \sum_{i=0}^{\infty} A_i z^i = \{ \phi(z) \theta(z) \}^{-1}$$

Ahora con estos resultados es posible reenumerar el teorema en la forma :

" La distribución de las autocorrelaciones muestrales \hat{r} del modelo (1.1) tiende asintóticamente, cuando n y m son suficientemente grandes y bajo la hipótesis nula de que dicho modelo es correcto, a una distribución normal de vector de medias $\vec{0}$ y matriz de covarianzas $n^{-1} (I - X(X'X)^{-1}X')$ ".

Es fácil analizando dicha matriz comprobar que es idempotente y de rango $m-p-q$, ya que se verifica que : $\hat{r}' X = \vec{0}$

esto es

$$(1.3.5.1) \quad (\hat{r}_1, \hat{r}_2, \dots, \hat{r}_m) \cdot \begin{bmatrix} 1 & 0 & \dots & 0 \\ A_1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{m-1} & \dots & \dots & A_{m-p-q} \end{bmatrix} = \bar{\sigma}_{1 \times (p+q)}$$

1.4.- TESTS ESTADÍSTICOS BASADOS EN \hat{r} , $Q(\hat{r})$ Y $\tilde{Q}(\hat{r})$.-

$$Q(\hat{r}) = n \sum_{k=1}^m \hat{r}_k^2$$

es el estadístico "portmanteau" desarrollado por (Box-Pierce 1970) para contrastar la bondad del modelo propuesto, mientras que:

$$\tilde{Q}(\hat{r}) = n(n+2) \sum_{k=1}^m \frac{\hat{r}_k^2}{n-k}$$

es el estadístico "portmanteau" modificado, propuesto por (Ljung-Box (1978)) con idéntico propósito, de contrastar la correcta especificación del modelo (1.1).

1.4.1.- Cálculo de las medias de los estadísticos $Q(\hat{r})$ y $\tilde{Q}(\hat{r})$.-

Vamos a analizar brevemente la conveniencia de utilizar uno u otro, para ello vamos a calcular sus esperanzas matemáticas primero para muestras finitas :

$$(1.4.1.1) \quad \begin{aligned} & a) \\ E(Q(\hat{r})) &= E\left(n \sum_{k=1}^m \hat{r}_k^2\right) = E(n \hat{r}' \hat{r}) = E[n \hat{r}' \{I - X(X'X)^{-1} X'\}' \\ & \{I - X(X'X)^{-1} X'\} \hat{r}] = E[n \hat{r}' \{I - X(X'X)^{-1} X'\} \hat{r}] = \\ & = \text{traza} [n \{I - X(X'X)^{-1} X'\} D] \end{aligned}$$

donde D es la matriz de covarianzas teóricas de \hat{r}

$$D = E(\bar{r}\bar{r}') = \begin{pmatrix} \frac{n-1}{n(n+2)} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \\ & & & \frac{n-m}{n(n+2)} \end{pmatrix}$$

La igualdad (1.4.1.1) es cierta, teniendo en cuenta :

- i) La matriz $I - X(X'X)^{-1}X'$ es idempotente y simétrica .
- ii) A partir del resultado (1.3.4.1), $\hat{r} = \bar{r} + X(\hat{\beta} - \bar{\beta})$ salvo un infinitesimo de orden superior a $\frac{1}{n}$, y premultiplicando por $X(X'X)^{-1}X'$ se tiene que:

$$\begin{aligned} X(X'X)^{-1}X'(\hat{r} - \bar{r}) &= X(X'X)^{-1}X'X(\hat{\beta} - \bar{\beta}) = \\ &= X(\hat{\beta} - \bar{\beta}) = \hat{r} - \bar{r} \end{aligned}$$

luego

$$X(X'X)^{-1}X'\hat{r} - X(X'X)^{-1}X'\bar{r} = \hat{r} - \bar{r}$$

y de (1.3.5.1), se deduce que $X(X'X)^{-1}X'\hat{r} = \bar{r}$ entonces

$$\hat{r} = (I - X(X'X)^{-1}X')\bar{r} + O(1/n)$$

Teniendo en cuenta, ahora la estructura de la matriz D, y que los elementos de la matriz $X(X'X)^{-1}X'$ tienden a cero cuando i ó j aumentan, podemos hacer la aproximación de tomar $n^{-1}X(X'X)^{-1}X'$ en vez de $X(X'X)^{-1}X'$ y con esta aproximación, volviendo a la igualdad (1.4.1.1) se tiene que:

$$\begin{aligned} E(Q(\hat{r})) &\simeq \text{traza}(nD) - \text{traza}[(nX(X'X)^{-1}X')D] \simeq \\ &\simeq E(Q(\bar{r})) - \text{traza}[X(X'X)^{-1}X'] = E(Q(\bar{r})) - (p+q) \end{aligned}$$

Ahora sustituyendo el valor de $E(Q(\bar{r}))$ obtenido en (1.2.2.1) se sigue que aproximadamente

$$(1.4.1.1) \quad E(Q(\hat{\bar{r}})) = \frac{mn}{n+2} \left(1 + \frac{m+1}{2n}\right) - p - q$$

b) El cálculo de $E(\tilde{Q}(\bar{r}))$ se hace de forma análoga teniendo en cuenta la misma aproximación anterior y se deduce que

$$(1.4.1.2) \quad E(\tilde{Q}(\hat{\bar{r}})) = E(\tilde{Q}(\bar{r})) - p - q = m - p - q$$

1.4.2.-Distribuciones asintóticas de $Q(\hat{\bar{r}})$ y $\tilde{Q}(\hat{\bar{r}})$.-

a) Vamos a probar que bajo la hipótesis de que el modelo (1.1) es correcto el estadístico $Q(\hat{\bar{r}})$ se distribuye aproximadamente según una χ^2_{m-p-q} cuando m y n sean suficientemente grandes.

Basándonos en la consecuencia del teorema 1, como :

$$\hat{\bar{r}} \xrightarrow{L} N(\bar{0}, n^{-1} (I - X(X'X)^{-1}X'))$$

si llamamos a la matriz de covarianzas de $\hat{\bar{r}}$, C_* , entonces, debido a que es idempotente y de rango $m-p-q$ se tiene que

$$S^* = n \hat{\bar{r}}' \hat{C}_*^{-1} \hat{\bar{r}} \xrightarrow{L} \chi^2_{m-p-q}$$

donde

$$\hat{C}_* = I - X(X'X)^{-1}X'$$

pero para valores grandes de m de orden superior a \sqrt{n} , la matriz C_* se aproxima a la matriz identidad, luego

$$Q(\hat{\bar{r}}) = n \hat{\bar{r}}' \hat{\bar{r}}$$

se aproxima a una χ^2_{m-p-q} para m y n suficientemente grandes.

Ahora para valores pequeños de n , la distribución de $Q(\hat{\bar{r}})$ parece desviarse bastante de una χ^2_{m-p-q} puesto que $E(Q(\hat{\bar{r}})) = m-p-q$, para n

suficientemente grande, mientras que la esperanza de $Q(\hat{r})$ toma el valor aproximado dado por (1.4.1.1) cuando n es pequeño.

Las propiedades de $Q(\hat{r})$ han sido discutidas en diferentes trabajos (Davies, N., Triggs, C.M. and Newbold, P. (1977); Davies, N. and Newbold, P. (1979)) siendo su principal punto de discusión que la distribución de $Q(\hat{r})$ para muestras pequeñas no es buena aproximación a una distribución χ^2_{m-p-q} .

b) El test "portmanteau" modificado $\tilde{Q}(\hat{r})$, trata de suplir a $Q(\hat{r})$ para el caso de muestras finitas y m suficientemente grande, pues su distribución será bastante cercana a una χ^2_{m-p-q} cuando n es pequeño, ya que hemos visto en (1.4.1.2) que $E(\tilde{Q}(\hat{r})) = m-p-q$.

1.5.- TEST PARA UN m NO NECESARIAMENTE DE ORDEN SUPERIOR A \sqrt{n} .

Otro test basado también en las m primeras autocorrelaciones residuales es de la forma

$$S = n \hat{r}' C_1^{-1} \hat{r}$$

con
$$C_1 = I_m - X \Lambda^{-1} X'$$

donde m no tiene porque ser de orden superior a \sqrt{n} en este caso, si usamos el Teorema 1^o entonces

$$S \xrightarrow{L} \chi^2_m$$

En este caso bajo la hipótesis nula de que el modelo (1.1) es correcto, la matriz C_1 será regular siempre que m no sea de orden superior a \sqrt{n} .

1.6.-CONCLUSIONES.-

Tanto $Q(\hat{r})$, como $\tilde{Q}(\hat{r})$ y S son estadísticos que sirven para diagnosticar la bondad de un modelo ARMA (p,q), ya fijado mediante el preblanqueado de los residuos.

Para valores de m de orden superior a \sqrt{n} los estadísticos que se usan son $Q(\hat{r})$ y $\tilde{Q}(\hat{r})$ en el caso de que n sea suficientemente grande, en caso de que no lo sea, se usará prioritariamente $\tilde{Q}(\hat{r})$, sobre todo, si n es de tamaño moderado comparado con m , ya que la potencia de $Q(\hat{r})$ es pequeña frente a alternativas que difieran seriamente del modelo de la hipótesis nula (Ljung-Box, 1978).

En el caso de que m sea suficientemente grande ya hemos visto en la consecuencia del Teorema 1^a que entonces $X'X$ tiende a Λ y $X(X'X)^{-1}X'$ tiende a ser singular en el caso en que $p+q \neq 0$, luego la matriz C_1 definida en (1.5.) , tiende a ser singular, en cuyo caso \hat{C}^{-1} se comporta como un mal estimador de C^{-1} , es por ello que en general será preferible usar un test basado en un número reducido de autocorrelaciones residuales.

C A P I T U L O I I

METODO DE LOS SCORES EN EL CONTRASTE DE HIPOTESIS :
· APLICACION DEL TEST DE LOS MULTIPLICADORES DE
LANGRANGE PARA CONTRASTAR LA BONDAD DE MODELOS
ARMA

"METODO DE LOS SCORES EN EL CONTRASTE DE HIPOTESIS : APLICACION
DEL TEST DE LOS MULTIPLICADORES DE LAGRANGE PARA CONTRASTAR LA
BONDAD DE MODELOS ARMA"

2.0.-	Sumario	25
2.1.-	Introducción.....	26
2.1.1.-	Planteamiento del problema.....	27
2.1.2.-	Hipótesis de partida.....	28
2.2.-	Proposiciones y lemas previos para hallar la distribución asintótica del test de los multiplicadores de Lagrange	29
2.2.1.-	Proposición 1 ^a	29
2.2.2.-	Proposición 2 ^a	30
2.2.3.-	Proposición 3 ^a	31
2.2.4.-	Proposición 4 ^a	34
2.2.5.-	Proposición 5 ^a	35
2.2.6.-	Lema 1 ^a	35
2.2.7.-	Lema 2 ^a	37
2.3.-	Estimaciones restringidas.....	39
2.3.1.-	Lema 3 ^a	40
2.3.2.-	Test de los multiplicadores de Lagrange.....	42
2.3.3.-	Planteamiento del problema.....	42
2.3.4.-	Teorema.....	44
2.4.-	Test de los multiplicadores de Lagrange para matrices de información singulares.....	46
2.4.1.-	Planteamiento del problema.....	46
2.4.2.-	Distribuciones asintóticas de $\hat{\beta}_n$ y $\hat{\lambda}_n$	47
2.4.3.-	Lema previo.....	49
2.4.4.-	Teorema.....	51
2.4.5.-	Consecuencia del Teorema.....	56

2.5.- Aplicación del Test de los multiplicadores de Lagrange para contrastar la bondad de modelos ARMA.....	57
2.5.1.- Introducción y planteamiento.....	57
2.5.2.- Construcción de la regresión auxiliar.....	61
2.5.3.- Aplicación a un modelo autoregresivo.....	64
2.5.4.- Aplicación a un contraste de un modelo ARMA (p,q) frente a un ARMA (p+r,q+s).....	66
2.5.5.- Teorema.....	69
2.5.6.- Teorema.....	72
2.5.7.- Consecuencias	74
2.6.- Equivalencia asintótica del Test del estadístico portmanteau y del "score" test en un caso particular.....	75
2.6.1.- Introducción.....	75
2.6.2.- Equivalencia.....	75
2.6.3.- Comentario.....	80

2.0.- SUMARIO.

En este capítulo abordamos la teoría general del contraste de hipótesis, usando el método de los scores, desarrollada por Fisher (1946), Silvey (1959), Aitchison y Silvey (1960) y Rao (1973), donde hemos tratado de homogeneizar dicha teoría, reestructurando, en algunos casos, resultados ya conocidos y en otros dando una demostración particular de los mismos, a fin de poder aplicar dicha teoría al contraste de modelos ARMA de series temporales, extendiéndola al caso, en el que la matriz de información de los parámetros del modelo es singular.

Continúa el capítulo, con aplicaciones del test para modelos ARMA, AR y ARMA donde se aumentan a la vez los parámetros de la parte autoregresiva y de medias móviles del modelo.

Finalizamos el capítulo con un estudio de la equivalencia asintótica del test basado en el estadístico portmanteau, y del score test para un caso particular de modelos alternativos ARMA $(p+m, q)$ ó ARMA $(p, q+m)$, donde se ha tenido en cuenta, el trabajo desarrollado por Newbold (1980).

METODO DE LOS "SCORES" EN EL CONTRASTE DE HIPOTESIS.

2.1.- INTRODUCCION.

Sea \vec{x} una variable aleatoria que $\in \mathbb{R}^\alpha$ con función de distribución $F(\vec{x}, \vec{\beta})$ donde $\vec{\beta}$ es un vector de parámetros

$$\vec{\beta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) \in \mathcal{R} \subset \mathbb{C}^{p'+q'}$$

Si tomamos una muestra aleatoria simple de tamaño n , entonces el log de la función de verosimilitud vendrá dado como una función de $\vec{\beta}$ para cada muestra \vec{x}_n

$$\log L_n(\vec{x}_n, \vec{\beta}) = \sum_{i=1}^n \log f(\vec{x}_i, \vec{\beta})$$

Trataremos la contrastación de la hipótesis simple

$$H_0 \equiv \{\vec{\beta} = \vec{\beta}^0\} = \{\phi_1 = \phi_1^0, \dots, \phi_p = \phi_p^0, \theta_1 = \theta_1^0, \dots, \theta_q = \theta_q^0\}$$

frente a la alternativa

$$H_1 \equiv \{\vec{\beta} = \vec{\beta}^1\} = \{\phi_1 = \phi_1^0 + L_1, \dots, \phi_p = \phi_p^0 + L_p, \theta_1 = \theta_1^0 + L_{p+1}, \dots, \theta_q = \theta_q^0 + L_{p+q}\}$$

o mejor, contrastar la hipótesis nula compuesta de que los conjuntos admisibles de soluciones para $\vec{\beta}$ pertenezcan a una intersección de superficies

$$H_0 \equiv \{\bar{h}(\vec{\beta}) = \vec{0}\} = \{h_j(\vec{\beta}) = 0 \text{ para } j=1, \dots, r\}$$

siendo \bar{h} una función bien definida de $\mathbb{C}^{p'+q'}$ en \mathbb{R}^r con $r < p'+q'$

Así $\vec{\beta}^0$, será el valor a contrastar de los parámetros,

donde

$$\vec{\beta}^0 \in \mathcal{R} \cap \text{Ker } \bar{h}$$

2.1.1.- Plantèamiento del problema.

Sea : $\vec{\beta}_n(\vec{x}_n, \tau^*)$
 la estimación máximo verosimil de $\vec{\beta}$ en algún subconjunto τ^* de \mathcal{R}

si $\hat{\beta}_n(\cdot, \tau^*)$, tiene un único valor definido para casi todo \vec{x}_n , entonces $\hat{\beta}_n(\cdot, \tau^*)$ es una v.a. llamada estimador máximo verosimil de $\vec{\beta}^0$ en τ^* , donde $\vec{\beta}^0$ es el valor verdadero de $\vec{\beta}$.

Supuesto que los parámetros $\vec{\beta}$ están sujetos a r restricciones independientes, $\vec{h}(\vec{\beta}) = \vec{0}$, habrá que probar que el estimador máximo verosimil $\hat{\beta}_n(\vec{x}_n, \tau)$, al que llamaremos $\hat{\beta}_n$ por brevedad, verifica las ecuaciones restringidas de verosimilitud (E.R.V)

$$(2.1.1.1) \begin{cases} n^{-1} D \log L_n(\vec{x}_n, \vec{\beta}) + H_{\vec{\beta}} \vec{\lambda} = \vec{0} \\ \vec{h}(\vec{\beta}) = \vec{0} \end{cases}$$

donde $\vec{\lambda}$ es el multiplicador lagrangiano en \mathbb{R}^r ,

$$(2.1.1.2) D \log L_n(\vec{x}_n, \vec{\beta}) = \left(\frac{\partial \log L}{\partial \beta_i} \right)'$$

es un vector $(p'+q') \times 1$.

$$H_{\vec{\beta}} = (h_{ij})_{(p'+q') \times r}$$

es una matriz, donde

$$h_{ij} = \frac{\partial h_j(\vec{\beta})}{\partial \beta_i}$$

para $i = 1, \dots, p'+q'$
 $j = 1, \dots, r$

Como el conjunto de ecuaciones (2.1.1.1) tiene $(p'+q'-r)$

restricciones lineales en

$$\left. \frac{\int \log L}{\int \bar{\beta}} \right|_{\bar{\beta} = \hat{\bar{\beta}}_n}$$

se espera que el estadístico

$$S_n = \sum_i \sum_j \alpha^{ij} (\hat{\bar{\beta}}_n) \frac{\int \log L(\bar{x}, \hat{\bar{\beta}}_n)}{\int \beta_i} \frac{\int \log L(\bar{x}, \hat{\bar{\beta}}_n)}{\int \beta_j}$$

se distribuya asintóticamente según una χ^2_r , donde el número de grados de libertad es $p'+q'-r$ unidades menor que para los verdaderos valores de $\bar{\beta}$.

Las derivadas $\frac{\int \log L}{\int \beta_i}$, $i = 1, \dots, p'+q'$ son los eficientes "scores" definidos en (Fisher 1946) y donde el vector definido en (2.1.1.2)

$$D \log L_n(\bar{x}_n, \bar{\beta})$$

es una variable aleatoria con vector de medias nulo y matriz de covarianzas $J = (\alpha_{ij})$ siendo $J^{-1} = (\alpha^{ij})$ su inversa.

Para demostrar que el estadístico $S_n \xrightarrow{L} \chi^2_r$, hacen falta unas hipótesis de partida.

También demostraremos en (2.2.6) que bajo estas hipótesis, la distribución asintótica de la variable $n^{-1} D \log L_n(x_n, \bar{\beta})$ es una $N(0, J)$ y entonces

$$n^{-1} \sum_{i=1}^{p'+q'} \sum_{j=1}^{p'+q'} \alpha^{ij} \frac{\int \log L}{\int \beta_i} \frac{\int \log L}{\int \beta_j} \Bigg|_{\bar{\beta} = \bar{\beta}^0} \xrightarrow{L} \chi^2_{p'+q'}$$

2.1.2.-Hipótesis de partida.-

1ª) Las derivadas

$$\frac{\int \log f(\bar{x}, \bar{\beta})}{\int \beta_i}, \quad \frac{\int^2 \log f(\bar{x}, \bar{\beta})}{\int \beta_i \int \beta_j} \quad y \quad \frac{\int^3 \log f(\bar{x}, \bar{\beta})}{\int \beta_i \int \beta_j \int \beta_k}$$

existen para casi todo $\vec{x} \in R^\alpha$ y para cada $\vec{\beta} \in \mathcal{R}$, siendo \mathcal{R} un conjunto abierto, convexo y compacto en \mathbb{C}^{p+q} que incluye el verdadero valor del parámetro $\vec{\beta}^0$.

2ª) Para $\vec{\beta} = \vec{\beta}^0$ se tiene que

$$i) \quad E \left(\frac{1}{f(\vec{x}, \vec{\beta})} \frac{\int f(\vec{x}, \vec{\beta})}{\int \beta_i} \right) = E \left(\frac{\int \text{Log } f(\vec{x}, \vec{\beta})}{\int \beta_i} \right) = 0$$

$$ii) \quad E \left(\frac{1}{f(\vec{x}, \vec{\beta})} \frac{\int^2 f(\vec{x}, \vec{\beta})}{\int \beta_i \int \beta_j} \right) = 0 \quad \forall i, j$$

$$iii) \quad E \left\{ \left(\frac{\int \text{Log } f(\vec{x}, \vec{\beta})}{\int \beta_i} \right)^2 \right\} > 0$$

$$iv) \quad \left| \frac{\int^3 \text{Log } f(\vec{x}, \vec{\beta})}{\int \beta_i \int \beta_j \int \beta_k} \right| < M(\vec{x})$$

donde

$E (M(\vec{x}) / \vec{\beta} = \vec{\beta}^0) < k$, siendo k una constante independiente del valor de $\vec{\beta}$.

3ª) Las funciones h_1, \dots, h_r , tales que $h_i(\vec{\beta}) = 0$, son funciones que admiten derivadas parciales de primer y segundo orden acotadas en \mathcal{R} .

2.2.- PROPOSICIONES Y LEMAS PREVIOS.-

2.2.1.-Proposición 1ª.-

Sea $\vec{\beta}^0$ el verdadero valor del parámetro y consideremos dos valores $\vec{\beta}^0 \pm \vec{J}$, donde \vec{J} es un vector de incrementos positivos. Las funciones $f(\vec{x}, \vec{\beta}^0)$ y $f(\vec{x}, \vec{\beta}^0 - \vec{J})$ son no negativas e integrables con respecto a una medida μ definida por $dF(\vec{x}, \vec{\beta})$.

Sea S^α la región donde f es mayor que cero, como

$$\int_{S^\alpha} \{f(\vec{x}, \vec{\beta}^0 - \vec{J}) - f(\vec{x}, \vec{\beta}^0)\} dF(\vec{x}, \vec{\beta}) \leq 0$$

Entonces se verifica que

$$\int_{S^\alpha} f(\bar{x}, \vec{\beta}^0 - \vec{J}) \log \frac{f(\bar{x}, \vec{\beta}^0 - \vec{J})}{f(\bar{x}, \vec{\beta}^0)} dF \leq 0$$

Demostración.-

Aplicando la desigualdad de Jensen, al ser f convexa en un entorno de $\vec{\beta}^0$, se tiene (RAO, 1973)

$$f(\bar{x}, \vec{\beta}^0) \geq f(\bar{x}, \vec{\beta}^0 - \vec{J}) + \left. \frac{Jf}{J\vec{\beta}} \right|_{\vec{\beta} = \vec{\beta}^0 - \vec{J}} \vec{J}$$

Al ser f no negativa, tomamos logaritmos y tenemos

$$\log f(\bar{x}, \vec{\beta}^0) \geq \log \left\{ f(\bar{x}, \vec{\beta}^0 - \vec{J}) + \left. \frac{Jf}{J\vec{\beta}} \right|_{\vec{\beta} = \vec{\beta}^0 - \vec{J}} \vec{J} \right\} \geq \log f(\bar{x}, \vec{\beta}^0 - \vec{J})$$

ya que

$$\left. \frac{Jf}{J\vec{\beta}} \right|_{\vec{\beta} = \vec{\beta}^0 - \vec{J}} \vec{J} \geq 0$$

Multiplicando ahora ambos miembros de la inecuación por $f(\bar{x}, \vec{\beta}^0 - \vec{J})$

se obtiene

$$f(\bar{x}, \vec{\beta}^0 - \vec{J}) \log f(\bar{x}, \vec{\beta}^0) \geq f(\bar{x}, \vec{\beta}^0 - \vec{J}) \log f(\bar{x}, \vec{\beta}^0 - \vec{J})$$

luego

$$f(\bar{x}, \vec{\beta}^0 - \vec{J}) \log \frac{f(\bar{x}, \vec{\beta}^0 - \vec{J})}{f(\bar{x}, \vec{\beta}^0)} \leq 0$$

e integrando respecto a μ , se tiene

$$\int_{S^\alpha} f(\bar{x}, \vec{\beta}^0 - \vec{J}) \log \frac{f(\bar{x}, \vec{\beta}^0 - \vec{J})}{f(\bar{x}, \vec{\beta}^0)} dF \leq 0 \quad \text{como queríamos demostrar}$$

2.2.2-Proposición 2^a.-

Por hipótesis la función $\log f(\bar{x}, \vec{\beta}^0)$ es diferenciable en \mathcal{R} y se deduce que la ecuación de máxima verosimilitud

$$\log L_n(\bar{x}_n, \hat{\vec{\beta}}_n) = \sup_{\vec{\beta} \in \mathcal{C}} \log L_n(\bar{x}_n, \vec{\beta})$$

tiene una raíz $\hat{\vec{\beta}}_n$ con probabilidad 1, y es $\hat{\vec{\beta}}_n$, además, consistente hacia $\vec{\beta}^0$

Demostración.-

Por la proposición 1.- se deduce que

$$E \left(\log \frac{f(\vec{x}, \vec{\beta}^0 - \vec{J})}{f(\vec{x}, \vec{\beta}^0)} \right) < 0$$

análogamente se puede demostrar que

$$E \left(\log \frac{f(\vec{x}, \vec{\beta}^0 + \vec{J})}{f(\vec{x}, \vec{\beta}^0)} \right) < 0$$

donde las desigualdades estrictas surgen del hecho de suponer que $f(\vec{x}, \vec{\beta})$ es independiente de $\vec{\beta}$ en un conjunto que contenga a $\vec{\beta}^0$.

A partir de ambas inecuaciones y aplicando la ley fuerte de los grandes números, se tiene que

$$\frac{1}{n} \left\{ \log L_n(\vec{x}_n, \vec{\beta}^0 \pm \vec{J}) - \log L_n(\vec{x}_n, \vec{\beta}^0) \right\} < 0$$

con probabilidad 1.. Lo que significa que para casi todo \vec{x}_n , la sucesión $\{\log L_n(\vec{x}_n, \vec{\beta})\}$ será mayor en $\vec{\beta}^0$ que en $\vec{\beta}^0 + \vec{J}$ ó $\vec{\beta}^0 - \vec{J}$

Como la función $\log L_n(\vec{x}_n, \vec{\beta})$ es continua en $(\vec{\beta}^0 \pm \vec{J})$ entonces existe un máximo local de $\log L_n(\vec{x}_n, \vec{\beta})$ en el intervalo $(\vec{\beta}^0 - \vec{J}, \vec{\beta}^0 + \vec{J})$, y como $\log L_n(\vec{x}_n, \vec{\beta})$ es diferenciable, su derivada respecto de $\vec{\beta}$, se anula en ese máximo local. al cual llamaremos $\hat{\vec{\beta}}_n$.

Además como el radio \vec{J} del entorno es arbitrario, la raíz $\hat{\vec{\beta}}_n$ será consistente hacia $\vec{\beta}^0$.

2.2.3.-Proposición 3^a.-

Dadas las hipótesis de partida se verifica que

$$\sqrt{n} \left[\left(\hat{\beta}_i - \beta_i^0 \right) \left\{ -E \left(\frac{\partial^2 \log f(\vec{x}, \vec{\beta})}{\partial \beta_i \partial \beta_j} \right) \Big|_{\vec{\beta} = \vec{\beta}^0} \right\} - \frac{1}{n} \frac{\partial \log L_n}{\partial \beta_i} \Big|_{\vec{\beta} = \vec{\beta}^0} \right]$$

$$\forall i=1, \dots, p+q'$$

converge a 0 con probabilidad 1.

Demostración.-

Como $\hat{\beta}$ satisface la ecuación de verosimilitud se verifica que

$$\left. \frac{\partial \log L}{\partial \beta_i} \right|_{\hat{\beta} = \hat{\beta}} = 0 \quad \forall i=1, \dots, p+q'$$

Si desarrollamos cada componente de la función vectorial

$$\frac{d \log L}{d \hat{\beta}}$$

en serie de Taylor, en un entorno de β^0 se tiene

$$(2.2.3.1) \quad 0 = \left. \frac{\partial \log L}{\partial \hat{\beta}_i} \right|_{\hat{\beta} = \hat{\beta}} + (\hat{\beta}_i - \beta_i^0) \left. \frac{\partial^2 \log L}{\partial \hat{\beta}_i \partial \hat{\beta}_j} \right|_{\hat{\beta} = \hat{\beta}} + \frac{\partial^3 \log L}{\partial \beta_i \partial \beta_j \partial \beta_k} \left. \right|_{\hat{\beta} = \hat{\beta}} \cdot \frac{(\hat{\beta}_i - \beta_i^0)^2}{2}$$

donde $\beta_i^1 \in (\beta_i^0, \hat{\beta}_i), \forall i$

Ahora despejando $\hat{\beta}_i - \beta_i^0$ se deduce que

$$\sqrt{n} (\hat{\beta}_i - \beta_i^0) = \frac{-\frac{1}{\sqrt{n}} \frac{\partial \log L}{\partial \beta_i^0}}{\frac{1}{n} \left(\frac{\partial^2 \log L}{\partial \beta_i^0 \partial \beta_j^0} + \frac{\hat{\beta}_i - \beta_i^0}{2} \frac{\partial^3 \log L}{\partial \beta_i^0 \partial \beta_j^0 \partial \beta_k^0} \right)}$$

$\forall i, j, k$

Si consideramos la igualdad

$$(2.2.3.2) \quad \sqrt{n} J_{\hat{\beta}^0} (\hat{\beta} - \hat{\beta}^0) - \frac{1}{\sqrt{n}} \frac{d \log L}{d \hat{\beta}^0} = -(B_n + I) \frac{1}{\sqrt{n}} \frac{d \log L}{d \hat{\beta}^0}$$

donde

$$B_n = \frac{J_{\hat{\beta}^0}}{\frac{1}{n} \frac{\partial^2 \log L}{\partial \beta_i^0 \partial \beta_j^0} + J_p}$$

Si se demuestra que $B_n \xrightarrow{p} -I$ con probabilidad 1, la convergencia de $1/\sqrt{n} \left(\frac{d \log L}{d \hat{\beta}^0} \right)$ a una distribución asintótica implica que el lado derecho de la igualdad (2.2.3.2) converge a 0 con probabilidad 1, con lo que se establecería el resultado deseado.

Bajo la hipótesis 2^a anterior:

$$E\left(\frac{d^2 \log f}{d\vec{\beta}^2}\right)_{\vec{\beta}=\vec{\beta}^0} = E\left(-\frac{1}{f^2} \left(\frac{df}{d\vec{\beta}^2}\right)^2 + \frac{1}{f} \frac{d^2 f}{d\vec{\beta}^2}\right) = E\left(-\frac{1}{f^2} \left(\frac{df}{d\vec{\beta}^2}\right)^2\right)$$

puesto que

$$E\left(\frac{1}{f} \frac{d^2 f}{d\vec{\beta}^2}\right) = 0 \text{ p' + q'}$$

por hipótesis

luego

$$E\left(\frac{d^2 \log f}{d\vec{\beta}^2}\right)_{\vec{\beta}=\vec{\beta}^0} = -E\left(\frac{J \log f}{J \beta_i} \frac{J \log f}{J \beta_j}\right)_{\vec{\beta}=\vec{\beta}^0} = -J_{\vec{\beta}=\vec{\beta}^0}$$

Ahora si aplicamos la ley fuerte de los grandes números se tiene que la sucesión matricial

$$(2.2.3.3) \quad \left\{ \frac{1}{n} \frac{d^2 \log L}{d\vec{\beta}^2} \right\} \longrightarrow -J_{\vec{\beta}^0}$$

casi seguro (ó con probabilidad 1).

Ahora bajo la hipótesis 3^a, $\forall i, j, k$

$$-\frac{1}{n} \left(\frac{\int^3 \log L}{\int \beta_i \int \beta_j \int \beta_k} \right) < K \quad \text{con probabilidad 1}$$

cuando $n \rightarrow \infty$, y en la Proposición 2., vimos que $\hat{\beta}_n$ es consistente hacia $\vec{\beta}^0$, luego

$$(2.2.3.4) \quad \left(\hat{\beta}_i - \beta_i^0 \right) \frac{1}{n} \left(\frac{\int^3 \log L}{\int \beta_i^1 \int \beta_j^1 \int \beta_k^1} \right) \longrightarrow 0$$

con probabilidad 1
(ó casi seguro)

$\forall i, j, k$

Y así la matriz $\mathcal{H} \longrightarrow \underline{0}$ casi seguro.

Ahora usando teoremas de convergencia usuales (RAO, 1973) y combinando los resultados (2.2.3.3) y (2.2.3.4) se tiene que $B_n \xrightarrow{P} -I$ con probabilidad 1, luego queda probada la proposición.

2.2.4.- Proposición 4^a.-

Si $\hat{\beta}_n^1$ y $\hat{\beta}_n^2$ son dos raíces consistentes de la ecuación de verosimilitud, se verifica que

$$\sqrt{n}(\hat{\beta}_n^1 - \hat{\beta}_n^2) \xrightarrow{P} \bar{0} \quad \text{con probabilidad 1}$$

Demostración.-

Si tenemos en cuenta el resultado de la proposición 3^a en donde

$$\sqrt{n} \left(J(\hat{\beta}_n - \bar{\beta}) - \frac{1}{n} \frac{d \log L}{d \bar{\beta}} \right) \xrightarrow{P} \bar{0} \quad \text{con probabilidad 1, entonces}$$

en particular

$$\sqrt{n} \left(J(\hat{\beta}_n^1 - \bar{\beta}) - \frac{1}{n} \frac{d \log L}{d \bar{\beta}} \right) \xrightarrow{P} \bar{0} \quad \text{con probabilidad 1,}$$

$$y \quad \sqrt{n} \left(J(\hat{\beta}_n^2 - \bar{\beta}) - \frac{1}{n} \frac{d \log L}{d \bar{\beta}} \right) \xrightarrow{P} \bar{0} \quad \text{con probabilidad 1}$$

luego

$$\sqrt{n} J(\hat{\beta}_n^1 - \hat{\beta}_n^2) \xrightarrow{P} \bar{0} \quad \text{con probabilidad 1}$$

y de aquí

$$\sqrt{n}(\hat{\beta}_n^1 - \hat{\beta}_n^2) \xrightarrow{P} \bar{0} \quad \text{con probabilidad 1}$$

al considerar J como una matriz regular, con lo que queda demostrada la proposición.

2.2.5.- Proposición 5ª.-

Sea x_1, x_2, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas con función de densidad $f(x, \vec{\beta})$.

Sea

$$\int \frac{d f(x, \vec{\beta})}{d \vec{\beta}} dx = \vec{0} \quad \text{e} \quad \int \frac{1}{f(x, \vec{\beta})} \left(\frac{d f(x, \vec{\beta})}{d \vec{\beta}} \right)^2 dx = J(\vec{\beta})$$

donde $J(\vec{\beta})$ es una matriz definida positiva. Se tiene entonces que si $\hat{\beta}_n$ es un estimador consistente de $\vec{\beta}$,

$$\sqrt{n} (\hat{\beta}_n - \vec{\beta}) \xrightarrow{L} N(\vec{0}, J(\vec{\beta}))$$

Demostración.-

Por la proposición 3ª, se tiene que la distribución asintótica de $\sqrt{n} J(\hat{\beta}_n - \vec{\beta})$ es la misma que la de $\frac{1}{n} \frac{d \log l}{d \vec{\beta}}$ y por el teorema central del límite esta última sucesión de variables aleatorias converge en ley a una $N(\vec{0}, J(\vec{\beta}))$ luego sí

$$\sqrt{n} J(\hat{\beta}_n - \vec{\beta}) \longrightarrow N(\vec{0}, J(\vec{\beta}))$$

entonces

$$\sqrt{n} (\hat{\beta}_n - \vec{\beta}) \longrightarrow N(\vec{0}, J^{-1}(\vec{\beta}))$$

Basándose en estas cinco proposiciones se enuncian los siguientes lemas.

2.2.6.- LEMA 1.-

Sea $f(\vec{x}, \vec{\beta})$ la función de densidad de una variable aleatoria que satisface las hipótesis de partida iniciales y sea $\hat{\beta}$ una raíz consistente de la ecuación de verosimilitud basada en n observaciones con

$$\vec{d}' = (\sqrt{n} (\hat{\beta}_1 - \beta_1), \dots, \sqrt{n} (\hat{\beta}_{p+q} - \beta_{p+q}))$$

Si J es la matriz de información no singular de $\vec{\beta}$, entonces se tienen las siguientes equivalencias asintóticas

$$\vec{v} \xrightarrow{L} J \vec{d} \quad \text{ó} \quad \vec{d} \xrightarrow{L} J^{-1} \vec{v}$$

donde

$$\vec{v}' = (v_1(\vec{\beta}), \dots, v_{p'+q'}(\vec{\beta}))$$

siendo

$$v_i(\vec{\beta}) = \frac{1}{\sqrt{n}} \frac{\int \text{Log } L(\vec{x}, \vec{\beta})}{\int \beta_i} \quad \text{para } i = 1, \dots, p'+q'$$

Demostración.-

Previamente es preciso demostrar que la distribución asintótica de \vec{v} es una $N(\vec{0}, J)$.

Para ello basta con probar que la distribución asintótica de una función lineal de las v_i es una normal, esto es:

$$\vec{b}' \vec{v} = \sum_{i=1}^{p'+q'} b_i v_i(\vec{\beta}) \xrightarrow{L} N(0, \vec{b}' J \vec{b})$$

Demostración previa.-

Sean las variables

$$y_i = \sum_{j=1}^{p'+q'} b_j \frac{\int \text{Log } f(x_i, \vec{\beta})}{\int \beta_j}$$

para $i = 1, \dots, n$

teniendo en cuenta la hipótesis 2(i), se tiene que $E(y_i) = 0$

además

$$\begin{aligned} V(y_i) &= \sum_K \sum_L b_K b_L E\left(\frac{\int \text{Log } f(x_i, \vec{\beta})}{\int \beta_K} \frac{\int \text{Log } f(x_i, \vec{\beta})}{\int \beta_L}\right) = \\ &= \sum_K \sum_L b_K b_L \alpha_{KL} = \vec{b}' J \vec{b} \end{aligned}$$

(Contando con las hipótesis 2 y el resultado (2.2.33)).

Así las variables y_i son independientes e idénticamente distribuidas, con media nula y varianza $\bar{b}' J \bar{b}$, luego por el teorema central del límite

$$\frac{\sum_{i=1}^n y_i}{\sqrt{n}} \xrightarrow{L} N(0, \bar{b}' J \bar{b})$$

Pero como

$$\begin{aligned} n^{-\frac{1}{2}} \sum_{i=1}^n y_i &= \sum_{i=1}^n \left(\sum_{j=1}^{p+q'} b_j \frac{1}{\sqrt{n}} \frac{\partial \text{Log } f(x_i, \bar{\beta})}{\partial \beta_j} \right) = \\ &= \sum_{j=1}^{p+q'} b_j \frac{1}{\sqrt{n}} \frac{\partial \sum_{i=1}^n \text{Log } f(x_i, \bar{\beta})}{\partial \beta_j} = \sum_{j=1}^{p+q'} b_j v_j(\bar{\beta}) = \bar{b}' \bar{v} \end{aligned}$$

luego si

$$\bar{b}' \bar{v} \xrightarrow{L} N(0, \bar{b}' J \bar{b})$$

entonces

$$(2.2.6.1) \quad \bar{v} \xrightarrow{L} N(\bar{\sigma}, J)$$

como previamente queriamos probar.

Estamos ya en disposición de probar el lema, ya que de las proposiciones 3^a y 5^a se deduce que, para $\alpha(\bar{\beta}) = 1$

$$\bar{d} = \sqrt{n} (\hat{\beta} - \bar{\beta}) \xrightarrow{L} N(\bar{\sigma}, J^{-1})$$

siempre que $\hat{\beta}$ sea una raíz consistente, de primer orden, de $\bar{\beta}$.

de lo anterior se deduce que

$$J \bar{d} \xrightarrow{L} N(\bar{\sigma}, J' J^{-1} J)$$

con lo que queda probado.

2.27.- LEMA 2.-

La distribución asintótica de

$$-2 (\text{Log } L_n(\bar{x}, \hat{\beta}) - \text{Log } L_n(\bar{x}, \bar{\beta}))$$

es igual a la de

$$\bar{d}' J_{\bar{\beta}_0} \bar{d}$$

o lo que es lo mismo

$$-2 \log \frac{L_n(\bar{x}, \hat{\beta})}{L_n(\bar{x}, \dot{\beta})} = -2 \log \frac{\sup_{\beta \in \mathcal{C}} L_n(\bar{x}, \beta)}{\sup_{\beta \in \mathcal{R}} L_n(\bar{x}, \beta)}$$

tiene una distribución asintóticamente equivalente a la de

$$n(\hat{\beta} - \dot{\beta})' J_{\dot{\beta}} (\hat{\beta} - \dot{\beta})$$

donde $\dot{\beta}$ es el estimador no restringido de máxima verosimilitud.

Demostración.-

Por la proposición 4^a sabemos que como $\hat{\beta}$ y $\dot{\beta}$ son raíces consistentes de la ecuación de verosimilitud, entonces

$$\sqrt{n} (\hat{\beta} - \dot{\beta}) \xrightarrow{\text{c.s.}} \vec{0} \quad (\text{con probabilidad } 1)$$

o lo que es igual

$$\|\hat{\beta} - \dot{\beta}\| = o_p(n^{-1/2})$$

donde

$$o_p(n^{-1/2})$$

indica el orden en probabilidad de la convergencia de $\hat{\beta}$ a $\dot{\beta}$.

Si desarrollamos en serie de Taylor $\log L(\bar{x}, \hat{\beta})$ en un entorno de $\hat{\beta} = \dot{\beta}$, entonces por definición de $\dot{\beta}$ se verifica que

$$\left. \frac{d \log L(\bar{x}, \hat{\beta})}{d \hat{\beta}} \right|_{\hat{\beta} = \dot{\beta}} = \vec{0}$$

si además tenemos en cuenta la hipótesis 2^a (iv), entonces

$$(2.2.7.1) \log L(\bar{x}, \hat{\beta}) = \log L(\bar{x}, \dot{\beta}) + \frac{1}{2} (\hat{\beta} - \dot{\beta})' \left(\frac{n d^2 \log L(\bar{x}, \hat{\beta})}{n d \hat{\beta}^2} \right) \Big|_{\hat{\beta} = \dot{\beta}} (\hat{\beta} - \dot{\beta}) + o_p(1)$$

Ahora si desarrollamos en serie de Taylor la función

$$n^{-1} \frac{d^2 \log L(\bar{x}, \hat{\beta})}{d \hat{\beta}^2}$$

en un entorno de $\vec{\beta} = \vec{\beta}^0$, y tenemos en cuenta el resultado (2.2.3.3) y la hipótesis 2^a(ii)

$$n^{-1} \frac{d^2 \text{Log } L(\vec{x}, \vec{\beta})}{d \vec{\beta}^2} = n^{-1} \left. \frac{d^2 \text{Log } L(\vec{x}, \vec{\beta})}{d \vec{\beta}^2} \right|_{\vec{\beta} = \vec{\beta}^0} + O_p(1) = -J_{\vec{\beta}^0} + O_p(1)$$

luego si definimos

$$\mu = \frac{\sup_{\beta \in \mathcal{Z}} L(\vec{x}, \vec{\beta})}{\sup_{\beta \in \mathcal{R}} L(\vec{x}, \vec{\beta})} = \frac{L(\vec{x}, \hat{\vec{\beta}})}{L(\vec{x}, \vec{\beta}^0)}$$

y tomamos logaritmos, se tiene que

$$(2.2.7.2) \text{Log } \mu = -\frac{1}{2} n (\hat{\vec{\beta}} - \vec{\beta}^0)' J_{\vec{\beta}^0} (\hat{\vec{\beta}} - \vec{\beta}^0) + O_p(1), \text{ luego sus distribuciones serán asintóticamente equivalentes, como queríamos demostrar.}$$

2.3.- ESTIMACIONES RESTRINGIDAS.-

Sea $f(\vec{x}_n, \vec{\beta})$ la función de densidad de probabilidad, donde \vec{x}_n representa la variable aleatoria n dimensional de la muestra aleatoria simple, y $\vec{\beta}$ es el vector de los $p'+q'$ parámetros desconocidos del modelo. Sea $L(\vec{\beta}, \vec{x})$ la función de verosimilitud.

Hemos definido los eficientes "scores" de Fisher

$$\frac{\partial \text{Log } L(\vec{\beta}, \vec{x})}{\partial \beta_i} \quad i = 1, \dots, p'+q'$$

y hemos visto que la matriz de información de $\vec{\beta}$, se define mediante

$$J_{\vec{\beta}} = E \left(\frac{d \text{Log } L}{d \vec{\beta}} \left(\frac{d \text{Log } L}{d \vec{\beta}} \right)' \right)$$

Las estimaciones máximo verosímiles de $\vec{\beta}$ se obtienen usualmente a partir de la ecuación

$$\frac{d \text{Log } L}{d \vec{\beta}} = \vec{0}$$

y la teoria asintótica de la estimación es conocida cuando J es no singular en un entorno del verdadero valor de $\vec{\beta}$, esto es en $\vec{\beta}^0$.

El lema 1 anterior se ha desarrollado teniendo en cuenta que J es no singular.

Abordamos ahora el problema inicial, que consiste en contrastar la hipótesis de que $\vec{h}(\vec{\beta}) = \vec{0}$, esto es equivalente a suponer que los $p'+q'$ parámetros del modelo $\beta_1, \dots, \beta_{p'+q'}$ están sujetos a $r < p'+q'$ restricciones independientes $h_j(\vec{\beta}) = 0$, para $j = 1, \dots, r$, donde las funciones h_j cumplen las hipótesis 3^a de partida.

Estas restricciones son equivalentes a especificar cada β_i como una función de $s = p'+q'-r$ nuevos parámetros $\delta_1, \dots, \delta_s$ entonces J será singular y no todos los parámetros $\beta_1, \dots, \beta_{(p'+q')}$ se podrán estimar

En este caso se puede definir una función $\log L(\vec{\delta}, \vec{x})$ función de los s parámetros, de forma tal que la matriz de información para $\vec{\delta}$ sea no singular en un entorno del verdadero valor de los parámetros.

Así definimos

$$\beta_i = g_i(\delta_1, \dots, \delta_s) \quad \text{para } i = 1, \dots, p'+q'$$

donde las funciones g_i admiten derivadas parciales continuas de primer y segundo orden.

2.3.1.-LEMA 3.-

Sea $\hat{\delta}$ el estimador de máxima verosimilitud de $\vec{\delta}$ y sea

$$\vec{c}' = (\sqrt{n}(\hat{\delta}_1 - \delta_1), \dots, \sqrt{n}(\hat{\delta}_s - \delta_s))$$

Si N es la matriz de información de $\vec{\delta}$, se obtienen las siguientes equivalencias asintóticas

i) $\vec{u} \longrightarrow N \vec{c}$

ii) $2 (\log L(\vec{x}, \hat{\delta}) - \log L(\vec{x}, \vec{\delta})) \longrightarrow \vec{c}' N \vec{c}$

donde

$$\vec{u}' = (u_1(\vec{\delta}), \dots, u_s(\vec{\delta}))$$

siendo

$$u_j(\vec{\delta}) = \frac{1}{\sqrt{n}} \frac{\int \log L(\vec{x}, \vec{\delta})}{\int \delta_j} \quad \text{para } j = 1, \dots, s.$$

Estos resultados son similares a los de los lemas 1 y 2, pero además se verifica que

$$\text{iii) } \vec{u} = M' \vec{v} \quad \text{donde } M = \begin{pmatrix} \int g_i \\ \int \delta_j \end{pmatrix} \quad \begin{matrix} i = 1, \dots, p'+q' \\ j = 1, \dots, s \end{matrix}$$

y

$$\text{iv) } N = M' J M$$

Demostración.-

Las equivalencias asintóticas i), e ii) se demuestran de forma análoga a lo visto en los lemas 1 y 2.

El resultado iii), se obtiene al considerar las derivadas parciales

$$\frac{1}{\sqrt{n}} \frac{\int \log L}{\int \delta_j}$$

y tener en cuenta que

$$\beta_i = g_i(\delta_1, \dots, \delta_s)$$

pues

$$\frac{1}{\sqrt{n}} \frac{\int \log L}{\int \delta_j} = \sum_{i=1}^{p'+q'} \frac{1}{\sqrt{n}} \frac{\int \log L}{\int \beta_i} \frac{\int \beta_i}{\int \delta_j} = \sum_{i=1}^{p'+q'} \frac{1}{\sqrt{n}} \frac{\int \log L}{\int \beta_i} \frac{\int g_i}{\int \delta_j}$$

De iguales consideraciones se obtiene el resultado iv), puesto que

$$\begin{aligned} E \left(\frac{\int \log L}{\int \delta_i} \frac{\int \log L}{\int \delta_j} \right) &= E \left(\sum_{k=1}^{p'+q'} \sum_{l=1}^{p'+q'} \frac{\int \log L}{\int \beta_k} \frac{\int g_k}{\int \delta_i} \frac{\int \log L}{\int \beta_l} \frac{\int g_l}{\int \delta_j} \right) = \\ &= \sum_k \sum_l \frac{\int g_k}{\int \delta_i} \frac{\int g_l}{\int \delta_j} E \left(\frac{\int \log L}{\int \beta_k} \frac{\int \log L}{\int \beta_l} \right) \quad \text{Luego} \\ E \left(\frac{\int \log L}{\int \delta_i} \frac{\int \log L}{\int \delta_j} \right) &= M' E \left(\frac{\int \log L}{\int \beta_k} \frac{\int \log L}{\int \beta_l} \right) M \end{aligned}$$

como queríamos demostrar

2.3.2.- Test de los multiplicadores de Lagrange.-

Definiciones.-

Llamaremos $\hat{\beta}_n$ ó $\hat{\beta}$ al estimador de máxima verosimilitud restringido de β y $\tilde{\beta}$ al no restringido de β , donde $\tilde{\beta}_i = g_i(\hat{\delta}_1, \dots, \hat{\delta}_s)$, siendo $\hat{\delta}$ el estimador de máxima verosimilitud de δ .

2.3.3.- Planteamiento del problema.-

La idea para contrastar la hipótesis nula

$$H_0 \equiv \{ \tilde{\beta}^0 \in \tau = \mathcal{R} \cap \text{Ker } h \}$$

consiste en que, supuesta la existencia de $\hat{\beta}_n$ y de la variable aleatoria

$$\hat{\lambda}_n(\bar{x})$$

aceptamos la hipótesis si $\hat{\lambda}_n(\bar{x})$ está "suficientemente" próxima a cero

puesto que si $\tilde{\beta}^0 \in \tau \rightarrow \tilde{\lambda}_0 = \bar{0}$ y como

$$E(\sqrt{n}(\hat{\lambda}_n - \tilde{\lambda}_0)) = 0$$

entonces $\hat{\lambda}_n \xrightarrow{P} \bar{0}$

La medida de la distancia a cero de $\hat{\lambda}$ en este test propuesto por (Aitchison y Silvey, 1960) para la hipótesis anterior es

$$S_c = \tilde{v}'(\hat{\beta}) J_{\hat{\beta}}^{-1} \tilde{v}(\hat{\beta}) = \sqrt{n} \hat{\lambda}' (-R_{\hat{\beta}}) \sqrt{n} \hat{\lambda}$$

donde, en primer lugar

$$\tilde{\lambda}_0 = (\lambda_1^0, \dots, \lambda_r^0)$$

es un vector de multiplicadores de Lagrange tal que

$$n^{-1} D \log L(\tilde{\beta}^0) + H_{\tilde{\beta}^0} \tilde{\lambda}_0 = \bar{0}$$

En segundo lugar $R_{\hat{\beta}}$ es tal que

$$\therefore \begin{bmatrix} J_{\bar{\beta}^0} & -H_{\bar{\beta}^0} \\ -H'_{\bar{\beta}^0} & \underline{0} \end{bmatrix}^{-1} = \begin{bmatrix} \dots & \dots \\ \dots & R_{\bar{\beta}^0} \end{bmatrix}$$

luego

$$R_{\bar{\beta}^0} = -H'_{\bar{\beta}^0} J_{\bar{\beta}^0}^{-1} H_{\bar{\beta}^0}$$

siempre que $J_{\bar{\beta}^0}$ sea regular.

Este test, según veremos en el Teorema 1 (2.3.4), es asintóticamente equivalente al test del cociente de verosimilitudes y al test secuencial de Wald.

En el primero se acepta la hipótesis si

$$\mu(\bar{x}) = \frac{\sup_{\bar{\beta} \in \mathcal{C}} L(\bar{x}, \bar{\beta})}{\sup_{\bar{\beta} \in \mathcal{R}} L(\bar{x}, \bar{\beta})}$$

está "suficientemente cerca" de 1, y la medida de la cercanía a 1, viene dada por $-2 \log \mu$.

En el segundo, suponiendo la existencia de $\hat{\beta}_n(\bar{x}, \mathcal{R})$ aceptamos la hipótesis si $\bar{h}(\hat{\beta}_n)$ está "suficientemente cerca" de cero.

La medida de la proximidad a $\bar{0}$ de $\bar{h}(\hat{\beta}_n)$ usada por Wald es,

$$-n (\bar{h}(\hat{\beta}_n))' R_{\hat{\beta}_n} (\bar{h}(\hat{\beta}_n))$$

donde $\hat{\beta}_n$ es el estimador, no restringido, de máxima verosimilitud y $R_{\hat{\beta}_n}$ es la asintótica matriz de covarianzas de $\bar{h}_j(\hat{\beta}_n)$, de forma tal que

$$R_{\hat{\beta}_n} = -H'_{\hat{\beta}_n} J_{\hat{\beta}_n}^{-1} H_{\hat{\beta}_n}$$

donde

$$H = \left(\frac{\partial h_i(\bar{\beta})}{\partial \beta_j} \right)_{\bar{\beta} = \hat{\beta}_n} (p'+q') \times r$$

2.3.4.-TEOREMA.-

Los tres estadísticos tienden asintóticamente a una misma distribución $\chi^2_{p'+q'-s}$

Demostración.-

A partir de los lemas 1 y 2 se sigue que asintóticamente

$-2(\log L(\bar{x}, \hat{\beta}) - \log L(\bar{x}, \bar{\beta})), \bar{d}' J \bar{d}$ y $\bar{v}' J^{-1} \bar{v}$ tienen una misma distribución.

Análogamente del lema 3, también asintóticamente

$-2(\log L(\bar{x}, \hat{\delta}) - \log L(\bar{x}, \bar{\delta})), \bar{c}' N \bar{c}$, $\bar{u}' N^{-1} \bar{u}$ y $\bar{v}' M N^{-1} M' \bar{v}$

tienen una misma distribución.

De ambos resultados asintóticos se deduce que

$-2 \log \mu = 2(\log L(\bar{x}, \hat{\beta}) - \log L(\bar{x}, \hat{\delta}))$ y $\bar{v}' (J^{-1} - M N^{-1} M') \bar{v}$ tiene la misma distribución asintótica.

Ahora para que la variable aleatoria segunda tenga una distribución χ^2 es condición necesaria y suficiente (Rao, (1973)), que se verifica que la igualdad matricial siguiente, en virtud del resultado (2.2.6.1)

$$(2.3.4.1) \quad J (J^{-1} - M N^{-1} M') J (J^{-1} - M N^{-1} M') J = J (J^{-1} - M N^{-1} M') J$$

Esta igualdad se verifica al desarrollar:

$$\begin{aligned} & (J^{-1} - M N^{-1} M') J (J^{-1} - M N^{-1} M') = \\ & = J^{-1} J J^{-1} - J^{-1} J M N^{-1} M' - M N^{-1} M' J J^{-1} + M N^{-1} (M' J M) N^{-1} M = \\ & = J^{-1} - 2 M N^{-1} M' + M N^{-1} N N^{-1} M' = J^{-1} - M N^{-1} M' \end{aligned}$$

donde hemos tenido en cuenta el resultado (iv) del Lema 3

Así, como

$$(J^{-1}MN^{-1}M')J(J^{-1}MN^{-1}M') = J^{-1}MN^{-1}M'$$

,se verifica la igualdad (2.3.4.1).

Para hallar el número de grados de libertad de la distribución, se calcula la traza de la matriz

$$\begin{aligned} \text{tr}(J^{-1}MN^{-1}M')J &= \text{tr}(I - MN^{-1}M'J) = p' + q' - \text{tr}MN^{-1}M'J = \\ &= p' + q' - \text{tr}N^{-1}M'JM = p' + q' - \text{tr}N^{-1}N = p' + q' - s = r \end{aligned}$$

Luego el estadístico; $-2\log \mu$, del cociente de verosimilitudes se distribuye según una $\chi^2_{p'+q'-s}$.

Se sigue del trabajo de (Silvey, 1947) en los lemas 3 y 4 que

$$(2.3.4.2) \quad 2 \log \mu = n \hat{\lambda}' R_{\hat{\beta}} \hat{\lambda} + O_p(1)$$

y que

$$(2.3.4.3) \quad 2 \log \mu = n (\vec{h}(\vec{\beta}))' R_{\vec{\beta}} (\vec{h}(\vec{\beta})) + O_p(1)$$

De forma tal que los tres estadísticos asintóticamente tienen la misma distribución $\chi^2_{p'+q'-s=r}$, supuestas r restricciones, de los parámetros, linealmente independientes.

La elección natural de las regiones críticas de tamaño α para contrastar la hipótesis nula, $\vec{\beta}^0 = \gamma$ serán

$$C_1 = \{ \vec{x} / -2\log \mu > k_\alpha \}$$

$$C_2 = \{ \vec{x} / -n (\vec{h}(\vec{\beta}))' R_{\vec{\beta}} (\vec{h}(\vec{\beta})) > k_\alpha \}$$

$$C_3 = \{ \vec{x} / -n \hat{\lambda}' R_{\hat{\beta}} \hat{\lambda} > k_\alpha \}$$

En el mismo trabajo, se demuestra que los tests basados en las regiones críticas C_1 y C_3 tienen para $n \rightarrow \infty$ la misma potencia, así si $\bar{\beta}^0$ está cerca de τ , la potencia de cada test estará próxima a 1, mientras que si $\bar{\beta}^0$ no está cerca de τ , las variables $-2 \log \mu$ y $-n \hat{\lambda}' R_{\hat{\beta}} \hat{\lambda}$ tenderán asintóticamente a una distribución χ_r^2 , pero no centrada en r .

Análogo razonamiento hace Wald con los tests basados en las regiones críticas C_1 y C_2 , (Wald (1943)).

2.4.- TEST DE LOS MULTIPLICADORES DE LAGRANGE PARA MATRICES DE INFORMACION SINGULAR.-

2.4.1.- Planteamiento del problema.-

Ya hemos visto en el lema 2 que

$$-2 \log \mu = n (\hat{\beta} - \bar{\beta})' J_{\bar{\beta}^0} (\hat{\beta} - \bar{\beta}) + O_p(1)$$

luego asintóticamente se tiene que

$$(2.4.1.1) E(\log f(\bar{x}, \hat{\beta})) - E(\log f(\bar{x}, \bar{\beta}^0)) = -\frac{1}{2} (\bar{\beta} - \bar{\beta}^0)' J_{\bar{\beta}^0} (\bar{\beta} - \bar{\beta}^0) + O_p(1)$$

cuando $\bar{\beta}$ está próximo a $\bar{\beta}^0$ y se verifica la hipótesis 2(i).

Así si $J_{\bar{\beta}^0}$ no es definida positiva, puede ocurrir que $E(\log f(\bar{x}, \hat{\beta}))$ no presente un máximo en un entorno de $\bar{\beta}^0$ en \mathcal{R} y entonces no se pueda asegurar la consistencia de $\hat{\beta}_n$ a $\bar{\beta}^0$, con lo que parte de la estimación no restringida se vendría abajo.

En el caso de la estimación restringida, $\bar{\beta}^0 \in \tau = \mathcal{R} \cap \text{Ker } h$, aunque $J_{\bar{\beta}^0}$ no sea definida positiva, ni $E(\log f(\bar{x}, \bar{\beta}^0))$ sea un máximo de la función $E(\log f(\bar{x}, \bar{\beta})) \quad \forall \bar{\beta} \in \mathcal{R}$

se puede dar el caso de que

$$E(\log f(\bar{x}, \bar{\beta}^0))$$

si sea un máximo de la función anterior para $\bar{\beta} \in \mathcal{Z}$, de forma tal que para la estimación restringida solo sean necesarias algunas revisiones de las hipótesis de partida para desarrollar el test.

2.4.2.- Distribuciones asintóticas de $\hat{\beta}_n$ y $\hat{\lambda}_n$.

Hipótesis 1.-

si $\bar{\beta}^0 \in \mathcal{Z}$, $\forall \bar{\beta} \in \mathcal{Z}$ se tiene que $F(\bar{x}, \bar{\beta}) \neq F(\bar{x}, \bar{\beta}^0)$ para al menos un \bar{x} .

La suposición es necesaria para que los parámetros $\bar{\beta}$ sean identificables en \mathcal{R} , esto es, que no existan diferentes valores de $\bar{\beta}$ en \mathcal{R} que den la misma distribución en \bar{x} cuando lo que deseamos es que $\bar{\beta}$ sea identificable en \mathcal{Z} , así mediante esta suposición se puede asegurar que

$$E(\log f(\bar{x}, \bar{\beta}^0)) > E(\log f(\bar{x}, \bar{\beta})) \quad \forall \bar{\beta} \in \mathcal{Z}$$

Teniendo en cuenta la hipótesis de partida y esta hipótesis, los estimadores $\hat{\lambda}_n$ y $\hat{\beta}_n$, para casi todo \bar{x} , verifican asintóticamente las ecuaciones restringidas de verosimilitud, en virtud del resultado (2.2.3) de la proposición 3^a y bajo hipótesis nula de que $\bar{\beta} = \bar{\beta}^0$

$$(2.4.2.1) \quad \begin{cases} J_{\bar{\beta}^0}(\hat{\beta}_n - \bar{\beta}^0) - H_{\bar{\beta}^0} \hat{\lambda}_n = \frac{d n^{-1} \log L(\bar{x}, \bar{\beta})}{d \bar{\beta}} \Big|_{\bar{\beta} = \bar{\beta}^0} + O_p(1) \\ -H'_{\bar{\beta}^0}(\hat{\beta}_n - \bar{\beta}^0) = \bar{o} + O_p(1) \end{cases}$$

o matricialmente

$$\begin{bmatrix} J_{\bar{\beta}^0} & -H_{\bar{\beta}^0} \\ -H'_{\bar{\beta}^0} & \underline{0} \end{bmatrix} \begin{bmatrix} \hat{\beta}_n - \bar{\beta}^0 \\ \hat{\lambda}_n \end{bmatrix} = \begin{bmatrix} D n^{-1} \log L(\bar{x}, \bar{\beta}) + O_p(1) \\ \bar{o} + O_p(1) \end{bmatrix}$$

Al tener que hacer uso de la matriz inversa de J, ya no puede suponerse que la matriz

$$\left[\begin{array}{c|c} J_{\beta^0} & -H_{\beta^0} \\ \hline -H'_{\beta^0} & \underline{0} \end{array} \right]$$

sea regular, luego tenemos en su lugar la

Hipótesis 2.-

La matriz H_{β^0} es de rango r, o sea es de rango completo por columnas y la matriz J es de rango $p'+q'-t$ donde $t \leq r$, de forma tal que exista una submatriz H_1 de H, de dimensión $(p'+q') \times t$, tal que $A = J_{\beta^0} + H_1 H_1'$, sea una matriz definida positiva.

Ahora teniendo en cuenta la definición del vector v del Lema 1 (2.2.6), las ecuaciones (2.4.2.1) se pueden poner en la forma:

$$(2.4.2.2) \quad \left[\begin{array}{c|c} J_{\beta^0} & -H_{\beta^0} \\ \hline -H'_{\beta^0} & \underline{0} \end{array} \right] \begin{bmatrix} \hat{\beta} - \beta^0 \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} n^{-1/2} \bar{v} + O_p(1) \\ \bar{0} + O_p(1) \end{bmatrix}$$

Si particionamos H en la forma

$$H_{\beta^0} = \begin{bmatrix} H_1 & H_2 \\ t & r-t \end{bmatrix} \begin{matrix} p'+q' \\ \cdot \end{matrix}$$

donde H_1 está compuesta por las t primeras columnas de H, entonces a partir de las ecuaciones que surgen de (2.4.2.2) se tiene

$$(2.4.2.3) \quad \left[\begin{array}{c|c} A & -H_{\beta^0} \\ \hline -H'_{\beta^0} & \underline{0} \end{array} \right] \begin{bmatrix} \hat{\beta} - \beta^0 \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} n^{-1/2} \bar{v} + O_p(1) \\ \bar{0} + O_p(1) \end{bmatrix}$$

Pero si tenemos en cuenta la hipótesis 2.- entonces A es definida positiva y H es de rango r, luego la matriz del lado izquierdo de la igualdad es no singular. (Rao 1973).

Estamos ahora en condiciones de probar el siguiente lema :

2.4.3.- LEMA PREVIO.-

Teniendo en cuenta las hipótesis de partida y las hipótesis 1 y 2 se tiene:

i) La distribución asintótica de la variable

$$\sqrt{n} (\hat{\beta} - \beta^0) \text{ es una } N(\vec{0}, P_{\beta^0})$$

ii) La distribución asintótica de la variable

$$\sqrt{n} \hat{\lambda} \text{ es una } N(\vec{0}, K_{\beta^0})$$

siendo

$$K_{\beta^0} = H'_{\beta^0} A^{-1} H_{\beta^0} - \begin{bmatrix} I_t & | & \underline{0} \\ \underline{0} & | & \underline{0} \end{bmatrix} \begin{matrix} t \\ r-t \end{matrix}$$

y donde

$$(2.4.3.1) \quad \begin{bmatrix} A & | & -H_{\beta^0} \\ -H'_{\beta^0} & | & \underline{Q}_r \end{bmatrix}^{-1} = \begin{bmatrix} P_{\beta^0} & | & Q_{\beta^0} \\ \underline{Q}'_{\beta^0} & | & R^*_{\beta^0} \end{bmatrix} \begin{matrix} p+q' \\ r \end{matrix}$$

iii) las variables

$$\sqrt{n} (\hat{\beta} - \beta^0) \text{ y } \sqrt{n} \hat{\lambda} \text{ son independientes}$$

Demostración.-

De la ecuación matricial (2.4.2.3) se deduce que asintóticamente

$$\begin{bmatrix} \hat{\beta} - \beta^0 \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} A & | & -H_{\beta^0} \\ -H'_{\beta^0} & | & \underline{Q}_r \end{bmatrix}^{-1} \begin{bmatrix} n^{1/2} \vec{v} \\ \vec{0} \end{bmatrix} = \begin{bmatrix} P_{\beta^0} & | & Q_{\beta^0} \\ \underline{Q}'_{\beta^0} & | & R^*_{\beta^0} \end{bmatrix} \begin{bmatrix} n^{1/2} \vec{v} \\ \vec{0} \end{bmatrix}$$

y como \vec{v} se distribuye asintóticamente según una $N(\vec{0}, J_{\beta^0})$ se sigue que

$$\sqrt{n} \begin{bmatrix} \hat{\beta} - \beta^0 \\ \hat{\lambda} \end{bmatrix} \xrightarrow{L} N(\vec{0}, U' \begin{bmatrix} J & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} U)$$

donde

$$U' = \begin{bmatrix} P & Q \\ Q' & R^* \end{bmatrix} = U \quad \text{y} \quad -R^* = -H'A^{-1}H$$

ya que las matrices P y R son simétricas al serlo las matrices A y 0,

luego

$$U' \begin{bmatrix} J & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} U = \begin{bmatrix} PJP & PJQ \\ Q'JP & Q'JQ \end{bmatrix}$$

Ahora de la igualdad matricial

$$\begin{bmatrix} A & -H \\ -H' & \underline{0} \end{bmatrix} \begin{bmatrix} P & Q \\ Q' & R \end{bmatrix} = \begin{bmatrix} I & \underline{0} \\ \underline{0} & I \end{bmatrix}$$

se deduce que

$$(2.4.3.2) \quad HQ' = AP - I_{p'+q'}$$

$$(2.4.3.3) \quad AQ = HR^*$$

$$(2.4.3.4) \quad -PH = \underline{0}$$

$$(2.4.3.5) \quad -Q'H = I_r$$

De (2.4.3.2) si premultiplicamos por P se tiene $PAP = P + PHQ'$ y teniendo en cuenta el resultado (2.4.3.4) se sigue que $PAP = P$. (2.4.3.6).

Desarrollando la matriz

$$U' \begin{bmatrix} J & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} U$$

se tiene que

$$1^2) \quad PJP = PAP - PH_1 H_1' P = P - PH_1 H_1' P$$

pero como por (2.4.3.4) $PH = \underline{0}$, en particular, se tiene que $PH_1 = \underline{0}$

luego $PJP = P$ (2.4.3.7) y se obtiene el resultado i de (2.4.3

2º)

$$\begin{aligned} K &= Q'JQ = Q'AQ - Q'H_1H_1'Q = Q'HR^* - Q'H_1H_1'Q = \\ &= -I_r R^* - Q'H_1H_1'Q = -R^* - Q'H_1H_1'Q \end{aligned}$$

donde se han empleado los resultados (2.4.3.3) y (2.4.3.5)

Ahora si la ecuación (2.4.3.5) la particularizamos cuando particionamos H se tiene que

$$Q'H_1 = - \left[\begin{array}{c|c} I_t & \underline{0} \end{array} \right]'$$

luego

$$Q'H_1H_1'Q = \left[\begin{array}{c|c} -I_t & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array} \right]$$

y

$$K = Q'JQ = -R^* - \left[\begin{array}{c|c} -I_t & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array} \right] = H'A^{-1}H - \left[\begin{array}{c|c} -I_t & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array} \right] \text{ como se pedía en ii.}$$

3º)

$$PJP = PAQ - PH_1H_1'Q = PHR^* - PH_1H_1'Q = \underline{0} - \underline{0} = \underline{0}$$

donde se ha echo uso de los resultados (2.4.3.3), (2.4.3.4) y que $PH_1 = \underline{0}$

Análogamente

$$Q'JP = Q'AP - Q'H_1H_1'P = R^*H'P - Q'H_1H_1'P = \underline{0} - \underline{0} = \underline{0}$$

puesto que si $PH = \underline{0}$ también $H'P = \underline{0}$ y si $PH_1 = \underline{0}$ también $H_1'P = \underline{0}$

con lo que se completa la demostración.

2.4.4.- TEOREMA.-

Sujetos a las hipótesis iniciales junto con las hipótesis 1 y 2)

el estadístico $S_c^* = -\sqrt{n} \hat{\lambda}' R_{\beta}^* \sqrt{n} \hat{\lambda}$

se distribuye asintóticamente según una χ^2_{r-t}

Demostración.-

Del lema previo se tiene que $\sqrt{n} \hat{\lambda} \xrightarrow{L} N(\bar{\sigma}, K_{\beta_0})$
 si definimos K_1 matriz cuadrada y no singular tal que $K = K_1 K_1'$

entonces

$$(2.4.4.1) \quad \sqrt{n} K_1^{-1} \hat{\lambda} \xrightarrow{L} N(\bar{\sigma}, I_r)$$

luego

$$\begin{aligned} \sqrt{n} \hat{\lambda}' (K_1')^{-1} K_1^{-1} \hat{\lambda} \sqrt{n} &= n \hat{\lambda}' (K_1 K_1')^{-1} \hat{\lambda} = \\ &= n \hat{\lambda}' K^{-1} \hat{\lambda} \end{aligned}$$

se distribuye asintóticamente según una χ^2_r (2.4.4.2).

Lo anterior es cierto siempre que $K = Q' J Q$ sea invertible.

Ahora en el caso en que J no sea invertible habrá que hacer una transformación.

Si tenemos en cuenta la hipótesis 2, al ser la matriz A definida positiva y el rango de J , $p'+q'-t$, existirá una matriz C , no singular, cuadrada de dimensión $p'+q'$, tal que

$$(2.4.4.3) \quad C' A C = I_{p'+q'} \quad , \quad C C' A = I_{p'+q'} \quad , \quad A' C C' = I_{p'+q'}$$

y

$$C' J C = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} p'+q'-t \\ t \end{matrix}$$

donde D , es una matriz diagonal de dimensión $p'+q'-t$

luego

$$(2.4.4.4) \quad C' H_1 H_1' C = C' A C - C' J C = I_{p'+q'} - \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I_t \end{bmatrix}$$

La igualdad anterior última se debe a que como el rango de $H_1 H_1'$ es t , entonces hacemos $D = I_{p'+q'-t}$ (2.4.4.5)

Por otra parte a partir del resultado (2.2.6.1), $\vec{v} \xrightarrow{L} N(\vec{0}, J\vec{\beta}^0)$

y si hacemos la transformación $\vec{m} = C'\vec{v}$, entonces la variable vectorial \vec{m} de dimensión $p'+q'$, verifica que

$$(2.4.4.6) \quad \vec{m} \xrightarrow{L} N(\vec{0}, C'JC)$$

pero si tenemos en cuenta los resultados (2.4.4.3) y (2.4.4.4) se sigue que

$$C'JC = \begin{bmatrix} I_{p'+q'-t} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix}$$

luego las variables $m_1, m_2, \dots, m_{p'+q'-t}$ son, asintóticamente, variables aleatorias independientes $N(0,1)$, mientras que las restantes

$$m_{p'+q'-t+j} = 0 \quad \forall j = 1, \dots, t$$

De lo anterior se deduce que

$$(2.4.4.7) \quad \vec{m}'\vec{m} \xrightarrow{L} \chi^2_{p'+q'-t}$$

luego $\vec{m}'\vec{m} = \vec{v}'CC'\vec{v}$, y si tenemos en cuenta la primera ecuación del sistema (2.4.2.3), se tiene que salvo un infinitesimo $O_p(1)$ es

$$(2.4.4.8) \quad \vec{m}'\vec{m} = n \{ (\hat{\beta} - \beta^0)' A' - \hat{\lambda}' H' \} C C' \{ A(\hat{\beta} - \beta^0) - H \hat{\lambda} \} = \\ = n(\hat{\beta} - \beta^0)' A' C C' A(\hat{\beta} - \beta^0) - n(\hat{\beta} - \beta^0)' A' C C' H \hat{\lambda} - \\ - n \hat{\lambda}' H' C C' A(\hat{\beta} - \beta^0) + n \hat{\lambda}' H' C C' H \hat{\lambda}$$

Desarrollando los cuatro términos de la derecha de la igualdad y teniendo en cuenta las igualdades (2.4.4.3)

1)

$$n(\hat{\beta} - \beta^0)' A' C C' A(\hat{\beta} - \beta^0) = n(\hat{\beta} - \beta^0)' A' (\hat{\beta} - \beta^0) = \\ = n(\hat{\beta} - \beta^0)' J (\hat{\beta} - \beta^0) + n(\hat{\beta} - \beta^0)' H_1 H_1' (\hat{\beta} - \beta^0)$$

pero la segunda igualdad de la ecuación matricial (2.4.2.3), nos dice que

$$H' (\hat{\beta} - \beta^0) = \vec{0}$$

salvo un infinitesimo $O_p(1)$, luego

$$(2.4.4.9) \quad H' (\hat{\beta} - \beta^0) = \vec{0}$$

así

$$n (\hat{\beta} - \beta^0)' A' C C' A (\hat{\beta} - \beta^0) = n (\hat{\beta} - \beta^0)' J_{\hat{\beta}} (\hat{\beta} - \beta^0)$$

De igual forma

ii)

$$-n (\hat{\beta} - \beta^0)' A' C C' H \hat{\lambda} = -n (\hat{\beta} - \beta^0)' H \hat{\lambda} = 0$$

e

iii)

$$-n \hat{\lambda}' H' C C' A (\hat{\beta} - \beta^0) = -n \hat{\lambda}' H' (\hat{\beta} - \beta^0) = 0$$

con lo que sustituyendo en (2.4.4.8) se obtiene que

$$(2.4.4.10) \quad \vec{m}' \vec{m} = n (\hat{\beta} - \beta^0)' J (\hat{\beta} - \beta^0) + n \hat{\lambda}' H' C C' H \hat{\lambda}$$

salvo un infinitésimo $o_p(1)$

Por otra parte, a partir del resultado (ii) del Lema previo vamos a probar que

$$(2.4.4.11) \quad n \hat{\lambda}' H' C C' H \hat{\lambda} \xrightarrow{L} \chi^2_{r-t}$$

Demostración.-

Para ello vamos a comprobar que la forma cuadrática

$$n \hat{\lambda}' K (H' C C' H) K (H' C C' H) K \hat{\lambda}$$

es igual a

$$n \hat{\lambda}' K (H' C C' H) K \hat{\lambda}$$

ó más brevemente que

$$n \hat{\lambda}' (H' C C' H) K (H' C C' H) \hat{\lambda}$$

es igual a la forma cuadrática

$$n \hat{\lambda}' (H' C C' H) \hat{\lambda} \quad (\text{Rao, 1973})$$

En primer lugar, desarrollamos

$$(H' C C' H) K (H' C C' H)$$

teniendo en cuenta que

$$K = -R^* - \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix}$$

y de los resultados

(2.4.3.5), (2.4.4.3), (2.4.4.4) se sigue que

$$\begin{aligned}
 (2.4.4.12) \quad & H' C C' H (-R) H' C C' H - H' C C' H \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix} H' C C' H = \\
 & = -H' C C' H R H' C C' H - H' C C' \begin{bmatrix} H_1 & H_2 \end{bmatrix} \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H_1' \\ H_2' \end{bmatrix} C C' H = \\
 & = -H' C C' A Q H' C C' H - H' C C' H_1 H_1' C C' H = -H' Q H' C C' H - \\
 & -H' C C' H_1 H_1' C C' H = H' C C' H - H' C C' H_1 H_1' C C' H
 \end{aligned}$$

En segundo lugar, tenemos que de (2.4.4.4'), se sigue

$$\bar{m}' C' H_1 H_1' C \bar{m} = \bar{m}' \begin{bmatrix} 0 & 0 \\ 0 & I_t \end{bmatrix} \bar{m}_{(p'+q'-t) \times 1} = 0 \quad \text{lo que implica que}$$

$$H_1' C \bar{m} = \vec{0}_{t \times 1}$$

de donde

$$(2.4.4.13) \quad H_1' C \bar{m} = H' C C' \vec{v} = -\sqrt{n} H_1' C C' H \hat{\lambda} = \vec{0}_{t \times 1}$$

Luego teniendo en cuenta los resultados (2.4.4.12) y (2.4.4.13)

$$\begin{aligned}
 & (-\sqrt{n}) (-\sqrt{n}) \hat{\lambda}' (H' C C' H) K (H' C C' H) \hat{\lambda} = n \hat{\lambda}' H' C C' H \hat{\lambda} - \\
 & - (-\sqrt{n}) \hat{\lambda}' H' C C' H_1 (-\sqrt{n}) H_1' C C' H \hat{\lambda} = n \hat{\lambda}' H' C C' H \hat{\lambda} - \\
 & = \vec{0}'_{1 \times t} \vec{0}_{t \times 1} = n \hat{\lambda}' H' C C' H \hat{\lambda}
 \end{aligned}$$

Con lo que la forma cuadrática $n \hat{\lambda}' H' C C' H \hat{\lambda}$ tiene, asintóticamente, una distribución χ^2 .

Ahora para hallar el número de grados de libertad estudiamos la traza de la matriz

$$\begin{aligned}
 \text{traza } (H' C C' H) K &= \text{traza } (-H' C C' H R - H' C C' H \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix}) = \text{traza } (-H' C C' A Q - \\
 -H' C C' H \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix}) &= \text{traza } (-H' Q - H' C C' H \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix}) = \text{traza } I_r - \text{traza } C' H \begin{bmatrix} I_t & 0 \\ 0 & 0 \end{bmatrix} \\
 H' C &= r - \text{traza } C' H_1 H_1' C = r - \text{traza } \begin{bmatrix} 0 & 0 \\ 0 & I_t \end{bmatrix} = r - t.
 \end{aligned}$$

Luego queda probado el resultado (2.4.4.11).

Ahora para terminar, observamos que de (2.4.4.4')

$$\begin{aligned}
 H'_{\beta^0} C C' H_{\beta^0} &= H' (C'A)^{-1} (AC)^{-1} H = H'A^{-1} (C')^{-1} C^{-1} A^{-1} H = \\
 &= H' (ACC'A)^{-1} H = H'A^{-1} H = H' (J + H_1 H_1')^{-1} H = -R_{\beta^0}^*
 \end{aligned}$$

ya que si $CC'A = I$ entonces $C = (C'A)^{-1}$

y además $R_{\vec{\beta}^0}^*$ tiende a $R_{\vec{\beta}}^*$

luego $-n\hat{\lambda}' R_{\hat{\beta}}^* \hat{\lambda} \xrightarrow{L} \chi^2_{r-t}$ como queriamos demostrar

2.4.5. Consecuencia del Teorema.-

Como una consecuencia de lo anterior, se sigue, si aplicamos el teorema de Cochran a la igualdad (2.4.410), que

$$(2.4.4.1 \ 4) \quad n(\hat{\beta} - \vec{\beta}^0)' J (\hat{\beta} - \vec{\beta}^0) \xrightarrow{L} \chi^2_{p'+q'-t-(r-t)}$$

en el caso de que J sea singular y de rango $p'+q'-t$, y existan r restricciones en los parámetros $\vec{\beta}$.

2.5.- APLICACION DEL TEST DE LOS MULTIPLICADORES DE LANGRANGE PARA CONTRASTAR LA BONDAD DE MODELOS ARMA.-

2.5.1.- Introducción y Planteamiento.-

Sea el proceso ARMA (p,q) generador de la serie

$$(1.1) \quad \{x_t\} = \{x_1, \dots, x_n\}$$

$$\phi(B)x_t = \Theta(B)a_t$$

Si los datos x_t siguen de hecho un modelo ARMA (p',q') aplicando la técnica usual desarrollada por (Box-Jenkins,1976) para obtener estimaciones mínimo cuadráticas no lineales se producirán estimadores $\hat{\phi}_1, \dots, \hat{\phi}_p$ y $\hat{\theta}_1, \dots, \hat{\theta}_q$ inconsistentes si $p < p'$ y/o $q < q'$ y en tal situación el modelo (1.1) estará incorrectamente especificado ó identificado.

La bondad del modelo (1.1) se contrasta mediante la hipótesis nula:

$$H_0 \equiv \{x_t \in \text{ARMA}(p,q)\}$$

frente a las alternativas

$$H_A \equiv \{x_t \in \text{ARMA}(p+r,q)\} \quad \text{ó} \quad \{H_B \equiv x_t \in \text{ARMA}(p,q+r)\}$$

ó también

$$H_0 \equiv \{(1 + \phi_1 B + \dots + \phi_p B^p)x_t = (1 + \theta_1 B + \dots + \theta_q B^q)a_t\}$$

frente a las alternativas

$$\text{ó} \quad H_A \equiv \{(1 + \phi_1 B + \dots + \phi_{p+r} B^{p+r})x_t = (1 + \theta_1 B + \dots + \theta_q B^q)a_t\}$$

$$H_B \equiv \{(1 + \phi_1 B + \dots + \phi_p B^p)x_t = (1 + \theta_1 B + \dots + \theta_{q+r} B^{q+r})a_t\}$$

El modelo (1.1) se obtiene imponiendo en H_A ó H_B las restricciones

$$\phi_{p+1} = \dots = \phi_{p+r} = 0 \quad \text{ó} \quad \theta_{q+1} = \dots = \theta_{q+r} = 0$$

La especificación del test se basa en el test del multiplicador de Lagrange que ya hemos visto que necesita solo la estimación del modelo asociado a la hipótesis nula y que es asintóticamente equivalente a los tests de la razón de verosimilitudes y al test de Wald.

Para obtener el test, consideramos el log. de la función de verosimilitud de un proceso ARMA(p', q') condicionada a que las perturbaciones muestrales sean números fijos, (BOX-JENKINS, 1976).

Desarrollando el modelo (1.1) tenemos:

$$a_t = x_t + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

supuesto que las a_t son variables aleatorias independientes e idénticamente distribuidas $N(0, \sigma)$ entonces

$$\log L_n(a_1, a_2, \dots, a_n) = \log L_n(\vec{\phi}, \vec{\theta}, \sigma^2 / \vec{x}_*, \vec{a}_*, \vec{x})$$

donde

$$\vec{x}_* = (x_{1-p}, \dots, x_{1-p+q})', \vec{a}_* = (a_{1-q}, \dots, a_{1-q+q})' \text{ y } \vec{x} = (x_1, x_2, \dots, x_n)'$$

luego

$$\log L_n = -\frac{n}{2} \log \sigma^2 - \frac{\sum_{t=1}^n a_t^2}{2\sigma^2}$$

para concentrarla, hallamos el estimador de máxima verosimilitud de σ^2

$$\frac{d \log L_n}{d \sigma^2} = \frac{-n}{2\sigma^2} + \frac{\sum_{t=1}^n a_t^2}{2(\sigma^2)^2} = 0, \text{ Luego } \hat{\sigma}^2 = \frac{\sum_{t=1}^n a_t^2}{n}$$

Si llamamos

$$L = \sup_{\sigma} \log L_n(\vec{x}, \vec{\beta}, \sigma^2) = \log L_n(\vec{x}, \vec{\beta}, \hat{\sigma}^2)$$

tenemos que

$$L = -\frac{n}{2} \log \frac{\sum a_t^2}{n} - \frac{\sum a_t^2}{2n^{-1} \sum a_t^2} = -\frac{n}{2} \left(\log \frac{\sum a_t^2}{n} + 1 \right)$$

$$= -\frac{n}{2} \left[\log \sum a_t^2 + \log \frac{1}{n} + 1 \right] = K - \frac{n}{2} \log \sum_{t=1}^n a_t^2 (\vec{\beta} / \vec{x}_*, \vec{a}_*, \vec{x})$$

ó también

$$L = K - \frac{n}{2} \log \frac{\sum a_t^2(\vec{\beta} / \vec{x}_*, \vec{a}_*, \vec{x})}{n} = K - \frac{n}{2} \log \{s(\vec{\beta})\}$$

dónde

$$s(\vec{\beta}) = \frac{\vec{a}' \vec{a}}{n}$$

siendo

$$\vec{\beta}' = \left(\begin{array}{c} p' \\ \phi' \\ q' \\ \Theta' \end{array} \right)'$$

el vector compuesto por los parámetros de ϕ y Θ .

Para calcular los "eficientes scores" hallaremos el vector de las derivadas parciales primeras de $L(\vec{x}, \vec{\beta})$ respecto de $\vec{\beta}$.

$$\frac{dL(\vec{\beta})}{d\vec{\beta}} = \sqrt{n} \vec{v}(\vec{\beta}) = -\frac{1}{2} n \{s(\vec{\beta})^{-1}\} \left(\frac{d\vec{a}}{d\vec{\beta}} \right)' \vec{a} \frac{2}{n}$$

donde

$$\frac{d\vec{a}}{d\vec{\beta}} = \left(\frac{d a_t}{d \beta_i} \right)_{n \times (p'+q')}$$

para

$$t = 1, \dots, n$$

$$i = 1, \dots, p'+q'$$

Si aplicamos el Lema 1 (2.2.6), supuesto que la hipótesis nula es cierta se tiene que

$$\vec{v}(\vec{\beta})$$

se distribuye asintóticamente según una $N(0, J_{\vec{\beta}^0})$.

También en la proposición 3 se demuestra que

$$\left\{ n^{-1} \frac{d^2 \log L_n}{d\vec{\beta} d\vec{\beta}} \right\} \longrightarrow -J_{\vec{\beta}^0}$$

En nuestro caso particular el test estadístico para contrastar la hipótesis de que alguno de los ϕ_i ó Θ_i son iguales a cero, es

$$S_c = -n \hat{\lambda}' R_{\hat{\beta}} \hat{\lambda}$$

donde

$$-\sqrt{n} H \hat{\lambda} = \vec{v}(\hat{\beta})$$

$$y \quad R_{\hat{\beta}} = -H' J_{\hat{\beta}}^{-1} H$$

luego en particular

$$\hat{\lambda} = -n^{-1/2} \bar{v}(\hat{\beta}) \quad \text{y} \quad R_{\hat{\beta}} = -J_{\hat{\beta}}^{-1}$$

puesto que

$$H_{\hat{\beta}} = \begin{bmatrix} r \\ \underline{0} \\ \underline{I}_r \\ \underline{0} \end{bmatrix} \begin{matrix} p \\ r \\ q \end{matrix}$$

en el caso de ser la hipótesis alternativa H_A ó

$$H_{\hat{\beta}} = \begin{bmatrix} \underline{0} \\ \underline{I}_r \end{bmatrix} \begin{matrix} p+r \\ r \end{matrix}$$

en el caso de ser la hipótesis alternativa H_B .

Sustituyendo los valores de $\hat{\lambda}$ y R en S_c se tiene que

$$(2.5.1.1) \quad S_c = \bar{v}'(\hat{\beta}) J_{\hat{\beta}}^{-1} \bar{v}(\hat{\beta}) = \bar{v}'(\hat{\beta}) n \left\{ \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right) \right\}_{\hat{\beta}=\hat{\beta}}^{-1} \\ \left\{ S(\hat{\beta})^{-1} \right\}^{-1} \bar{v}(\hat{\beta}) = n^{-1} \left[(-1) \bar{a}' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right) (\bar{a}' \bar{a})^{-1} n n \left\{ \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right) \right\}_{\hat{\beta}=\hat{\beta}}^{-1} S(\hat{\beta}) \right. \\ \left. (-1) S(\hat{\beta})^{-1} \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)' \bar{a} \right]_{\hat{\beta}=\hat{\beta}} = n \bar{a}' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)_{\hat{\beta}=\hat{\beta}} \left\{ \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right) \right\}_{\hat{\beta}=\hat{\beta}}^{-1} \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)'_{\hat{\beta}=\hat{\beta}} \bar{a} (\bar{a}' \bar{a})^{-1}$$

donde $\hat{\beta}$ es el vector de los estimadores máximo verosímiles, obtenido cuando se verifica la hipótesis nula.

La ecuación (2.5.1.1) , revela que S_c es n veces el coeficiente de determinación de la regresión de \hat{a} , vector de residuos asociado a $\hat{\beta}$, en la matriz de derivadas parciales $\left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)_{\hat{\beta}=\hat{\beta}}$

Así

$$(2.5.1.2) \quad S_c = R^2 \left\{ \hat{a}, \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)_{\hat{\beta}=\hat{\beta}} \right\} = \frac{\hat{\beta}' \left\{ \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right) \right\}_{\hat{\beta}=\hat{\beta}} \hat{\beta}}{\bar{a}' \bar{a}}$$

en la regresión

$$(2.5.1.3) \quad \hat{a} = \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)_{\hat{\beta}=\hat{\beta}} \hat{\beta} + \bar{e}$$

donde

$$(2.5.1.4) \quad \hat{\beta} = \left\{ \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)' \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right) \right\}_{\hat{\beta}=\hat{\beta}}^{-1} \left(\frac{\partial \bar{a}}{\partial \hat{\beta}} \right)'_{\hat{\beta}=\hat{\beta}} \hat{a}$$

es el estimador mínimo cuadrático de la regresión (2.5.13)

Así el resultado (2.5.12) se aplicará para contrastar la bondad del modelo ARMA especificado en la hipótesis nula.

2.5.2.-Construcción de la regresión auxiliar.-

Partiendo del modelo (1.1), una vez obtenidos los estimadores máximo verosímiles y sustituidos estos en el modelo tenemos:

(2.5.2.1) $\hat{\phi}(B) x_t = \hat{\theta}(B) \hat{a}_t$

Definimos las variables $\bar{\mu}$ y $\bar{\nu}$, en la forma siguiente, en el supuesto de que hacemos nulos los valores premuestrales \bar{x}_* y \bar{a}_*

(2.5.2.2) $\mu_t = \nu_t = 0$ si $t \leq 0$

y

$\hat{\phi}(B) \mu_t = -x_t$

$\hat{\theta}(B) \nu_t = \hat{a}_t$

$\forall t = 1, \dots, n$

ó lo que es igual

$\mu_t = \nu_t = 0$

si $t \leq 0$

y

$\mu_t = -x_t - \hat{\phi}_1 \mu_{t-1} - \dots - \hat{\phi}_p \mu_{t-p}$

$\nu_t = \hat{a}_t - \hat{\theta}_1 \nu_{t-1} - \dots - \hat{\theta}_q \nu_{t-q}$

$\forall t = 1, \dots, n$

Despejando en (2.5.2.1) el valor \hat{a}_t , tenemos que

(2.5.2.3) $\hat{a}_t = x_t + \hat{\phi}_1 x_{t-1} + \dots + \hat{\phi}_p x_{t-p} - \hat{\theta}_1 a_{t-1} - \dots - \hat{\theta}_q a_{t-q}$

$\forall t = 1, \dots, n$

donde

$\hat{\beta} = (\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)'$

son los estimadores máximo verosímiles del modelo ARMA(p,q).

Si calculamos las derivadas parciales en (2.5.2.2), obtenemos

(2.5.2.4) $\frac{\partial \hat{a}_t}{\partial \hat{\phi}_i} = +x_{t-i} = -\hat{\phi}(B) \mu_{t-i}$ para $i = 1, \dots, p$

y

$\frac{\partial \hat{a}_t}{\partial \hat{\theta}_i} = -\hat{a}_{t-i} = -\hat{\theta}(B) \nu_{t-i}$ para $i = 1, \dots, q$

Desarrollando ((2.5.2.3) y teniendo en cuenta (2.5.2.2), se verifica

que

$$\begin{aligned}
 (2.5.2.5) \quad \hat{a}_t &= -\hat{\phi}(B)\mu_t - \hat{\phi}_1 \hat{\phi}(B)\mu_{t-1} - \dots - \hat{\phi}_p \hat{\phi}(B)\mu_{t-p} - \\
 &- \hat{\theta}_1 \hat{\theta}(B)\nu_{t-1} - \hat{\theta}_2 \hat{\theta}(B)\nu_{t-2} - \dots - \hat{\theta}_q \hat{\theta}(B)\nu_{t-q} = \\
 &= \hat{\phi}(B)(-\mu_t - \hat{\phi}_1 \mu_{t-1} - \dots - \hat{\phi}_p \mu_{t-p}) + \hat{\theta}(B)(\hat{\theta}_1 \nu_{t-1} - \dots - \hat{\theta}_q \nu_{t-q})
 \end{aligned}$$

pero como de (2.5.2.4) se deduce

$$\mu_{t-i} = -\frac{1}{\hat{\phi}(B)} \frac{\int \hat{a}_t}{\int \hat{\phi}_i} \quad y \quad \nu_{t-i} = -\frac{1}{\hat{\theta}(B)} \frac{\int \hat{a}_t}{\int \hat{\theta}_i}$$

los sustituimos en (2.5.2.5) y nos dan la relación

$$\hat{a}_t = \frac{\int \hat{a}_t}{\int \hat{\phi}_1} \hat{\phi}_1 + \dots + \frac{\int \hat{a}_t}{\int \hat{\phi}_p} \hat{\phi}_p + \frac{\int \hat{a}_t}{\int \hat{\theta}_1} \hat{\theta}_1 + \dots + \frac{\int \hat{a}_t}{\int \hat{\theta}_q} \hat{\theta}_q + \tau_t$$

$$\forall t=1, \dots, n$$

Luego para contrastar la hipótesis

$$H_0 \equiv \{x_t \in \text{ARMA}(p, q)\}$$

frente a la alternativa

$$H_A \equiv \{x_t \in \text{ARMA}(p+r, q)\}$$

solo será necesario estimar

$$(2.5.2.6) \quad \hat{a}_t = \alpha_1 \mu_{t-1} + \dots + \alpha_{p+r} \mu_{t-(p+r)} + \psi_1 \nu_{t-1} + \dots + \psi_q \nu_{t-q} + \tau_t$$

para $t=1, \dots, n$

por el método de mínimos cuadrados y calcular el estadístico multiplicador de Lagrange como n veces el coeficiente de determinación de la regresión

(2.5.2.6) donde se ha tenido en cuenta la igualdad (2.2.2.5)

Análogamente si la alternativa a la hipótesis H_0 fuera

$$H_B \equiv \{x_t \in \text{ARMA}(p, q+r)\}$$

la regresión en este caso sería

$$(2.5.2.7) \quad \hat{a}_t = \alpha'_1 \mu_{t-1} + \dots + \alpha'_p \mu_{t-p} + \rho'_1 \nu_{t-1} + \dots + \rho'_{q+r} \nu_{t-(q+r)} + \varepsilon_t$$

para $t = 1, \dots, n$

El estadístico S_c , ya se ha visto en el Teorema 1 (2.3.4) que asintóticamente está distribuido según una χ_r^2 cuando el modelo (1.1) es correcto.

Valores significativamente grandes indican que la especificación del modelo (1.1) no es apropiada.

2.5.3.- Aplicación a un modelo autorregresivo AR.-

Si en el modelo ARMA (p,q) , hacemos q=0, tenemos un modelo autorregresivo AR(p) ó ARMA(p,0)

$$(2.5.3.1) \quad \hat{\phi}(B)x_t = a_t$$

Si obtenemos los estimadores máximo verosímiles del modelo, bajo la hipótesis nula, se tiene que

$$(2.5.3.2) \quad \hat{\phi}(B)x_t = \hat{a}_t$$

En este caso, para construir la regresión auxiliar se define la variable μ_t , en la forma análoga a (2.5.2.2)

$$(2.5.3.3) \quad \mu_t = 0 \quad \text{si } t = 0$$

$$y \quad \mu_t = -x_t - \hat{\phi}_1 \mu_{t-1} - \dots - \hat{\phi}_p \mu_{t-p}$$

para $t = 1, 2, \dots, n$.

Desarrollando (2.5.3.2) se tiene que

$$\hat{a}_t = x_t + \hat{\phi}_1 x_{t-1} + \dots + \hat{\phi}_p x_{t-p}$$

y sustituyendo la transformación (2.5.3.3) y teniendo en cuenta el resultado (2.5.2.3) se sigue que

$$(2.5.3.4) \quad \begin{aligned} \hat{a}_t &= -\hat{\phi}(B)\mu_t - \hat{\phi}_1 \hat{\phi}(B)\mu_{t-1} - \dots - \hat{\phi}_p \hat{\phi}(B)\mu_{t-p} + z'_t = \\ &= \hat{\phi}(B)(-\mu_t - \hat{\phi}_1 \mu_{t-1} - \dots - \hat{\phi}_p \mu_{t-p}) \end{aligned}$$

Ahora si queremos contrastar la hipótesis

$$H_0 \equiv \{x_t \in \text{ARMA}(p,0)\}$$

frente a una hipótesis alternativa

$$H_A \equiv \{x_t \in \text{ARMA}(p+r,0)\}$$

la regresión auxiliar a construir será teniendo en cuenta (2.5.3.4)

$$(2.5.3.5) \quad \hat{a}_t = \alpha_1 \mu_{t-1} + \dots + \alpha_{p+r} \mu_{t-(p+r)} + z'_t$$

para $t = 1, \dots, n$

mientras que si H_0 es la misma, pero la alternativa es

$$H_B \equiv \{x_t \in \text{ARMA}(p, r)\}$$

entonces la regresión será

$$(2.5.3.6) \quad \hat{a}_t = \alpha_1 \mu_{t-1} + \dots + \alpha_p \mu_{t-p} + \varphi_1 v_{t-1} + \dots + \varphi_r v_{t-r} + \varepsilon'_t$$

para $t=1, \dots, n$

Pero en el caso autoregresivo; bajo la hipótesis nula, $q = 0$ y tenemos

$$\text{que } v_{t-1} = \hat{a}_{t-1}, v_{t-2} = \hat{a}_{t-2}, \dots, v_{t-r+1} = \hat{a}_{t-r+1} \text{ y } v_{t-r} = \hat{a}_{t-r}$$

en las ecuaciones (2.5.2.2), luego (2.5.3.6) se convierte en:

$$(2.5.3.7) \quad \hat{a}_t = \alpha_1 \mu_{t-1} + \dots + \alpha_p \mu_{t-p} + \varphi_1 \hat{a}_{t-1} + \dots + \varphi_r \hat{a}_{t-r} + \varepsilon'_t$$

para $t=1, \dots, n$

Luego la estimación mínimo cuadrática de (2.5.3.5) y de (2.5.3.7) dará el mismo coeficiente de determinación, lo que implica que el test del multiplicador de Lagrange de H_0 frente a H_A , es el mismo que el de H_0 frente a H_B .

2.5.4.- CONTRASTE DE H_0 FRENTE A LA ALTERNATIVA $H_c = \{x_t \in \text{ARMA}(p+r, q+s)\}$

En este caso existen problemas de singularidad en la matriz de varianzas - covarianzas de los parámetros

$$\bar{\beta}' = (\beta_1, \beta_2, \dots, \beta_{p'+q'}) = (\phi' : \theta')$$

con lo que el procedimiento estándar del test, como ya hemos visto en la sección 4ª(24), no puede usarse y es necesario imponer restricciones en el espacio paramétrico del modelo alternativo para obtener identificabilidad.

Lo anterior es cierto puesto que al anularse a la vez los parámetros AR y MA del modelo alternativo la matriz J es singular.

Singularidad de la matriz de covarianzas.

Definimos la función generadora teórica de autocovarianzas de x_t y de a_t en la forma

$$(2.5.4.1) \quad g_{xx}(z) = \sum_{k=-\infty}^{\infty} \gamma_{xx}(k) z^k \quad \text{ó} \quad g_{xx}(B) = \sum_{k=-\infty}^{\infty} \gamma_{xx}(k) B^k$$

y

$$g_{aa}(z) = \sum_{k=-\infty}^{\infty} \gamma_{aa}(k) z^k = \gamma_{aa}(0) z^0 = \sigma^2$$

$$g_{xa}(z) = \sum_{k=-\infty}^{\infty} \gamma_{xa}(k) z^k$$

donde

$$\gamma_{xx}(k) = E(x_t x_{t-k}) \quad \text{pues suponemos que } E(x_t) = 0$$

$$\text{Sea el modelo } \phi_{p+r}(B)x_t = \theta_{q+s}(B)a_t \quad (2.5.4.2)$$

$$t = 1, \dots, n$$

bajo la hipótesis nula y reestructurado en otra forma el modelo puede ponerse como:

$$(2.5.4.3) \quad x_t = \frac{\theta_{q'}(B)}{\phi_{p'}(B)} a_t \quad \text{y} \quad a_t = a_t$$

donde

$$\text{para } t = 1, \dots, n$$

$$p' = p+r$$

y

$$q' = q+s$$

Multiplicando (2.5.4.3) por a_{t-k} y tomando esperanzas se obtiene

$$\delta_{xa}(k) = \Theta_{q'}(B) \phi_{p'}^{-1}(B) \delta_{aa}(k) \quad \text{para } k = -\infty \dots \infty$$

luego

$$(2.5.4.4) \quad \delta_{xa}(k) \phi_{p'}(B) = \Theta_{q'}(B) \sigma^2$$

Análogamente multiplicando por x_{t-k} (2.5.4.3) y tomando esperanzas

$$\delta_{xx}(k) = \Theta_{q'}^2(B) \phi_{p'}^{-2}(B) \sigma^2 = \Theta_{q'}(B) \phi_{p'}^{-1}(B) \delta_{xa}$$

luego

$$(2.5.4.5) \quad \delta_{xx}(k) \phi_{p'}(B) = \Theta_{q'}(B) \delta_{xa}(k)$$

y de (2.5.4.4) y (2.5.4.5) obtenemos el sistema de ecuaciones

$$(2.5.4.6) \quad \begin{cases} \delta_{xx}(k) \phi_{p'}(B) - \delta_{xa}(k) \Theta_{q'}(B) = 0 & \text{para } k = 0, \dots, p' \\ \delta_{xa}(k) \phi_{p'}(B) + \sigma^2 \Theta_{q'}(B) = 0 & \text{para } k = 0, \dots, q' \end{cases}$$

donde

$$\phi_{p'}(B) = \sum_{i=0}^{p'} \phi_i B^i, \quad \Theta_{q'}(B) = \sum_{i=0}^{q'} \theta_i B^i$$

y

$$\phi_{p'}^{-1}(B) = \sum_{i=0}^{\infty} \phi_i^* B^i = \frac{1}{\phi_{p'}(B)} \quad \text{según definimos en (1.3.3.3)}$$

El sistema (2.5.4.6) en forma matricial es de la forma

$$(2.5.4.7) \quad \begin{matrix} p' \\ q' \end{matrix} \left[\begin{array}{c|c} \delta_{xx} & -\delta'_{xa} \\ \hline -\delta_{xa} & \sigma^2 I \end{array} \right] \begin{matrix} \bar{\phi}_{p'}^0 \\ \bar{\theta}_{q'}^0 \end{matrix} = \bar{\sigma}$$

donde

$$\bar{\phi}_{p'}^0 = (1, \phi_1, \dots, \phi_p, \overbrace{0, \dots, 0}^r)' \quad \text{y} \quad \bar{\theta}_{q'}^0 = (1, \theta_1, \dots, \theta_q, \overbrace{0, \dots, 0}^s)'$$

$$\delta_{xx} = \begin{bmatrix} \delta_{xx}(0) & \dots & \delta_{xx}(p-r) \\ \vdots & & \vdots \\ \delta_{xx}(p+r) & \dots & \delta_{xx}(0) \end{bmatrix}, \quad \delta_{xa} = \begin{bmatrix} \delta_{xa}(0) & \dots & \delta_{xa}(p') \\ \vdots & & \vdots \\ \delta_{xa}(-q-r) & \dots & \delta_{xa}(p'-q) \end{bmatrix} \quad \text{con} \quad \delta_{xx}(-k) = \delta_{xx}(k)$$

y llamando

$$\bar{\beta}^0 = (\bar{\phi}_{p'}^0 : \bar{\theta}_{q'}^0) \quad \text{y} \quad J = \left[\begin{array}{c|c} \delta_{xx} & -\delta'_{xa} \\ \hline -\delta_{xa} & \sigma^2 I \end{array} \right]$$

$$\text{entonces (2.5.4.8)} \quad J \bar{\beta}^0 = \bar{\sigma}$$

Para que el sistema homogéneo (25.4.8) tenga solución distinta de la trivial, el rango $J < n^2$ de incógnitas, luego $|J| = 0$ con lo que J es singular como queríamos probar.

Cálculo del score test.-

Para contrastar la hipótesis H_0 , frente a H_c , consideramos

$$\vec{\beta} = (\phi_1, \dots, \phi_{p'}, \theta_1, \dots, \theta_{q'})'$$

y tomando el logaritmo de la función de verosimilitud del modelo (2.5.4.2)

se obtiene

$$\log L_n(\vec{\beta}) = K - \frac{n}{2} \log \sigma^2 - \frac{\sum_{t=1}^n a_t^2}{2\sigma^2}$$

A partir de ella se construye el vector de los "eficientes scores"

y la matriz de varianzas-covarianzas de los parámetros $\vec{\beta}$, $J_{\vec{\beta}}$, a ésta matriz, converge con probabilidad 1 a la matriz de información de Fisher

$$\frac{1}{n} \left\{ \frac{\int^2 \log L_n(\vec{\beta})}{\int \vec{\beta}^2} \right\}$$

según vimos en la proposición 3 de la sección 2.

La matriz J se obtiene para las observaciones correspondientes a los parámetros del modelo ARMA (p+r, q+s).

En la sección 4^a se ha visto que en general el test para contrastar H_0 frente a H_c , se basa en el estadístico

$$S_c = \vec{v}'(\hat{\beta}) J_{\hat{\beta}}^{-1} \vec{v}(\hat{\beta}) = \sqrt{n} \hat{\lambda}' (-R_{\hat{\beta}}) \sqrt{n} \hat{\lambda}$$

donde

$$\vec{v} = \left(\frac{1}{\sqrt{n}} \frac{\int \log L_n}{\int \beta_1}, \dots, \frac{1}{\sqrt{n}} \frac{\int \log L_n}{\int \beta_{p'+q'}} \right)$$

$\hat{\lambda}$ y $R_{\hat{\beta}}$ se calculan usando las estimaciones de máxima verosimilitud de los parámetros del modelo ARMA (p,q) de la hipótesis nula.

Ahora como en este caso $|J| = 0$, si suponemos que

$$\text{rango } J = p'+q'-g \quad \text{siendo} \quad g \leq \min(r,s)$$

el estadístico a construir será

$$(2.5.4.9) \quad S_c^* = n \hat{\lambda}' (-R_{\hat{\beta}}^*) \hat{\lambda}$$

siendo

$$-R_{\hat{\beta}}^* = H_{\hat{\beta}}' (J + H_1 H_1')^{-1} H_{\hat{\beta}} = H_{\hat{\beta}}' J^{-1} H_{\hat{\beta}}$$

donde H_1 es una submatriz de orden $(p'+q) \times g$, elegida apropiadamente, de la matriz H , de rango $r+s$, de las derivadas parciales de las $r+s$ restricciones implícitas bajo la hipótesis nula.

Se ha visto en el teorema 2.4.4) que este estadístico se distribuye asintóticamente según una χ^2_{r+s-g} .

Aitchison y Silvey (1960) demuestran que la adaptación del "score test" a este caso usando la matriz H_1 para salvar el problema de la singularidad de J , es equivalente a imponer g restricciones en los parámetros en orden a obtener la identificabilidad de $\vec{\phi}_p$ y $\vec{\Theta}_q$.

Cuando la hipótesis nula es cierta, valores significativamente grandes de S_c^* indican que las restricciones impuestas no son consistentes con los datos.

La elección de H_1 no es arbitraria; puesto que hay que asegurar que $J + H_1 H_1'$ sea no singular en orden a emplear S_c^* , pero esto no es un obstáculo si desarrollamos el siguiente teorema:

2.5.5.-TEOREMA

El estadístico S_c^* es asintóticamente invariante respecto a la elección de H_1 .

Demostración.-

Nos basaremos en principio en el Lema 2.2.4 (ii) de (Rao y Mitra (1971)).

Definición.-

Una hipótesis alternativa admisible resulta cuando al imponer restricciones al espacio paramétrico del modelo alternativo, existe una matriz H_1 , tal que $J+H_1H_1'$ es no singular cuando es cierta la hipótesis nula.

Proposición Previa 1.-

Supuesto que la alternativa es admisible, se tiene que $(J+H_1H_1')^{-1} = J^-$ es una matriz generalizada inversa de J y entonces \hat{v} pertenece asintóticamente a $\mathcal{H}(J)$, que es el subespacio vectorial engendrado por las columnas de J .

Lo anterior es cierto debido a que siguiendo (2.4.2.3)

$$\sqrt{n} \begin{bmatrix} J+H_1H_1' & -H \\ -H' & 0 \end{bmatrix} \begin{bmatrix} \hat{\beta} - \beta^0 \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \hat{v} + O_p(1) \\ \hat{\sigma} + O_p(1) \end{bmatrix}$$

Ahora si

$$(J+H_1H_1')^{-1} = A^{-1} = J^-$$

es una matriz generalizada inversa, entonces existe

$$(2.5.5.1) \begin{bmatrix} A & -H \\ -H' & 0 \end{bmatrix}^{-1} = \begin{bmatrix} P & Q \\ Q' & R \end{bmatrix} \quad \text{tal que} \quad \sqrt{n} \begin{bmatrix} \hat{\beta} - \beta^0 \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} P \hat{v} \\ Q' \hat{v} \end{bmatrix}$$

asintóticamente

luego

$$\sqrt{n} (\hat{\beta} - \beta^0) = P \hat{v} \quad \text{y} \quad \sqrt{n} \hat{\lambda} = Q' \hat{v}$$

de donde se deduce que asintóticamente $\hat{v} \in \text{Im}(P^{-1})$

Teniendo en cuenta el resultado (2.4.3.7), en donde $(PJ-I)P = 0$

se sigue que asintóticamente $\hat{v} \in \text{Im}(J)$.

Una vez probada esta primera proposición siguiendo a (Rao 1973) aplicamos la siguiente:

Proposición Previa 2.-

BJ^-C es invariante para cualquier elección de J^- si y solo si $\text{Im}(B') \subset \text{Im}(J')$ e $\text{Im}(C) \subset \text{Im}(J)$.

En nuestro caso particular hacemos $B = H'$ y $C = H$ con lo que la propo-

sición queda de la forma :

$$\text{Im}(H) \subset \text{Im}(J)$$

si y solo si

$$H'J^{-1}H$$

es invariante para cualquier elección de J^{-1} y entonces S_c^* es invariante respecto a la elección de H_1 .

Demostración.-

El A^{-1} proyector ortogonal del espacio euclideo $R^{p'+q'}$ sobre el subespacio de dimensión r , $\mathcal{H} \subset R^{p'+q'}$ viene dado por

$$(2.5.5.2) \quad p_1^* = H(H'A^{-1}H)^{-1}H'A^{-1}$$

donde H es cualquier matriz tal que $\text{Im}(H) = \mathcal{H}$ y $(H'A^{-1}H)^{-1}H'A^{-1}$ es la mínima A^{-1} distancia generalizada inversa de H .

En (Pollock, (1980)) se demuestra que p_1^* es A^{-1} simétrica e idempotente.

Si tenemos en cuenta la igualdad matricial(2.5.5.1) entonces

$$P = A^{-1} + A^{-1}(-H) \left\{ -(-H')A^{-1}(-H) \right\}^{-1} (-H')(A')^{-1}$$

luego

$$(2.5.5.3) \quad P = A^{-1} - A^{-1}H(H'A^{-1}H)^{-1}H'A^{-1}$$

donde la mínima A -norma generalizada inversa de H' es $A^{-1}H(H'A^{-1}H)^{-1}$ y premultiplicando por A se obtiene

$$(2.5.5.4) \quad AP = I - P_1^*$$

A partir del resultado (2.4.3.2) se sigue que

$$HQ' = AP - I_{p'+q'} = -P_1^*$$

y de (2.5.5.1) que

$$Q' \hat{v} = \sqrt{n} \hat{\lambda}$$

de ambos obtenemos

$$HQ' \hat{v} = -P_1^* \hat{v}$$

de donde

$$\sqrt{n} H \hat{\lambda} = -P_1^* \hat{v}$$

con lo que concluimos, asintóticamente, que una proyección ortogonal de $\hat{v} \in \text{Im}(H)$ \hat{y} como por la proposición previa 1, $\hat{v} \in \text{Im}(J)$, también asintóticamente, implica que una proyección ortogonal suya también $\in \text{Im}(J)$, luego queda demostrada la proposición.

Otro problema importante es la manera de elegir las submatrices H_1 . Para ello, se elegirán de forma tal que no se anule $H_1 H_1'$ para alguno de los autovectores $\vec{\beta}_j$ correspondientes a los autovalores nulos de J , o sea,

$$(J + H_1 H_1') \vec{\beta}_j = (H_1 H_1') \vec{\beta}_j \neq 0$$

puesto que J y $H_1 H_1'$ son semidefinidas positivas, entonces la forma cuadrática $\vec{b}' J \vec{b}$ tiene al menos un autovalor de J nulo y los demás son positivos.

Los $\min(r,s)$ autovectores nulos de J vienen dados por

$$\vec{\beta}_j = B^j \vec{\beta}^0$$

para $j = 0, \dots, \min(r,s) - 1$, luego g definido en (2.4.2) es igual a $\min(r,s)$, siendo $\vec{\beta}^0$ el vector definido en (2.5.4.8) y B la matriz diagonal por bloques en la forma $B = \text{diag.}(C_{p+r}, C_{q+s})$ siendo C_m una matriz circular de dimensión m con fila inicial $(0, 0, \dots, 0, 1)$.

Estamos en condiciones de probar el siguiente teorema :

2.5.6.- TEOREMA

El "score test" para contrastar

$$H_0 = \{x_t \in \text{ARMA}(p, q)\}$$

frente a

$$H_{A^*} = \{x_t \in \text{ARMA}(p+m, q)\} \quad \text{ó} \quad H_{B^*} = \{x_t \in \text{ARMA}(p, q+m)\}$$

usando el estadístico S_c , donde $c = \max.(r,s)$, es asintóticamente equi-

valente a contrastar H_0 frente a $H_{C^*} = \{x_t \in \text{ARMA}(p+r, q+s)\}$ usando el estadístico S_C^* .

Demostración.-

Según hemos visto en la definición dada en(2.5.5), tanto la alternativa

H_{A^*} como la H_{B^*} son admisibles puesto que

$$H_{\beta^0} = \begin{bmatrix} m \\ \underline{0} \\ \underline{I}_m \\ \underline{0} \end{bmatrix} = \begin{bmatrix} 0 & \underline{0} \\ \underline{I}_g & \underline{0} \\ \underline{0} & \underline{I}_{m-g} \\ 0 & \underline{0} \end{bmatrix} \begin{matrix} p \\ g \\ m-g \\ q+g \end{matrix} = \begin{bmatrix} g & m-g \\ H_1 & H_2 \end{bmatrix} \begin{matrix} p+q+m \\ 9 \end{matrix}$$

si la hipótesis alternativa es H_{A^*}

y

$$H_{\beta^0} = \begin{bmatrix} m \\ \underline{0} \\ \underline{I} \end{bmatrix} = \begin{bmatrix} 0 & \underline{0} \\ \underline{I}_g & \underline{0} \\ 0 & \underline{I}_{m-g} \end{bmatrix} \begin{matrix} p+q+g & g & m-g \\ g & & \\ m-g & & \end{matrix} = \begin{bmatrix} g & m-g \\ H_1 & H_2 \end{bmatrix} \begin{matrix} p+q+m \\ 9 \end{matrix}$$

si la hipótesis alternativa es H_{B^*}

lo que implica una matriz $H_1 H_1'$ de dimensión $p'+q' = p+q+m$ cuyas filas, pueden reordenarse para dar

$$H_1 H_1' = \begin{bmatrix} p'+q' & g \\ \underline{0} & \underline{0} \\ \underline{0} & \underline{I}_g \end{bmatrix} \begin{matrix} p'+q' \\ 9 \end{matrix}$$

ó en este caso

$$H_1 H_1' = \begin{bmatrix} p+q+m & g \\ \underline{0} & \underline{0} \\ \underline{0} & \underline{I}_g \end{bmatrix} \begin{matrix} p+q+m \\ 9 \end{matrix}$$

Análogamente J se puede reorganizar en la forma

$$J = \begin{bmatrix} p+q+m & g \\ \underline{J}_{11} & \underline{J}_{12} \\ \underline{J}'_{12} & \underline{J}_{22} \end{bmatrix} \begin{matrix} p+q+m \\ 9 \end{matrix}$$

donde la matriz J_{11} es la matriz de covarianzas de los $p+q+m$ parámetros de las alternativas H_{A^*} ó H_{B^*} implícitamente especificadas.

Si tenemos en cuenta que, asintóticamente, $\hat{V} \in \text{Im}(J)$ y que el rango de J es $p+q+m$ se tiene que asintóticamente

$$\hat{V} = \begin{bmatrix} \hat{V}_1 \\ \hat{V}_2 \end{bmatrix} = \begin{bmatrix} \underline{J}_{11} & \underline{J}_{12} \\ \underline{J}'_{12} & \underline{J}_{22} \end{bmatrix} \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} \underline{J}_{11} \hat{y}_1 + \underline{J}_{12} \hat{y}_2 \\ \underline{J}'_{12} \hat{y}_1 + \underline{J}_{22} \hat{y}_2 \end{bmatrix} = \begin{bmatrix} \underline{J}_{11} \hat{y}_1 \\ \underline{J}_{21} \hat{y}_1 \end{bmatrix}$$

Ahora el estadístico S_c^* puede desarrollarse en la forma:

$$\begin{aligned}
 S_c^* &= \hat{v}'(\hat{\beta}) J_{\hat{\beta}}^{-1} \hat{v}(\hat{\beta}) = \begin{bmatrix} \hat{v}_1' \\ \vdots \\ \hat{v}_2' \end{bmatrix} J_{\hat{\beta}}^{-1} \begin{bmatrix} \hat{v}_1 \\ \vdots \\ \hat{v}_2 \end{bmatrix} = \\
 &= \begin{bmatrix} \hat{y}_1' J_{11} & \hat{y}_1' J_{12} \end{bmatrix} (J + H_1 H_1')^{-1} \begin{bmatrix} J_{11} \hat{y}_1 \\ J_{12} \hat{y}_1 \end{bmatrix} = \\
 &= \begin{bmatrix} \hat{y}_1' J_{11} & \hat{y}_1' J_{12} \end{bmatrix} \begin{bmatrix} J_{11} & J_{12} \\ J_{12}' & J_{22} + I_9 \end{bmatrix}^{-1} \begin{bmatrix} J_{11} \hat{y}_1 \\ J_{12} \hat{y}_1 \end{bmatrix} = \\
 &= \begin{bmatrix} \hat{y}_1' J_{11} & \hat{y}_1' J_{12} \end{bmatrix} \begin{bmatrix} A & B \\ B' & D \end{bmatrix} \begin{bmatrix} J_{11} \hat{y}_1 \\ J_{12} \hat{y}_1 \end{bmatrix} = \\
 &= \hat{y}_1' J_{11} A J_{11} \hat{y}_1 + \hat{y}_1' J_{12} B' J_{11} \hat{y}_1 + \hat{y}_1' J_{11} B J_{12}' \hat{y}_1 + \\
 &+ \hat{y}_1' J_{12} D J_{12}' \hat{y}_1
 \end{aligned}$$

Pero si tenemos en cuenta los resultados (2.3.4.2) y (2.4.3.3) se sigue que:

$$J_{12} B' = -J_{11} A + I_{p+q+m}$$

y que

$$J_{11} B = -J_{12} D$$

Si sustituimos estos resultados en S_c^* se obtiene que:

$$S_c^* = \hat{y}_1' J_{11} \hat{y}_1 = \hat{v}_1' J_{11}^{-1} J_{11} J_{11}^{-1} \hat{v}_1 = \hat{v}_1' J_{11}^{-1} \hat{v}_1$$

como queriamos demostrar.

2.5.7. Consecuencia del Teorema.-

Particularizando el teorema, se puede afirmar que el score test para contrastar $H(p,q)$, frente a cualquier hipótesis alternativa $H(p+r,q)$ es, intóticamente equivalente a contrastar la misma hipótesis nula frente a la alternativa $H(p,q+r)$, así el estadístico S_c es numéricamente invariante para contrastar una hipótesis frente a una amplia clase de hipótesis alternativas.

2.6.- EQUIVALENCIA ASINTOTICA DEL TEST DEL ESTADISTICO PORTMANTEAU Y DEL SCORE TEST EN UN CASO PARTICULAR.-

2.6.1.- INTRODUCCION:

Vamos a probar que el test portmanteau S es en realidad un score test cuando el modelo ARMA (p,q) se contrasta frente a la alternativa, de que x_t sea un proceso ARMA (p+m,q) ó (p,q+m).

2.6.2.- Equivalencia.-

Sea el modelo ARMA (p,q) (1.1) y lo queremos contrastar frente al ARMA (p+m,q), así sobreparametrizamos el modelo añadiéndole m parámetros autoregresivos

$$\phi_{p+i} \quad (i = 1, \dots, m)$$

Para grandes muestras el log de la función de verosimilitud condicionada, salvo constantes, para el modelo alternativo, es

$$(2.6.2.1) \quad \log L_n(\bar{\beta}) = -\frac{1}{2} n \log \sigma^2 - \sum_{t=1}^n a_t^2 / 2 \sigma^2$$

Calculamos el vector de las derivadas de log L con respecto a los parámetros $\bar{\beta}$, para $\bar{\beta} = \hat{\beta}$, sujeto a las m restricciones

$$\hat{\phi}_{p+i} = 0 \quad (i = 1, \dots, m)$$

Esto es,

$$(2.6.2.2) \quad \left. \frac{\partial \log L}{\partial \theta_i} \right|_{\bar{\beta} = \hat{\beta}} = 0 \quad \forall i = 1, \dots, q$$

$$\left. \frac{\partial \log L}{\partial \phi_i} \right|_{\bar{\beta} = \hat{\beta}} = 0 \quad \forall i = 1, \dots, p$$

mientras que

$$\left. \frac{\partial \log L}{\partial \phi_{p+i}} \right|_{\bar{\beta} = \hat{\beta}} = \hat{b}_i = - \sum_{t=1}^n \hat{a}_t \hat{v}_{t-(p+i)} / \hat{\sigma}^2 \quad \forall i = 1, \dots, m$$

Para obtener las derivadas (2.6.2.2) se considera el modelo alternativo, donde se han estimado los parámetros $\vec{\beta}$ por el estimador máximo verosímil $\hat{\beta}$ que sustituido en el modelo, nos da

$$\hat{\phi}_{p+m}(B) x_t = \hat{\theta}(B) \hat{a}_t \quad \text{ó} \quad \frac{\hat{\phi}_{p+m}(B)}{\hat{\theta}(B)} x_t = \hat{a}_t$$

luego si derivamos respecto a ϕ_{p+i} la función de verosimilitud (2.6.2.1) tenemos que

$$(2.6.2.3) \quad \hat{b}_i = \frac{-\sum_{t=1}^n 2 \hat{a}_t \frac{\partial \hat{a}_t}{\partial \phi_{p+i}}}{2 \hat{\sigma}^2}$$

donde

$$\frac{\partial \hat{a}_t}{\partial \phi_{p+i}} = \hat{v}_{t-(p+i)}$$

pero ya definimos en (1.3.2.1) que $\hat{\phi}(B) \hat{v}_t = -\hat{a}_t$, luego

$$(2.6.2.4) \quad \hat{v}_{t-(p+i)} = \frac{\partial \hat{a}_t}{\partial \phi_{p+i}} \Big|_{\vec{\beta}=\hat{\beta}} = \frac{\partial \left[\frac{\phi_{p+m}(B)}{\theta(B)} x_t \right]}{\partial \phi_{p+i}} \Big|_{\vec{\beta}=\hat{\beta}} =$$

$$= \frac{1}{\hat{\theta}(B)} \frac{\partial \phi_{p+m}(B) x_t}{\partial \phi_{p+i}} \Big|_{\vec{\beta}=\hat{\beta}} = \frac{1}{\hat{\theta}(B)} \left[-x_{t-(p+i)} \right] \Big|_{\vec{\beta}=\hat{\beta}}$$

Ahora bajo H_0 y para $\vec{\beta}=\hat{\beta}$ el modelo es $\hat{\phi}(B) x_t = \hat{\theta}(B) \hat{a}_t$ luego

$$x_{t-(p+i)} = \hat{\phi}^{-1}(B) \hat{\theta}(B) \hat{a}_{t-(p+i)}$$

resultado que sustituido en (2.6.2.4) y este a su vez en (2.6.2.3) nos queda

$$(2.6.2.5) \quad \hat{v}_{t-(p+i)} = \frac{\hat{\phi}^{-1}(B) \hat{\theta}(B) (-\hat{a}_{t-(p+i)})}{\hat{\theta}(B)} \quad \text{ó} \quad \hat{\phi}(B) \hat{v}_{t-(p+i)} = -\hat{a}_{t-(p+i)}$$

luego sustituyendo (2.6.2.5) en (2.6.2.3) se tiene que

$$\hat{b}_i = \frac{\partial \log L}{\partial \phi_{p+i}} \Big|_{\vec{\beta}=\hat{\beta}} = -\frac{\sum_{t=1}^n 2 \hat{a}_t \hat{v}_{t-(p+i)}}{2 \hat{\sigma}^2} = -\frac{\sum_{t=1}^n \hat{a}_t \hat{v}_{t-(p+i)}}{\hat{\sigma}^2}$$

para $i = 1, \dots, m$ como queríamos probar.

Ahora aplicando a nuestro caso particular los resultados del Lema 2^a (1.3.2) y del Lema 1^a (2.2.6) se tiene que

$$(2.6.2.6) \quad n^{-1/2} \hat{b}' = n^{-1/2} (\hat{b}_1, \dots, \hat{b}_m) \xrightarrow{L} N(\bar{0}, J_m)$$

siendo

$$J_m = \frac{1}{n} E(\hat{b} \hat{b}')$$

la submatriz cuadrada inferior derecha, de orden m , de Λ_m^{-1} , donde Λ_m es la matriz de covarianzas de $(u_{t-1}, \dots, u_{t-q}, v_{t-1}, \dots, v_{t-(p+1)}, \dots, v_{t-(p+m)})$ estando u_t y v_t definidos en (1.3.2.1), con

$$(2.6.2.7) \quad \Lambda_m = \begin{bmatrix} \gamma_{uu}(i-j) & \gamma_{uv}(i-j) & \gamma_{uv_m}(i-j) & q \\ \gamma_{vu}(i-j) & \gamma_{vv}(i-j) & \gamma_{vv_m}(i-j) & p \\ \gamma_{v_mu}(i-j) & \gamma_{v_mv}(i-j) & \gamma_{v_mv_m}(i-j) & m \end{bmatrix}$$

donde γ_{uu} , γ_{uv} y γ_{vv} están definidas en (1.3.2.2) y

$$\gamma_{v_mv_m}(i-j) = E(v_{t-(p+i)} v_{t-(p+j)})$$

entonces el estadístico del multiplicador de lagrange será

$$S_c = n^{-1} \hat{b}' J_m^{-1} \hat{b} = n \left(\frac{\hat{b}}{n} \right)' J_m^{-1} \left(\frac{\hat{b}}{n} \right)$$

Ahora, ya se ha visto en el capítulo 1^a, que un estadístico "portmanteau" basado en la m primeras autocorrelaciones muestrales residuales es (1.5)

$$S = n \hat{r}' C_1^{-1} \hat{r}$$

donde

$$C_1 = I_m - X \Lambda^{-1} X'$$

y vamos a probar que ambos estadísticos son asintóticamente equivalentes

Si expresamos (2.6.2.5) en la forma

$$\sum_{k=0}^p \hat{\phi}_k \hat{v}_{t-j-k} = -\hat{a}_{t-j} \quad \text{para } j = 1, \dots, m$$

Si ahora multiplicamos ambos miembros por \hat{a}_t y sumamos en t se sigue

$$\sum_{t=1}^n (\hat{a}_t \sum_{k=0}^p \hat{\phi}_k \hat{v}_{t-j-k}) = - \sum_{t=1}^n \hat{a}_t \hat{a}_{t-j}$$

pero si cambiamos los índices y ponemos $i = j-p+k$ nos queda sumando en i

$$\sum_{t=1}^n (\hat{a}_t \sum_{i=j-p}^j \hat{\phi}_{p-j+i} \hat{v}_{t-j-(p-j+i)}) = - \sum_{t=1}^n \hat{a}_t \hat{a}_{t-j}$$

luego

$$(2.6.2.8) \quad \sum_{i=j-p}^j (\hat{\phi}_{p-j+i} \sum_t \hat{a}_t \hat{v}_{t-p-i}) = - \sum \hat{a}_t \hat{a}_{t-j}$$

para $j = 1, \dots, m$

si se divide por $n \hat{\sigma}^2$ se tiene que

$$(2.6.2.9) \quad \hat{r} = \frac{\sum_{t=1}^n \hat{a}_t \hat{a}_{t-j}}{n \hat{\sigma}^2} = \hat{M} \frac{\hat{b}}{n}$$

donde \hat{M} es una matriz cuadrada de dimensión m donde el elemento i, j es $+\hat{\phi}_{p-i+j}$

para $i \geq j$, con $\hat{\phi}_0 = 1$ y $\hat{\phi}_{-j} = 0$ para $j \geq 1$, luego

$$\hat{M} = \begin{bmatrix} +\hat{\phi}_p & 0 & \dots & \dots & 0 \\ +\hat{\phi}_{p-1} & +\hat{\phi}_p & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ +\hat{\phi}_{p-m+1} & +\hat{\phi}_{p-m+2} & \dots & \dots & \hat{\phi}_p \end{bmatrix}$$

Ahora bajo la hipótesis nula, $\hat{\phi}_{p+j} = 0$ para $j = 1, \dots, m$,
 luego \hat{M} es triangular inferior y tal que

$$\hat{M} \hat{b} = \begin{pmatrix} +\hat{\phi}_{p-1+1} \hat{b}_1 \\ +\hat{\phi}_{p-2+1} \hat{b}_1 + \hat{\phi}_p \hat{b}_2 \\ \text{---} \\ \text{---} \\ +\sum_{j=1}^m \hat{\phi}_{p-m+j} \hat{b}_j \end{pmatrix}$$

con $\hat{b}_i = -\frac{\sum_t \hat{a}_t \hat{v}_{t-(p+i)}}{\hat{\sigma}^2}$
 para $i = 1, \dots, m$, siempre
 que $i > j$

Ahora identificando se sigue que

$$-\sum_{j=1}^m \hat{\phi}_{p-i+j} \hat{b}_j = \sum_t \hat{a}_t \hat{a}_{t-i} / \hat{\sigma}^2 \quad \text{o cambiando } i \text{ por } j$$

$$-\sum_{i=1}^m \hat{\phi}_{p-j+i} \hat{b}_i = \sum_t \hat{a}_t \hat{a}_{t-j} / \hat{\sigma}^2 \quad \text{para } \forall j = 1, \dots, m$$

el sumatorio de la izquierda se anula para $i > j$ bajo la hipótesis nula, luego

$$-\sum_{i=1}^j \hat{\phi}_{p-j+i} \hat{b}_i = \sum_t \hat{a}_t \hat{a}_{t-j} / \hat{\sigma}^2$$

o mejor

$$-\sum_{i=j-p}^j \hat{\phi}_{p-j+i} \hat{b}_i = \frac{\sum_t \hat{a}_t \hat{a}_{t-j}}{\hat{\sigma}^2} \quad \text{para } j = 1, \dots, m$$

que es la expresión (2.6.2.8) puesto que

$$\sum_t \hat{a}_t \hat{v}_{t-j} = 0 \quad \text{para } j = 1, \dots, p$$

debido a que las derivadas

$$\hat{b}_{j-p} = \frac{-\sum_t \hat{a}_t \hat{v}_{t-j}}{\hat{\sigma}^2} \quad \text{son nulas para } j = 1, \dots, p$$

ya que

$$\left. \frac{\partial \log L}{\partial \beta_i} \right|_{\hat{\beta} = \hat{\beta}} = 0 \quad \text{para } i = 1, \dots, p$$

al ser $\hat{\beta}$ el estimador máximo verosímil que anula las derivadas primeras de $\log L$ respecto a $\hat{\beta}_i$ y a $\hat{\theta}_j$ para $i = 1, \dots, p$ y $j = 1, \dots, q$.

También, teniendo en cuenta el resultado (2.6.2.6), se deduce la varianza asintótica de $\hat{\bar{r}}$; ya que $n^{-1} \hat{\bar{b}}$ se distribuye asintóticamente con media nula

Así teniendo en cuenta (2.6.2.9) se deduce que

$$E(\hat{\bar{r}} \hat{\bar{r}}') = E \left(\hat{M} \frac{\hat{\bar{b}}}{n} \frac{\hat{\bar{b}}'}{n} \hat{M}' \right) = n^{-1} \hat{M} E \left(\frac{\hat{\bar{b}}}{n} \frac{\hat{\bar{b}}'}{n} \right) \hat{M}' = n^{-1} \hat{M} J_m \hat{M}'$$

asintóticamente.

Ahora, teniendo en cuenta el resultado del T^a (1.3.4) se sigue que

$$E(\hat{\bar{r}} \hat{\bar{r}}') = n^{-1} \hat{C}_1$$

Luego $\hat{M} J_m \hat{M}' = \hat{C}_1$ asintóticamente, por lo que los estadísticos S y S_C son asintóticamente equivalentes, como queríamos demostrar.

2.6.3. Comentario.-

No tiene sentido en general emplear ambos tests; la utilización de uno u otro, viene determinada por la necesidad ó no de plantear un modelo alternativo al especificado en la hipótesis nula.

En el caso en que $m \rightarrow \infty$, la matriz C , ya hemos visto que tiende a ser singular y en lugar del estadístico S , se puede emplear el de Ljung-Box, $\tilde{Q}(\hat{\bar{r}})$. También en este caso, en el test del multiplicador de Lagrange, la matriz de covarianzas J_m se hace singular, si se contrasta un modelo ARMA (p,q) frente a un modelo ARMA, donde se aumente a la vez p y q . Este problema se abordó en este capítulo, en la sección 4^a.

De lo anterior se deduce que para $m \rightarrow \infty$, ambos tests son diferentes y es preferible en este caso utilizar el test multiplicador de Lagrange.

CAPITULO III

SCORE TEST Y ESTADISTICO PORTADNEAD PARA DIAGNOSTICA
MODELOS MARKA.

"SCORE TEST Y ESTADISTICO PORTMANTEAU PARA DIAGNOSTICAR MODELOS MARMA "

3.0.-	Sumario.....	84
3.1.-	Score Test para realizar contrastes de especificación en modelos MARMA	85
3.1.1.-	Hipótesis de partida	86
3.1.2.-	Planteamiento del problema.....	87
3.1.3.-	Cálculo de la matriz de varianzas-covarianzas de Θ	89
3.1.4.-	Teorema.....	91
3.1.5.-	Construcción del estadístico S_c^*	93
3.1.6.-	Problemas de identificabilidad cuando se hacen extensiones simultáneas en los polinomios $\Theta(z)$ y $\phi(z)$	96
3.1.7.-	Proposición.....	99
3.1.8.-	Conclusiones	101
3.2.-	Estadístico Portmanteau para modelos MARMA	103
3.2.1.-	Cálculo del score test para contrastar un modelo MARMA (p ,q) frente a un modelo MARMA (p ,q +s)	103
3.2.2.-	Distribución asintótica del estadístico S_c^* ...	106
3.2.3.-	Equivalencia asintótica de los estadísticos S_c^* y Q	112
3.3.-	Cálculo práctico del estadístico del score test.....	113
3.3.1.-	Construcción de la regresión auxiliar.....	113
3.4.-	APENDICE.....	117
A.1.-	Transformación F	117
A.1.1.-	Definición	117
A.2.-	Propiedades de la transformación F.....	118
A.2.1.-	Propiedad 1 ^a	118
A.2.2.-	Propiedad 2 ^a	118
A.2.3.-	Propiedad 3 ^a	119
A.2.4.-	Propiedad 4 ^a	121
A.3.-	Algunos resultados de algebra tensorial.	122

3.0.- SUMARIO.

El capítulo es una generalización de buena parte de las ideas desarrolladas en los anteriores capítulos, en él tratamos de contrastar la bondad de modelos autorregresivos y de medias móviles, pero donde el n^2 de series temporales es múltiple.

Comenzamos con un planteamiento general de los modelos MARMA y con el fin de construir el estadístico del score test S , en forma análoga a lo visto en el capítulo 2^a, usamos el producto del Kronecker, y las derivadas matriciales como un algebra más potente y rápida, BASILEVSKY, A. (1983), ROGERS, G.S. (1980) y POLLOCK, D.S.G. (1979).

Continuamos con el contraste de modelos MARMA usando el estadístico portmanteau, y viendo su equivalencia asintótica con el score test para una clase de hipótesis alternativas, para a continuación calcular el estadístico del score test como una regresión generalizada mínimo-cuadrática de los residuos sobre las derivadas estimadas.

Terminamos el capítulo con un Apéndice, donde definimos una función en el cuerpo de las series de potencias con coeficientes matriciales, que nos sirve como una potente herramienta en gran parte de las demostraciones acerca de la equivalencia asintótica de ambos tests.

3.1.- SCORE TEST PARA DIAGNOSTICAR MODELOS MARMA.

Sea $\{\vec{x}_t\}$ un vector, v dimensional, asociado a v series temporales, donde el proceso $\{\vec{x}_t\}$ sigue el modelo vectorial

$$(3.1) \quad \sum_{j=0}^p \underline{\phi}_j \vec{x}_{t-j} = \sum_{j=0}^q \underline{\theta}_j \vec{a}_{t-j}$$

con

$$\underline{\phi}_0 = \underline{\theta}_0 = I_v$$

y las matrices

$$\underline{\phi}_j \text{ y } \underline{\theta}_j$$

cuadradas de dimensión v , son de parámetros desconocidos

Por facilidad operativa suponemos que

$$E(\vec{a}_t / \mathcal{F}_{t-1}) = \vec{0} \quad \text{y que} \quad E(\vec{a}_t \vec{a}_t' / \mathcal{F}_{t-1}) = \Sigma$$

donde

$$\mathcal{F}_{t-1}$$

es la σ -álgebra de Borel generada por la historia pasada de $\vec{a}_z, \forall z \leq t$ y $\{a_t\}$ es una sucesión de martingalas gaussianas estacionarias y diferenciadas.

Si definimos

$$\underline{\phi}(z) = \sum_{i=0}^p \underline{\phi}_i z^i$$

y

$$\underline{\theta}(z) = \sum_{i=0}^q \underline{\theta}_i z^i$$

entonces el modelo (3.1) se puede poner en la forma:

$$\vec{a}_t = \sum_{j=0}^{n-1} \underline{A}_j \vec{x}_{t-j}$$

donde

$$\underline{A}(z) = \sum_{i=0}^{\infty} \underline{A}_i z^i = \underline{\theta}^{-1}(z) \underline{\phi}(z)$$

Para que el modelo (3.1) sea estacionario e invertible haremos las siguientes hipótesis de partida sobre $\underline{\phi}(z)$ y $\underline{\theta}(z)$

3.1.1.- Hipótesis de partida.-

1^a $\det \phi(z) \neq 0, |z| \leq 1$ y $\det \Theta(z) \neq 0$ para $|z| < 1$

De esta 1^a suposición, se deduce que \vec{x}_t solo depende de \vec{a}_{t-j} para $j \geq 0$, esto asegura que el modelo (3.1) es estacionario e invertible.

2^a $\phi(z)$ y $\Theta(z)$ son primos por la izquierda

Por esta suposición si

$$\Theta(z) = Q(z) \Theta_1(z) \quad \text{y} \quad \phi(z) = Q(z) \phi_1(z)$$

donde

$$Q, \Theta_1 \text{ y } \phi_1$$

son polinomios matriciales, entonces $Q(z)$ tiene un determinante constante, o lo que es igual, es de módulo unidad.

Esto hace que no existan raíces comunes en ambos polinomios.

3^a $[\phi_p : \Theta_q]$

es de rango v , ó equivalentemente (Hannan(1970)), $\phi(z)$ es una matriz triángular inferior con $\phi(0) = I_v$ y los elementos bajo la diagonal principal no tienen un grado mayor que los que están en la diagonal principal en la misma columna

Estas 3 hipótesis hacen que el modelo (3.1), que está dentro de la clase de equivalencia definida por la matriz de densidad espectral

$$g(\omega) = (2\pi)^{-1} \phi^{-1} \Theta \Sigma \Theta^* \phi^{*-1}$$

donde Θ y ϕ se calculan para $z = e^{i\omega}$, sea único, esto es, que sea identificable (Dunsmuir y Hannan (1976) y Deistler, Dunsmuir y Hannan (1978)).

Así el proceso estocástico $\{\vec{x}_t\}$ estacionario y ergódico es identificable, cuando se verifica(3.1).

Definimos por Θ el espacio paramétrico admisible de los parámetros estructurales $\bar{\theta}$ donde

$$\bar{\theta} = (\bar{\alpha}'; \bar{\mu}')_{1 \times (p+q)v^2}$$

siendo

$$\bar{\alpha} = \text{vec}(\phi_1; \dots; \phi_p)_{pv^2 \times 1}$$

y

$$\bar{\mu} = \text{vec}(\theta_1; \dots; \theta_q)_{qv^2 \times 1}$$

3.1.2.- Planteamiento del problema.

El problema general es contrastar

$$H_0 \equiv \{ \bar{\theta} = \bar{\theta}^0 \}$$

frente a

$$H_1 \equiv \{ \bar{\theta} = \bar{\theta}^1 \}$$

ó también

$$H_0 \equiv \{ \bar{h}(\bar{\theta}^0) = \bar{\sigma} \}$$

donde los f valores de la función vectorial \bar{h} , siendo $f < (p+q)v^2 = m_1$, son funciones continuas y doblemente diferenciables.

El estadístico como ya hemos visto, en el capítulo anterior, se construye en función del vector gradiente de una función concentrada de la de verosimilitud, que salvo constantes es:

$$(3.1.2.1) \quad L(\bar{\theta}) = -\frac{n}{2} \ln \left| \frac{\sum_{t=1}^n \bar{a}_t \bar{a}_t'}{n} \right| = -\frac{n}{2} \ln \det V$$

donde

$$V = -\frac{\sum_{t=1}^n \bar{a}_t \bar{a}_t'}{n} = \hat{\Sigma}$$

está calculada en $\bar{\theta} = \hat{\bar{\theta}}$, siendo $\hat{\bar{\theta}}$ el vector de las estimaciones máximo verosimiles de los parámetros del modelo, bajo la hipótesis nula, esto es, $\hat{\bar{\theta}}$ minimiza a $l(\bar{\theta})$ sujeto a las restricciones de los parámetros.

Así, $l(\bar{\theta})$ es el log. de la función de verosimilitud condicionada a que los valores premuestrales \bar{a}_j^* y \bar{x}_j^* se consideren fijos y en concreto $\bar{a}_j^* = \bar{0}$, donde \bar{a}_j^* esta definido para $(j = 0, -1, \dots, -q+1)$ y \bar{x}_j^* para $(j = 0, -1, \dots, -p+1)$.

Para llegar al resultado (3.1.2.1) partimos del supuesto más general de que las $\bar{a}_t \in N(\bar{0}, \Sigma) \quad \forall t = 1, \dots, n$

luego

$$L_1(\vec{a}_t / \vec{x}_t, \vec{\theta}, \Sigma) = (2\pi)^{-\frac{nv}{2}} |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2} \sum_{t=1}^n \vec{a}_t' \Sigma^{-1} \vec{a}_t\right\}$$

y

$$(3.1.2.2) L_2 = \ln L_1 = -\frac{nv}{2} \log 2\pi - \frac{n}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^n \vec{a}_t' \Sigma^{-1} \vec{a}_t$$

si sustituimos Σ por su estimador máximo verosímil obtenido en la forma:

$$\frac{\partial L_2}{\partial \Sigma^{-1}} = -\frac{n}{2} \frac{\partial \log |\Sigma|}{\partial \Sigma^{-1}} - \frac{1}{2} \frac{\partial \sum_{t=1}^n \vec{a}_t' \Sigma^{-1} \vec{a}_t}{\partial \Sigma^{-1}} =$$

$$= -\frac{n}{2} (-\Sigma) - \frac{1}{2} \sum_{t=1}^n \vec{a}_t \vec{a}_t' = 0_{v \times v}$$

de donde

$$\hat{\Sigma} = \frac{\sum \vec{a}_t \vec{a}_t'}{n} = \frac{A A'}{n} = V$$

siendo

$$A = (\vec{a}_1 : \vec{a}_2 : \dots : \vec{a}_n)_{v \times n}$$

valor que sustituido en (3.1.2.2); nos dá:

$$L(\vec{a}_t / \vec{x}_t, \vec{\theta}) = K_1 - \frac{n}{2} \log \frac{|\sum \vec{a}_t \vec{a}_t'|}{n} - \frac{1}{2} \text{traza}(A' \hat{\Sigma}^{-1} A) =$$

$$= K_1 - \frac{n}{2} \log \frac{|\sum \vec{a}_t \vec{a}_t'|}{n} - \frac{1}{2} \text{traza}(A A' \Sigma^{-1}) = K_1 - \frac{n}{2} \log \frac{|\sum \vec{a}_t \vec{a}_t'|}{n} -$$

$$- \frac{1}{2} \text{traza}\{A A' (A A')^{-1} n\} = K_2 - \frac{n}{2} \log \frac{|\sum \vec{a}_t \vec{a}_t'|}{n}$$

como queríamos ver.

Al suponer que $\{\vec{a}_t\}$ es un proceso gaussiano, construimos el vector gradiente de la función criterio en la forma

$$\vec{s}'(\vec{\theta}) = \frac{\partial L(\vec{\theta})}{\partial \vec{\theta}} = -\frac{\partial \left[\frac{n}{2} \ln \frac{1}{n} |\sum \vec{a}_t \vec{a}_t'| \right]}{\partial \vec{\theta}} = -\frac{n}{2} \frac{\partial \ln |V|}{\partial \vec{\theta}} = -\frac{n}{2} \frac{\partial |V|}{\partial V^c} \frac{\partial V^c}{\partial \vec{\theta}}$$

$$= \frac{1}{|V|} - \frac{n}{2} \frac{|V|}{|V|} (V^{-1})^r \frac{\partial \sum (\vec{a}_t \vec{a}_t')^c}{\partial \vec{\theta}} = \frac{1}{2} (V^{-1})^r \sum_{t=1}^n \frac{\partial (\vec{a}_t \vec{a}_t')^c}{\partial \vec{\theta}}$$

y si tenemos en cuenta el resultado del apartado A.3.1 del Apéndice

entonces

$$\vec{s}'(\vec{\theta}) = -\sum_{t=1}^n \vec{a}_t' V^{-1} D_t$$

de donde se sigue que

$$(3.1.2.3) \vec{s}'(\vec{\theta}) = -\sum \frac{\partial \vec{a}_t'}{\partial \vec{\theta}} V^{-1} \vec{a}_t = -\sum D_t' V^{-1} \vec{a}_t$$

donde

$$D'_t = \frac{\int \vec{a}_t'}{\int \vec{\theta}} \quad \text{es una matriz de dimensión } m_1 \times v$$

El estimador gaussiano $\hat{\vec{\theta}}$ de $\vec{\theta}$ se obtiene siguiendo los resultados del capítulo 2º, sección 2ª, como solución del sistema de ecuaciones vectoriales

$$\begin{cases} \vec{s}(\vec{\theta}) + H'(\vec{\theta}) \vec{\lambda} = \vec{0} \\ \vec{h}(\vec{\theta}) = \vec{0} \end{cases}$$

siendo

$$H'(\vec{\theta}) = \frac{\int \vec{h}'(\vec{\theta})}{\int \vec{\theta}} = \left[\frac{\int h_j}{\int \theta_i} \right]_{(p+q)v^2 \times f} \quad \begin{array}{l} \text{para } j=1, \dots, f \\ \text{y } i=1, \dots, (p+q)v^2 \end{array}$$

una matriz de rango completo por columnas, lo que quiere decir que las f restricciones son linealmente independientes, además el vector $\vec{\lambda}$ contiene los f multiplicadores de Lagrange.

3.1.3.- Cálculo de la matriz de información

Ahora vamos a calcular la matriz de varianzas-covarianzas de los parámetros $\vec{\theta}_{(p+q)v^2 \times 1}$, construyendo antes la matriz de información de Fisher

$$\Omega_n = \frac{1}{n} \left[\frac{\int^2 L(\vec{\theta})}{\int \vec{\theta} \int \vec{\theta}} \right]$$

donde

$$\begin{aligned} (3.1.3.1) \quad \Omega_n &= \frac{1}{n} \frac{\int}{\int \vec{\theta}} \left(\frac{\int L(\vec{\theta})}{\int \vec{\theta}} \right) = \frac{1}{n} \frac{\int}{\int \vec{\theta}} \left(\sum_{t=1}^n \frac{\int \vec{a}_t'}{\int \vec{\theta}} V^{-1} \vec{a}_t \right) = \frac{1}{n} \sum \frac{\int}{\int \vec{\theta}} \left(\frac{\int \vec{a}_t'}{\int \vec{\theta}} V^{-1} \vec{a}_t \right) = \\ &= -\frac{1}{n} \sum_{t=1}^n \left(\frac{\int^2 \vec{a}_t'}{\int \vec{\theta} \int \vec{\theta}} V^{-1} \vec{a}_t + \frac{\int \vec{a}_t'}{\int \vec{\theta}} \frac{\int V^{-1} \vec{a}_t}{\int \vec{\theta}} + \frac{\int \vec{a}_t'}{\int \vec{\theta}} V^{-1} \frac{\int \vec{a}_t'}{\int \vec{\theta}} \right) = \\ &= -\frac{1}{n} \sum \left\{ (V^{-1} \vec{a}_t \otimes I) \frac{\int (\frac{\int \vec{a}_t'}{\int \vec{\theta}})^c}{\int \vec{\theta}} + (\vec{a}_t \otimes \frac{\int \vec{a}_t'}{\int \vec{\theta}}) \frac{\int (V^{-1})^c}{\int \vec{\theta}} + \right. \\ &\quad \left. + (I \otimes \frac{\int \vec{a}_t'}{\int \vec{\theta}} V^{-1}) \left(\frac{\int \vec{a}_t'}{\int \vec{\theta}} \right) \right\} \end{aligned}$$

pero desarrollando el 2º sumando obtenemos que

$$\begin{aligned} &-\frac{1}{n} \sum (\vec{a}_t' \otimes \frac{\int \vec{a}_t'}{\int \vec{\theta}}) \frac{\int (V^{-1})^c}{\int V^c} \frac{\int V^c}{\int \vec{\theta}} = -\frac{1}{n} \sum (\vec{a}_t' \otimes \frac{\int \vec{a}_t'}{\int \vec{\theta}}) \left\{ -((V^{-1}) \otimes V^{-1}) \right\} \frac{\int V^c}{\int \vec{\theta}} = \\ &= -\frac{1}{n} \sum (\vec{a}_t' \otimes \frac{\int \vec{a}_t'}{\int \vec{\theta}}) \left\{ -((V^{-1}) \otimes V^{-1}) \right\} \left\{ \frac{1}{n} \sum_{t=1}^n (\vec{a}_t \otimes I_v) D_t + \frac{1}{n} \sum_{t=1}^n (I_v \otimes \vec{a}_t) D_t \right\} = \\ &= n^{-2} (\vec{\psi}_n^* \vec{\psi}_n \otimes \vec{a}_t' V^{-1} \otimes D_t V^{-1}) \left\{ (\vec{\psi}_n^* \vec{\psi}_n \otimes \vec{a}_t \otimes D_t) + (\vec{\psi}_n^* \vec{\psi}_n \otimes D_t \otimes \vec{a}_t) \right\} = \end{aligned}$$

$$= n^2 \left(\bar{\psi}_n^* \bar{\psi}_n \bar{\psi}_n^* \bar{\psi}_n \otimes \bar{a}_t' V^{-1} \bar{a}_t \otimes D_t' V^{-1} D_t \right) + n^2 \left(\bar{\psi}_n^* \bar{\psi}_n \bar{\psi}_n^* \bar{\psi}_n \otimes \bar{a}_t' V^{-1} D_t \otimes D_t' V^{-1} \bar{a}_t \right) =$$

$$= \frac{n v}{n^2} \sum_{t=1}^n D_t' V^{-1} D_t + n^2 \left(\bar{s}'(\bar{\theta}) \otimes \bar{s}(\bar{\theta}) \right) = \frac{v}{n} \sum_{t=1}^n D_t' V^{-1} D_t \quad (3.1.3.2)$$

puesto que $\bar{s}'(\bar{\theta}) \otimes \bar{s}(\bar{\theta}) = \underline{0}_{m_1 \times m_1}$ y $\bar{a}_t' V^{-1} \bar{a}_t = \text{traza } I_v = v$

así
$$\Omega_n = -\frac{1}{n} \sum_{t=1}^n \left(\bar{a}_t' V^{-1} \otimes I_{m_1} \right) \frac{\int (D_t')^c}{\int \bar{\theta}} - \frac{v}{n} \sum_{t=1}^n D_t' V^{-1} D_t - \frac{1}{n} \sum_{t=1}^n D_t' V^{-1} D_t$$

Desarrollamos a continuación el primer sumando teniendo en cuenta

que :

$$\sum_{t=1}^n D_t' \bar{a}_t = \bar{0}_{m_1 \times 1} \quad (\text{Apéndice A.3.2})$$

y si observamos dicha igualdad respecto de $\bar{\theta}$ se obtiene

$$\frac{\int \sum_{t=1}^n D_t' I_v \bar{a}_t}{\int \bar{\theta}} = \underline{0}_{m_1 \times m_1} \quad \text{y partiendo de aquí}$$

$$\sum_{t=1}^n \left((I_v \bar{a}_t)' \otimes I_{m_1} \right) \frac{\int (D_t')^c}{\int \bar{\theta}} + \sum_{t=1}^n (\bar{a}_t' \otimes D_t') \frac{\int I_v^c}{\int \bar{\theta}} + \sum_{t=1}^n (1 \otimes D_t') D_t = \underline{0}$$

pero como el segundo sumando es la matriz nula se verifica que:

$$\sum_{t=1}^n (\bar{a}_t' \otimes I_{m_1}) \frac{\int (D_t')^c}{\int \bar{\theta}} = - \sum_{t=1}^n D_t' D_t$$

luego

$$(3.1.3.3) \quad \sum_{t=1}^n (\bar{a}_t' V^{-1} \otimes I_{m_1}) \frac{\int (D_t')^c}{\int \bar{\theta}} = - \sum_{t=1}^n D_t' V^{-1} D_t$$

por lo que el primer sumando de Ω_n es igual a

$$\frac{1}{n} \sum_{t=1}^n D_t' V^{-1} D_t$$

y teniendo en cuenta los resultados (3.1.3.2) y (3.1.3.3)

se sigue que

$$\Omega_n = \frac{v}{n} \sum_{t=1}^n D_t' V^{-1} D_t$$

Ahora veremos que Ω_n converge casi seguro a la matriz Ω , siendo el rango $\Omega = m_1 - g = (p+q)v^2 - g$ para $0 \leq g \leq f$, y teniendo en cuenta que es su rango por columnas, ($m_1 - g$ columnas linealmente independientes) y que no son las mismas que las de la matriz $H_{f \times m_1}^1$ (de rango f), ya podemos enunciar el Teorema siguiente debido a (Kohn, 1978 y 1979) que es una extensión de los vistos por Rao (1947) y Silvey (1959).

3.1.4.- TEOREMA 1.-

Bajo una serie de suposiciones análogas a las vistas en las hipótesis de partida (2.1.2) se sigue que:

1ª) para $\forall n \geq n_1$ existe casi seguro una sucesión $\hat{\theta}_n$ que maximiza $L_n(\hat{\theta})$ para $\hat{\theta} \in \Theta_n = \Theta \cap \text{Ker } h$

2ª) para cada n , se elige una sucesión $\hat{\theta}_n$ tal que

$$\hat{\theta}_n \xrightarrow{\text{c.s.}} \hat{\theta}^0$$

y además

$$\Omega_n(\hat{\theta}_n) \xrightarrow{\text{c.s.}} \Omega(\hat{\theta}^0)$$

donde se satisface que

$$\tilde{h}(\hat{\theta}^0) = \bar{0}$$

La demostración, es análoga a las vistas en (2.2.2) y en (2.2.3)

Ahora el "score test" que contrasta la hipótesis nula de que $\tilde{h}(\hat{\theta}^0) = \bar{0}$ será:

$$(3.1.4.1) \quad S = n^{-1} \hat{s}' \hat{\Omega}^{-1} \hat{s} = n^{-1} \hat{\lambda}' \hat{H}' \hat{\Omega}^{-1} \hat{H} \hat{\lambda}$$

donde la tilde denota que las matrices se calculan para $\hat{\theta} = \hat{\theta}_n$

y donde $\hat{\Omega}^{-1}$ es una matriz generalizada inversa.

Ahora ya estamos en condiciones de aplicar el Teorema 2.4.4, donde se demuestra que el estadístico S_c^* bajo la hipótesis nula $\bar{h}(\bar{\Theta}) = \bar{0}$ converge en distribución a una χ^2_{f-g} donde $f-g$ es el n° de grados de libertad.

También en Kohn (1979) se demuestra el siguiente Teorema.-

Bajo una sucesión Pitman de hipótesis alternativas $\bar{h}(\bar{\Theta}^0) = n^{-1}\bar{\xi}$, siendo $\bar{\xi}$ un vector constante no nulo, S_c^* converge en distribución a una $\chi^2(f-g, \mathcal{J}_0)$ siendo \mathcal{J}_0 el parámetro de descentralidad de la distribución, e igual a $\bar{\xi}' H(H' \Omega^{-1} H)^{-1} H' \bar{\xi}$ donde el parámetro se calcula para $\bar{\Theta} = \bar{\Theta}^0$ verdadero valor de los parámetros del modelo, y siendo $g = \min(r, s)v^2$ y $f = \text{rango } H$.

Para concluir el teorema, es necesario ver que S es invariante respecto a la elección de la matriz generalizada inversa, pero esto ya se ha visto en el capítulo 2º, Teorema 2.5.5, donde se aplican los resultados de Rao y Mitra (1971. Teorema 9.2.3). Así, si

$$n^{-1/2} \hat{s} \xrightarrow{L} N(\hat{\mu}, \hat{\Omega})$$

entonces

$$n^{-1} \hat{s}' \hat{\Omega}^{-1} \hat{s} \xrightarrow{L} \chi^2(k, \mathcal{J})$$

siempre que $\hat{\mu} \in \text{Imagen } \hat{\Omega}$ de forma tal que $\hat{s}' \hat{\Omega}^{-1} \hat{s}$ sea invariante, respecto a la elección de $\hat{\Omega}^{-1}$, o que $\hat{\Omega}^{-1}$ sea una matriz reflexiva y simétrica g -inversa de $\hat{\Omega}$ y donde $k = \text{rango } \hat{\Omega}$ y $\mathcal{J} = \hat{\mu}' \hat{\Omega}^{-1} \hat{\mu}$, siendo k el n° de grados de libertad y \mathcal{J} el parámetro de descentralización.

También se demuestra que $\hat{s}' \hat{\Omega}^{-1} \hat{s}$ es invariante respecto a la elección de $\hat{\Omega}^{-1}$ solo con probar que $\hat{s} \in \text{Imagen } (\hat{\Omega})$ y aplicar en la Proposición Previa 2 (2.5.5). que $\hat{s} \in \text{Imagen } (\hat{\Omega})$ está probado en la Proposición Previa 1 (2.5.5).

De todo lo anterior vemos que $S_c^* = n^{-1} \hat{s}' \hat{\Omega}^{-1} \hat{s}$ es invariante respecto a la elección de $\hat{\Omega}^{-1}$ y que converge en distribución a una $\chi^2(f-g, 0)$, (3.1.4.2)

3.1.5.-Construcción de la matriz de covarianzas de $\vec{\theta}$

Teniendo en cuenta el resultado (3.1.2.3) como el vector $\vec{\theta} = (\vec{\alpha}; \vec{\mu})$
 $(p+q)v^2 \times 1$

vamos primero a descomponer la matriz

$$D_t = \frac{\int \vec{a}_t}{\int \vec{\theta}}$$

a su vez en dos matrices $\frac{\int \vec{a}_t}{\int \vec{\alpha}}$ y $\frac{\int \vec{a}_t}{\int \vec{\mu}}$

Ahora como vimos en (3.1)

$$A(z, \vec{\theta}) = \theta^{-1}(z, \vec{\theta}) \phi(z, \vec{\theta})$$

ó también

$$A(B) = \theta^{-1}(B) \phi(B)$$

si convenimos en aplicar B a la componente precedente del producto de Kronecker, entonces, el modelo (3.1) se puede poner en la forma

$$(3.1.5.1) \quad \vec{a}_t = \theta^{-1}(B) \phi(B) \vec{x}_t$$

si derivamos \vec{a}_t respecto a $\vec{\alpha}$ teniendo en cuenta (3.1.5.1) se sigue que

$$(3.1.5.2) \quad \frac{\int \vec{a}_t}{\int \vec{\alpha}} = \theta^{-1}(B) \frac{\int \phi(B)}{\int \vec{\alpha}} \vec{x}_t = (\vec{x}_t' \otimes \theta^{-1}(B)) \frac{\int \text{vec } \phi(B)}{\int \vec{\alpha}}$$

ahora si tenemos en cuenta la notación

$$\begin{aligned} \text{vec } \phi(z) &= \text{vec } \sum_{j=0}^{p'} \phi_j z^j = \sum_{j=0}^{p'} \text{vec } \phi_j z^j = \text{vec } \phi_0 z^0 + \\ &+ \sum_{j=1}^{p'} \text{vec } \phi_j z^j = \text{vec } I_v + \sum_{j=1}^{p'} \text{vec } \phi_j z^j = \text{vec } I_v + (\vec{\psi}'_p(z) \otimes I_{v^2}) \vec{\alpha} \end{aligned}$$

donde

$$\vec{\psi}'_p(z) = (z, z^2, \dots, z^{p'}) \quad \text{y} \quad \vec{\alpha} = \text{vec}(\phi_1; \dots; \phi_{p'})_{p v^2 \times 1}$$

y la aplicamos en (3.1.5.2)

$$\begin{aligned} \frac{\int \text{vec } \phi(B)}{\int \vec{\alpha}} &= \frac{\int \{ \text{vec } I_v + (\vec{\psi}'_p(B) \otimes I_{v^2}) \vec{\alpha} \}}{\int \vec{\alpha}} = \underline{0}_{v^2 \times p v^2} + \\ &+ \frac{\int (\vec{\psi}'_p(B) \otimes I_{v^2}) \vec{\alpha}}{\int \vec{\alpha}} = (\vec{\psi}'_p(B) \otimes I_{v^2}) I_{p v^2} \end{aligned}$$

y sustituyendolo en (3.15.2) se tiene que

$$(3.15.3) \quad \frac{\int \vec{a}_t}{\int \vec{\alpha}} = (\vec{x}_t' \otimes \theta^{-1}(B)) (\vec{\psi}'_p(B) \otimes I_{v^2}) = \vec{\psi}'_p(B) \otimes \vec{x}_t' \otimes \theta^{-1}(B)$$

Análogamente se halla

$$\frac{\partial \bar{a}_t}{\partial \bar{\mu}}$$

siendo $\varphi(B) \bar{x}_t = \theta(B) \bar{a}_t$

$$(3.1.5.4) \quad \frac{\partial \bar{a}_t}{\partial \bar{\mu}} = \frac{\partial (\theta^{-1}(B) \varphi(B) \bar{x}_t)}{\partial \bar{\mu}} = ((\varphi(B) \bar{x}_t)' \otimes I_v) \frac{\partial (\theta^{-1}(B))'}{\partial \bar{\mu}} =$$

$$= ((\theta(B) \bar{a}_t)' \otimes I_v) \frac{\partial \text{vec } \theta^{-1}(B)}{\partial \bar{\mu}}$$

operando con la nomenclatura anterior

$$\text{vec } \theta(z) = \text{vec } \sum_{j=0}^{q'} \theta_j z^j = \text{vec } \theta_0 z^0 + \sum_{j=1}^{q'} \text{vec } \theta_j z^j = \text{vec } I_v +$$

$$+ \sum_{j=1}^{q'} \text{vec } \theta_j z^j = \text{vec } I_v + (\bar{\varphi}'_q(z) \otimes I_{v^2}) \bar{\mu}$$

siendo

$$\bar{\mu} = \text{vec}(\theta_1 : \dots : \theta_q) \quad \text{y} \quad \bar{\varphi}'_q(z) = (z, z^2, \dots, z^q)$$

ahora derivando

$$\text{vec}(\theta^{-1}(B))$$

respecto de $\bar{\mu}$ se tiene:

$$\frac{\partial \text{vec } \theta^{-1}(B)}{\partial \bar{\mu}} = \frac{\partial \text{vec } \theta^{-1}(B)}{\partial \text{vec } \theta(B)} \frac{\partial \text{vec } \theta(B)}{\partial \bar{\mu}} =$$

$$= -((\theta^{-1})'(B) \otimes \theta^{-1}(B)) (0_{v^2 \times q v^2} + (\bar{\varphi}'_q(B) \otimes I_{v^2})) =$$

$$= -((\theta^{-1})'(B) \otimes \theta^{-1}(B)) (\bar{\varphi}'_q(B) \otimes I_{v^2})$$

expresión que sustituimos en (3.1.5.4), y obtenemos

$$(3.1.5.5) \quad \frac{\partial \bar{a}_t}{\partial \bar{\mu}} = \{(\theta(B) \bar{a}_t)' \otimes I_v\} \{-((\theta^{-1})'(B) \otimes \theta^{-1}(B))\} (\bar{\varphi}'_q(B) \otimes I_{v^2}) =$$

$$= -\{\bar{a}_t' \theta'(B) (\theta^{-1}(B))' \otimes \theta^{-1}(B)\} (\bar{\varphi}'_q(B) \otimes I_{v^2}) =$$

$$= -\{\bar{a}_t' (\theta^{-1}(B) \theta(B))' \otimes \theta^{-1}(B)\} (\bar{\varphi}'_q(B) \otimes I_{v^2}) =$$

$$= -\bar{\varphi}'_q(B) \otimes \bar{a}_t' \otimes \theta^{-1}(B)$$

Teniendo en cuenta los valores obtenidos para

$$D_{t\alpha} = \frac{\partial \bar{a}_t}{\partial \bar{\alpha}} \quad \text{y} \quad D_{t\mu} = \frac{\partial \bar{a}_t}{\partial \bar{\mu}}$$

ahora vamos a ver la forma de construir la matriz de varianzas-covarianzas de los parámetros $\bar{\theta}$ del modelo (3.1). Ya hemos visto que la matriz

de información de Fisher

$$\Omega_n = \frac{1}{n} \left(\frac{\partial^2 L(\vec{\theta})}{\partial \vec{\theta} \partial \vec{\theta}} \right) = \frac{V}{n} \sum_{t=1}^n D_t' V^{-1} D_t$$

y por proposiciones anteriores (2.2.3)

$$\hat{\Omega}_n \xrightarrow{\text{c.s.}} \Omega(\vec{\theta}^0)$$

bajo la hipótesis nula.

Vamos a probar entonces que

$$\Omega = \begin{bmatrix} \Omega_{\alpha\alpha} & \Omega_{\alpha\mu} \\ \Omega_{\mu\alpha} & \Omega_{\mu\mu} \end{bmatrix}$$

siendo

$$\Omega_{\alpha\alpha} = (2\pi)^{-1} \int_{-\pi}^{\pi} \vec{\psi}_p \vec{\psi}_p^* \otimes K V K^* \otimes (\Theta V \Theta')^{-1} d\omega$$

$$\Omega_{\alpha\mu} = (2\pi)^{-1} \int_{-\pi}^{\pi} \vec{\psi}_p \vec{\psi}_q^* \otimes K V \otimes (\Theta V \Theta')^{-1} d\omega$$

y

$$\Omega_{\mu\mu} = (2\pi)^{-1} \int_{-\pi}^{\pi} \vec{\psi}_q \vec{\psi}_q^* \otimes V \otimes (\Theta V \Theta')^{-1} d\omega$$

Donde $\vec{\psi}_p^*$ es el vector conjugado traspuesto de $\vec{\psi}_p$, Θ^{-} es la matriz generalizada inversa de $\Theta(B)$ y $K(z, \vec{\theta}) = A^{-1}(z, \vec{\theta})$

siendo

$$V = \frac{\sum \vec{a}_t \vec{a}_t'}{n} \quad \text{con} \quad P \lim_{n \rightarrow \infty} V = \Sigma = E(\vec{a}_t \vec{a}_t')$$

donde hemos sustituido $B \delta z$ por e^{iw} , con $w = -2\pi f$

luego

$$\Omega(\vec{\theta}^0) = \Omega_0 = P \lim_{n \rightarrow \infty} \frac{V}{n} \sum_{t=1}^n D_t' V^{-1} D_t$$

de donde

$$\begin{aligned} (3.1.5.6) \quad \Omega_{\alpha\alpha} &= (2\pi)^{-1} \int_{-\pi}^{\pi} D_{t\alpha}' V^{-1} D_{t\alpha} d\omega = (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{\partial \vec{a}_t'}{\partial \vec{\alpha}} V^{-1} \frac{\partial \vec{a}_t}{\partial \vec{\alpha}} d\omega = \\ &= (2\pi)^{-1} \int_{-\pi}^{\pi} (\vec{\psi}_p^* \otimes \vec{x}_t' \otimes \Theta^{-1}(z))' V^{-1} (\vec{\psi}_p^* \otimes \vec{x}_t' \otimes \Theta^{-1}(z)) d\omega = \\ &= (2\pi)^{-1} \int_{-\pi}^{\pi} (\vec{\psi}_p \otimes \vec{x}_t \otimes \Theta^{*-1}(z)) (\vec{\psi}_p^* \otimes \vec{x}_t' \otimes \Theta^{-1}(z)) d\omega = \\ &= (2\pi)^{-1} \int_{-\pi}^{\pi} (\vec{\psi}_p \vec{\psi}_p^* \otimes \vec{x}_t \vec{x}_t' \otimes \Theta^{*-1}(z) V^{-1} \Theta^{-1}(z)) d\omega \end{aligned}$$

Ahora como

$$\vec{x}_t = \phi^{-1}(z) \Theta(z) \vec{a}_t$$

se obtiene que

$$\vec{x}_t = A^{-1}(z) \vec{a}_t = K(z) \vec{a}_t$$

luego

$$\vec{x}_t \vec{x}_t' = K(z) \vec{a}_t \vec{a}_t' K^*(z) = K(z) V K^*(z)$$

y sustituyendo éste resultado en (3.1.5.6) se sigue que

$$(3.1.5.7) \quad \Omega_{\alpha\alpha} = (2\pi)^{-1} \int_{-\pi}^{\pi} \vec{\varphi}_p \vec{\varphi}_p' \otimes K V K^* \otimes (\Theta(z) V \Theta^*(z))^{-1} d\omega$$

análogamente se obtienen los resultados de $\Omega_{\alpha\mu}$ y $\Omega_{\mu\mu}$.

3.1.6.-Problemas de identificabilidad cuando se hacen extensiones simultáneas en los polinómios $\Theta(z)$ y $\phi(z)$.

Definición:

Sea $\Gamma_{\eta\nu}(z)$ la función generatriz entre dos procesos cualesquiera $\{\eta_t\}$ y $\{\nu_t\}$

a saber,

$$(3.1.6.1) \quad \Gamma_{\eta\nu}(z) = \sum_{k=-\infty}^{\infty} \Gamma_k z^k = \sum_{k=-\infty}^{\infty} \gamma_{\eta\nu}(k) z^k$$

donde

$$\gamma_{\eta\nu}(i-j)$$

es el elemento de la matriz que está definido análogamente a (2.5.4.1)

Si la hipótesis nula es cierta, $\vec{h}(\vec{\Theta}^0) = \vec{0}$, entonces el modelo

(3.1) lo trasponemos y multiplicamos por la derecha por \vec{a}_{t-j}' , tomamos esperanzas y teniendo en cuenta que:

$$(3.1.6.2) \quad \begin{cases} \Gamma_{ax}(i-j) = \underline{0} & \text{si } i > j \\ \Gamma_{ax}(0) = \Sigma & \text{si } i = j \\ \Gamma_{aa}(i-j) = \Sigma & \text{para todo } i-j \end{cases}$$

entonces

$$(3.1.6.3) \quad \sum_{i=0}^j \Gamma_{ax}(i-j) \underline{\phi}'_i = \Gamma_{aa}(i-j) \underline{\theta}'_j \quad \text{si } i \leq j$$

y para $j = 1, \dots, q'$

con $\underline{\phi}'_0 = \underline{\theta}'_0 = I_V$ y con $\underline{\phi}'_i = \underline{0}$ si $i = p+1, \dots, q$

Análogamente si (3.1) lo multiplicamos por la derecha por \bar{x}'_{t-j} y tomamos esperanzas se sigue que

$$(3.1.6.4) \quad \sum_{i=0}^{p'} \underline{\phi}'_i \Gamma_{xx}(j-i) = \sum_{i=j}^{q'} \underline{\theta}'_i \Gamma_{ax}(j-i) \quad \text{si } i \geq j$$

y $j = 0, 1, \dots, p'$

siendo $\underline{\theta}'_i = \underline{0}$ si $p > q$ para $j = q+1, \dots, p$

De (3.1.6.3) y (3.1.6.4) se sigue que

$$(3.1.6.5) \quad \phi(z) \Gamma_{xx}(z) = \theta(z) \Gamma_{ax}(z)$$

donde $\phi(z)$ y $\theta(z)$ son de grado p y q respectivamente. si contrastamos un modelo MARMA (p, q) frente a un MARMA (p', q') con $p' = p+r$ y $q' = q+s$

Ahora desarrollamos (3.1.6.5) teniendo en cuenta que $\text{vec}(ABC)$

$= (C' \otimes A) \text{vec } B$, luego

$$(\Gamma'_{xx}(z) \otimes I_V) \text{vec } \phi(z) = (\Gamma'_{ax}(z) \otimes I_V) \text{vec } \theta(z)$$

pero como

$$\text{vec } \phi(z) = \text{vec } I_V + (\bar{\varphi}'_p(z) \otimes I_V) \bar{\alpha}$$

y

$$\text{vec } \theta(z) = \text{vec } I_V + (\bar{\varphi}'_q(z) \otimes I_V) \bar{\mu}$$

si lo sustituimos y operamos se sigue que

$$(\Gamma'_{xx}(z) \otimes I_V) \{ \text{vec } I_V + (\bar{\varphi}'_p(z) \otimes I_V) \bar{\alpha} \} = (\Gamma'_{ax}(z) \otimes I_V) \cdot \{ \text{vec } I_V + (\bar{\varphi}'_q(z) \otimes I_V) \bar{\mu} \}$$

de donde

$$(3.1.6.6) \quad (\bar{\varphi}'_p(z) \otimes \Gamma'_{xx}(z) \otimes I_V) \bar{\alpha} + (\Gamma'_{xx}(z) \otimes I_V) \text{vec } I_V = (\bar{\varphi}'_q(z) \otimes \Gamma'_{ax}(z) \otimes I_V) \bar{\mu} + (\Gamma'_{ax}(z) \otimes I_V) \text{vec } I_V$$

Si definimos los vectores $\vec{\beta}_1'$ y $\vec{\beta}_2'$ en la forma

$$\vec{\beta}_1' = ((\text{vec } I_V)' : \vec{\alpha}' : \vec{\sigma}')_{1 \times (v^2 + p \cdot v^2 + v^2(r-1))} = 1 \times v^2(p+r)$$

y

$$\vec{\beta}_2' = ((\text{vec } I_V)' : \vec{\mu}' : \vec{\sigma}')_{1 \times (v^2 + q \cdot v^2 + v^2(s-1))} = 1 \times v^2(q+s),$$

y tenemos en cuenta que las submatrices de Ω son los coeficientes de las potencias de z en las convoluciones $\Gamma_{xx}(z)$, $\Gamma_{xa}(z)$, $\Gamma_{aa}(z)$ del proceso multivariante de autocovarianzas, esto es

$$\Omega = \begin{bmatrix} \Omega_{\alpha\alpha} & \Omega_{\alpha\mu} \\ \Omega_{\alpha\mu} & \Omega_{\mu\mu} \end{bmatrix}$$

donde de forma análoga a (2.5.4.1)

$$\Omega_{\alpha\mu} = E(\vec{a}_i \vec{x}_j') = (\gamma_{ax}(i-j)),$$

$$\Omega_{\mu\mu} = E(\vec{a}_i \vec{a}_j') = (\gamma_{aa}(i-j)) = \text{diag}(\Sigma, \dots, \Sigma)$$

es una matriz diagonal por bloques y

$$\Omega_{\alpha\alpha} = E(\vec{x}_i \vec{x}_j') = (\gamma_{xx}(i-j)),$$

ya podemos poner (3.1.6.6) en la forma

$$(\text{vec } I_V)' (\Gamma_{xx}(z) \otimes I_V) + \vec{\alpha}' (\vec{\Psi}_p(z) \otimes \Gamma_{xx}(z) \otimes I_V) =$$

$$= (\text{vec } I_V)' (\Gamma_{ax}(z) \otimes I_V) + \vec{\mu}' (\vec{\Psi}_q(z) \otimes \Gamma_{ax}(z) \otimes I_V)$$

y se obtiene que

$$= ((\text{vec } I_V)' : \vec{\alpha}' : \vec{\sigma}') \begin{bmatrix} \Gamma_{xx}(z) \otimes I_V & & & & & \\ & \vec{\Psi}_p(z) \otimes \Gamma_{xx}(z) \otimes I_V & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \end{bmatrix} \begin{matrix} v^2 \\ p v^2 \\ v^2(r-1) \\ v^2 \\ v^2 \\ v^2(s-1) \end{matrix} =$$

$$= ((\text{vec } I_V)' : \vec{\mu}' : \vec{\sigma}') \begin{bmatrix} \Gamma_{ax}(z) \otimes I_V & & & & & \\ & \vec{\Psi}_q(z) \otimes \Gamma_{ax}(z) \otimes I_V & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \end{bmatrix} \begin{matrix} v^2 \\ q v^2 \\ v^2(s-1) \\ v^2 \\ v^2 \\ v^2(s-1) \end{matrix}$$

luego

$$(3.1.6.7) \quad \vec{\beta}_1' \Omega_{\alpha\alpha} = \vec{\beta}_2' \Omega_{\alpha\mu}$$

cuando

$$p' > p \quad \text{y} \quad q' > q$$

$$\text{siendo} \quad \underline{p'} = \underline{p+r} \quad \text{y} \quad \underline{q'} = \underline{q+s}$$

De forma similar se pueden manipular las ecuaciones generadoras

$$\phi(z) \Gamma_{\alpha\alpha}(z) = \Theta(z) \Sigma$$

y se obtiene que

$$(3.1.6.8) \quad \vec{\beta}_1' \Omega_{\alpha\mu} = \vec{\beta}_2' \Omega_{\mu\mu}$$

luego

$$(3.1.6.9) \quad \left. \begin{aligned} \vec{\beta}_1' \Omega_{\alpha\alpha} - \vec{\beta}_2' \Omega_{\alpha\mu} &= \vec{0} \\ \vec{\beta}_1' \Omega_{\alpha\mu} - \vec{\beta}_2' \Omega_{\mu\mu} &= \vec{0} \end{aligned} \right\} \vec{\beta}' \Omega = \vec{0}$$

con lo que se encuentran problemas de identificabilidad cuando se hacen extensiones simultáneas en los polinomios matriciales $\Theta(z)$ y $\phi(z)$.

3.1.7. Proposición .-

El rango de Ω es $((p'+q') - \min(r, s))v^2$

bajo la hipótesis nula

$$H_0 \equiv \vec{\Theta} = \vec{\Theta}^0 \quad \delta \quad H_0 \equiv \vec{h}(\vec{\Theta}^0) = \vec{0}$$

Demostración:

Dada la matriz

$$\Omega = \begin{bmatrix} p'v^2 & q'v^2 \\ \Omega_{\alpha\alpha} & \Omega_{\alpha\mu} \\ \Omega_{\alpha\mu}' & \Omega_{\mu\mu} \end{bmatrix} \begin{matrix} p'v^2 \\ q'v^2 \end{matrix}$$

si en $\Omega_{\alpha\alpha}$ eliminamos $v^2(p'-p) = v^2r$ filas y columnas, y en $\Omega_{\alpha\mu}$ $v^2(p'-p) = v^2r$ filas, se obtienen una submatriz cuadrada $\Omega_{\alpha\alpha}^1$ de orden $p \cdot v^2$ y una submatriz $\Omega_{\alpha\mu}^1$ de orden $p \cdot v^2 \times q'v^2$, con lo que se define una submatriz cuadrada Ω^I , de Ω , de orden $p \cdot v^2 + q'v^2$ y no singular

$$\Omega^I = \begin{bmatrix} p \cdot v^2 & q'v^2 \\ \Omega_{\alpha\alpha}^1 & \Omega_{\alpha\mu}^1 \\ (\Omega_{\alpha\mu}^1)' & \Omega_{\mu\mu} \end{bmatrix} \begin{matrix} p \cdot v^2 \\ q'v^2 \end{matrix}$$

luego, rango $\Omega^I = p'v^2 + q'v^2 - rv^2$

Si análogamente eliminamos en $\Omega_{\mu\mu} v^2(q'-q) = v^2s$, filas y columnas y en $\Omega_{\alpha\mu} v^2(q'-q)$, columnas, se tiene una submatriz cuadrada $\Omega_{\mu\mu}^{II}$ de orden

$q v^2$ y una submatriz $\Omega_{\alpha\mu}^{II}$ de orden $p'v^2 \times qv^2$ con lo que se puede definir una submatriz cuadrada Ω^{II} , de Ω , de orden $p'v^2 + qv^2$ y no singular

$$\Omega^{II} = \begin{array}{c} \begin{array}{cc} p'v^2 & qv^2 \\ \hline \Omega_{\alpha\alpha} & \Omega_{\alpha\mu}^{II} \\ \hline (\Omega_{\alpha\mu}^{II})' & \Omega_{\mu\mu}^{II} \end{array} \begin{array}{l} p'v^2 \\ qv^2 \end{array} \end{array}$$

luego

$$\text{rango } \Omega^{II} = p'v^2 + qv^2 - sv^2$$

ahora

$$\begin{aligned} \text{rango } \Omega &\geq \max(p'v^2 + qv^2 - rv^2, p'v^2 + qv^2 - sv^2) = (p'v^2 + qv^2) - \min(rv^2, sv^2) = \\ &= ((p'+q) - \min(r,s)) v^2 \end{aligned}$$

Si tenemos en cuenta las ecuaciones vectoriales (3.1.6.9)

$$\bar{\beta}' \Omega = \bar{0}$$

en ambas ecuaciones, al menos rv^2 de $\bar{\beta}'_1$ ó sv^2 de $\bar{\beta}'_2$ son autovectores asociados a rv^2 ó sv^2 autovalores nulos de la ecuación vectorial

$$\Omega \bar{\beta}_i = \lambda_i \bar{\beta}_i \quad \text{con} \quad \bar{\beta}_i \neq \bar{0}$$

de donde se deduce que el número de autovectores nulos es

$$\min(r,s)v^2$$

y aplicando una proposición conocida (Pollock(1979)) se deduce que

$$\text{rango } \Omega = ((p'+q) - \min(r,s)) v^2$$

como queríamos probar.

En este caso, en el que para obtener la hipótesis alternativa se hacen extensiones simultáneas en los polinomios $\Theta(z)$ y $\phi(z)$, en el estadístico S_c^* de (3.1.4.1), debe hacerse una elección apropiada de la matriz generalizada inversa $\hat{\Omega}^-$.

Ya hemos visto que existen submatrices Ω^I y Ω^{II} y en general Ω^* de la matriz Ω , de órdenes $(p+q+m)v^2$ con $m = \max(r,s)$, los cuales se obtienen eliminando $g = \min(r,s)v^2$ filas y columnas de Ω , entonces si permutamos las filas y columnas de Ω , esta matriz se puede poner en la forma

$$\Omega = \left[\begin{array}{c|c} \Omega^* & \Omega_2 \\ \hline \Omega_2' & \Omega_3 \end{array} \right] \begin{array}{l} (p+q+m)v^2 \\ (p'+q'-p-q-m)v^2 \end{array}$$

donde Ω^* es una matriz cuadrada no singular de rango $(p+q+m)v^2$ y de forma tal que $\Omega_3 = \Omega_2' \Omega^{*-1} \Omega_2$

con lo que una elección de la matriz generalizada inversa es

$$\Omega^- = \left[\begin{array}{c|c} (\Omega^*)^{-1} & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array} \right]$$

Así, en nuestro caso para elegir Ω^* se seleccionarían las filas y columnas de Ω correspondientes a un modelo alternativo MARMA $(p+r, q)$ cuando $r \geq s$ ó a un modelo MARMA $(p, q+s)$ cuando $s \geq r$.

3.1.8. Conclusiones.-

1a) La invarianza del estadístico S_c^* respecto a la elección de $\hat{\Omega}^-$ ya ha sido vista en el capítulo anterior; con lo que se llega a idénticas conclusiones en el sentido de que el "score test" para contrastar la hipótesis nula de que el modelo es MARMA (p, q) frente a la alternativa MARMA $(p+m, q)$ es asintóticamente equivalente al que resulta de contrastar la misma hipótesis nula frente a la alternativa MARMA $(p, q+m)$. (resultado del Teorema 2(2.5.6)).

2a) Teniendo en cuenta lo anterior, como el estadístico S no está afectado por la estructura del modelo alternativo, para estudiar la potencia del test, se hará usando la distribución de S_c^* bajo la sucesión Pitman de hipótesis alternativas, donde se hace uso del teorema enunciado en (3.1.4), (Kohn, 1979).

Dado m , la potencia del test dependerá solamente de

$$K_0 = \bar{\eta}' H (H' \Omega^- H)^{-1} H' \bar{\eta}$$

cuyo valor resulta al contrastar una hipótesis nula frente a cualquier hipótesis alternativa identificada implícitamente en la elección de la submatriz Ω^* de Ω , valor que sin embargo, es igual al que se obtiene cuando contrastamos H_0 frente a una hipótesis alternativa no identificable como MARMA (p+r, q+s).

Manipulando Ω^* y H, podremos elegir valores de ξ y de los coeficientes de los polinomios matriciales adicionales en la hipótesis alternativa implícita por la elección de Ω^* de forma tal que el valor de k_0 sea igual, sea cual sea, la matriz Ω^* seleccionada.

Así definiremos una clase de hipótesis alternativas locales que convergen a la hipótesis nula, por lo que la falta de identificabilidad de la hipótesis MARMA (p+r, q+s) no debe dar como resultado una pérdida de potencia del "score test", en este caso.

Así cualquier hipótesis alternativa, será localmente equivalente a otra que pertenezca a la misma clase, es por ello que la hipótesis alternativa MARMA (p+r, q+s) puede escribirse en función de solo $(p+q+m)v^2$ parámetros en vez de $(p+q+r+s)v^2$.

3a) Como conclusión final la conducta del estadístico S_c^* depende solo de la hipótesis nula fijada y del valor de m y no de la forma particular de la hipótesis alternativa.

3.2.- ESTADISTICO PORTMANTEAU PARA MODELOS MARMA.-

Sea el modelo (3,1) , la idea fundamental ahora es contrastar mediante el "estadístico portmanteau" la hipótesis nula de que el modelo es MARMA(p , q) frente a la hipótesis alternativa de que el modelo es MARMA(p , q +s) , para con posterioridad demostrar que este estadístico es asintóticamente equivalente al hallado mediante el "score test" donde

$$S_c^* = n^{-1} \hat{s}' \hat{\Omega} - \hat{s} = n^{-1} \hat{\lambda}' \hat{H}' \hat{\Omega} - \hat{H} \hat{\lambda}$$

es en este caso función de los últimos v_m^2 , con $m=s$, elementos del vector de "scores", pertenecientes a los coeficientes extras del modelo alternativo, y a su matriz de covarianzas condicional $\Omega_{\mu_e \mu_e}$

3.2.1.- CALCULO DEL SCORE TEST PARA CONTRASTAR UN MODELO MARMA (p , q) FRENTE A UN MODELO MARMA (p , q +s).-

En este caso si tenemos en cuenta el resultado (3.1.55) de $\frac{\int \bar{a}_t}{\int \bar{\mu}}$

solo se necesita la parte correspondiente a las derivadas de \bar{a}_t respecto a las matrices $\Theta_{-q+1}, \dots, \Theta_{-q+s}$.

Definición.-

Sean $C_k = n^{-1} \sum_{t=1}^n \bar{a}_t \bar{a}_{t-k}'$ matrices cuadradas para $k = 0, 1, \dots$

y $V^{1/2}$ una matriz triángular inferior tal que

$$C_0 = V^{1/2} (V^{1/2})'$$

3.2.1.- Proposición.-

El subvector de "scores", $\bar{s}' (\bar{\Theta})$, es asintóticamente igual a $n D \cdot \bar{r}$

donde

$$\bar{r} = \text{vec}(R_1' : \dots : R_m') = \bar{\Psi}_m(z) \otimes \text{vec} R_0'$$

siendo

$$R_k = V^{-1/2} C_k (V^{-1/2})' \quad \text{para } k = 1, \dots, m$$

y donde

$$D = F_{mm} (z^q V^{1/2} \otimes (\Theta^{-1}(z))' (V^{-1/2})')$$

siendo $m = s$, en este caso, $y \frac{m}{n} \rightarrow 0$ cuando m y n tienden a infinito

Demostración.-

El subvector de scores es

$$\bar{s}'(\bar{\Theta}^*) = \sum_{t=1}^n \underline{D}' V^{-1} \bar{a}_t$$

donde

$$\underline{D}' = \frac{\int \bar{a}_t}{\int \bar{\Theta}^{*'}} = -\bar{\Psi}'_{q,m}(z) \otimes \bar{a}_t' \otimes \Theta^{-1}(z)$$

siendo

$$-\bar{\Psi}'_{q,m}(z) = z^q \bar{\Psi}'_m(z) = (z^{q+1}, \dots, z^{q+m})$$

y

$$\bar{\Theta}^{*'} = \text{vec}(\underline{\Theta}_{q+1} \vdots \dots \vdots \underline{\Theta}_{q+m})$$

luego

$$\bar{s}'(\bar{\Theta}^*) = \sum_{t=1}^n (-\bar{\Psi}'_{q,m}(z) \otimes \bar{a}_t' \otimes (\Theta^{-1}(z))')$$

En todo lo anterior y en lo que sigue se supone que el vector \bar{a}_t se calcula supuesta la hipótesis nula cierta, es decir el modelo será

$$\begin{aligned} \text{con} \quad \hat{\Phi}_p(B) \hat{X}_t &= \Theta_q(B) \hat{a}_t \\ \sum_{k=q+1}^m \Theta_k \bar{a}_{t-k} &= \bar{0} \end{aligned}$$

ahora como

$$V = \frac{\sum \hat{a}_t \hat{a}_t'}{n} = C_0 = V^{1/2} (V^{1/2})'$$

se sigue que

$$(3.2.1.1) \quad \bar{s}' = \sum_{t=1}^n (-\bar{\Psi}'_{q,m}(z) \otimes \bar{a}_t' \otimes (\Theta^{-1}(z))') (V^{-1/2})' V^{-1/2} \bar{a}_t$$

ahora como, teniendo en cuenta la definición de F_{mm} (Apéndice, A.1.1)

$$\begin{aligned} D &= F_{mm} (z^q V^{1/2} \otimes (\Theta^{-1}(z))' (V^{-1/2})') = \\ &= \bar{\Psi}'_m \bar{\Psi}_m^* \otimes z^q V^{1/2} \otimes (\Theta^{-1}(z))' (V^{-1/2})' \end{aligned}$$

se tiene que

$$\begin{aligned}
 D\vec{r} &= (\vec{\varphi}_m \vec{\varphi}_m^* \otimes z^q V^{1/2} \otimes (\theta^{-1}(z))' (V^{-1/2})') (\vec{\varphi}_m(z) \otimes \text{vec } R_o') = \\
 &= (z^q \vec{\varphi}_m \vec{\varphi}_m^* \otimes V^{1/2} \otimes (\theta^{-1}(z))' (V^{-1/2})') (\vec{\varphi}_m \otimes \text{vec } R_o') = \\
 &= z^q \vec{\varphi}_m \vec{\varphi}_m^* \vec{\varphi}_m \otimes [(V^{1/2} \otimes (\theta^{-1}(z))' (V^{-1/2})') \text{vec } R_o'] = \\
 &= m \vec{\varphi}_{q,m} \otimes \text{vec} ((\theta^{-1}(z))' (V^{-1/2})' R_o' (V^{1/2})') = \\
 &= m \vec{\varphi}_{q,m} \otimes \text{vec} ((\theta^{-1}(z))' (V^{-1/2})' V^{-1/2} C_o (V^{-1/2})' V^{1/2}) = \\
 &= m \vec{\varphi}_{q,m} \otimes \text{vec} ((\theta^{-1}(z))' V^{-1} n^{-1} \sum_{t=1}^n \vec{a}_t \vec{a}_t') = \\
 &= \frac{m}{n} \sum_{t=1}^n \vec{\varphi}_{q,m} \otimes \text{vec} ((\theta^{-1}(z))' V^{-1} \vec{a}_t \vec{a}_t') = \\
 &= \frac{m}{n} \sum_{t=1}^n \vec{\varphi}_{q,m} \otimes [(\vec{a}_t \otimes (\theta^{-1}(z))') \text{vec} (V^{-1} \vec{a}_t)] = \\
 &= \frac{m}{n} \sum_{t=1}^n (\vec{\varphi}_{q,m}(z) \otimes \vec{a}_t \otimes (\theta^{-1}(z))') V^{-1} \vec{a}_t
 \end{aligned}$$

luego

$$n D\vec{r} = m \vec{s}' (\vec{\theta}^*)$$

asi cuando $n \rightarrow \infty$ y $\frac{m}{n} \rightarrow 0$ implica que $n D\vec{r}$ es asintoticamente igual

a \vec{S}_1 , luego : $n^{1/2} D\vec{r} = n^{-1/2} \vec{S}_1$ como queriamos demostrar

3.2.2.- Distribución asintótica del estadístico S_c^*

Al demostrar que \bar{s}^{-1} es asintóticamente igual a $n D\bar{r}$, aplicando el Teorema de Billingsley para martingalas (1968) se sigue que

como

$$n^{-1/2} \bar{s}^{-1} \xrightarrow{L} N(\bar{\sigma}, \Omega_{\mu_e \mu_e})$$

entonces también

$$n^{1/2} D\bar{r} \xrightarrow{L} N(\bar{\sigma}, \Omega_{\mu_e \mu_e})$$

luego

(3.2.2.1)
$$n^{1/2} \bar{r} \xrightarrow{L} N(\bar{\sigma}, D^{-1} \Omega_{\mu_e \mu_e} (D^{-1})')$$

teniendo en cuenta este resultado se sigue la siguiente

Proposición.-

La matriz de covarianzas de $n^{1/2} \bar{r}$ puede tomarse aproximadamente como la matriz idempotente

$$I_{v^2 m} - \hat{E}' (\hat{E} \hat{E}')^{-1} \hat{E}$$

siendo

$$E' = (E'_1 : E'_2)$$

donde

$$E_1 = F_{p, m} (-K(z) V^{1/2} \otimes (\theta^{-1}(z))' (V^{-1/2})')$$

y

$$E_2 = F_{q, m} (V^{1/2} \otimes (\theta^{-1}(z))' (V^{-1/2})')$$

Demostración.-

En este contraste de hipótesis la matriz de varianzas de los parámetros del modelo toma la forma particular

$$\Omega = \left[\begin{array}{c|c} F_{pp} (KVK' \otimes (\theta V \theta')^{-1}) & F_{pq} (KV \otimes (\theta V \theta')^{-1}) \\ \hline F_{pq}' (KV \otimes (\theta V \theta')^{-1}) & F_{q+m, q+m} (V \otimes (\theta V \theta')^{-1}) \end{array} \right]$$

donde

$$F_{q+m, q+m} = \left[\begin{array}{c|c} F_{qq} & F_{qm} \\ \hline F_{qm}' & F_{mm} \end{array} \right] \quad \text{pues } q' = q+m$$

Ahora la matriz inversa de Ω , en este caso existe, será

$$\Omega^{-1} = \left[\begin{array}{c|c} \text{---} & \text{---} \\ \hline & (-F_{pq}' F_{pp}^{-1} F_{pq} + F_{qq}') \end{array} \right]$$

pero al ser $\Omega_{\mu_e \mu_e}$ la matriz inversa de la submatriz de dimensión $m \times m$ inferior derecha de la inversa de Ω , se tiene que será una submatriz de

$$F_{qq}' - F_{pq}' F_{pp}^{-1} F_{pq}$$

Empezamos a desarrollar en primer lugar la matriz I_{V2m-P}

$$P = E'(EE')^{-1}E = E_1' \{E_1(I-P_2)E_1'\}^{-1} E_1(I-P_2) + E_2' \{E_2(I-P_1)E_2'\}^{-1} E_2(I-P_1) = P_{1/2} + P_{2/1}$$

donde

$$EE' = \left[\begin{array}{c|c} E_1 E_1' & E_1 E_2' \\ \hline E_2 E_1' & E_2 E_2' \end{array} \right]$$

siendo

$$E_1 E_1' = F_{pm}(\cdot) F_{qm}'(\cdot), \quad E_2 E_2' = F_{qm}(\cdot) F_{qm}'(\cdot)$$

$$P_1 = E_1' (E_1 E_1')^{-1} E_1 \quad \text{y} \quad P_2 = E_2' (E_2 E_2')^{-1} E_2$$

Ahora teniendo en cuenta las propiedades de la transformación F, dadas en el Apéndice se sigue que

$$(3.2.2.2) \quad E_1 E_1' = F_{pm}(-K(z)V^{1/2} \otimes (\Theta^{-1}(z))'(V^{-1/2})')$$

$$F_{mp}(-(V^{1/2})'K' \otimes V^{-1/2} \Theta^{-1}(z)) =$$

$$= m F_{pp}(K(z)VK'(z) \otimes (\Theta^{-1}(z))'V^{-1} \Theta^{-1}(z))$$

Análogamente

$$E_2 E_2' = (\bar{F}_q \bar{F}_m^* \otimes V^{1/2} \otimes (\Theta^{-1}(z))'(V^{-1/2})')$$

$$\cdot (\bar{F}_m \bar{F}_q^* \otimes (V^{1/2})' \otimes V^{-1/2} \Theta^{-1}(z))$$

luego

$$(3.2.2.3) \quad E_2 E_2' = m F_{q \ q} (V \otimes (\theta^{-1}(z))' V^{-1} \theta^{-1}(z))$$

$$E_1 E_2' = m F_{p \ q} (-K(z) V \otimes (\theta^{-1}(z))' V^{-1} \theta^{-1}(z))$$

y

$$E_2 E_1' = m F_{q \ p} (-V K'(z) \otimes (\theta^{-1}(z))' V^{-1} \theta^{-1}(z))$$

entonces

$$(3.2.2.4) \quad P_1 = E_1' (E_1 E_1')^{-1} E_1 = F_{m \ p} () \cdot \frac{1}{m} \cdot F_{p \ p} () \cdot F_{p \ m} ()$$

y aplicando las propiedades de la función F del Apéndice se obtiene que

$$P_1 = \frac{1}{m} (\bar{\psi}_m \bar{\psi}_p^* (\bar{\psi}_p \bar{\psi}_p^*)^{-1} \bar{\psi}_p \bar{\psi}_m^* \otimes I_v \otimes I_v) =$$

$$= m^{-1} (\bar{\psi}_{p \ m} \bar{\psi}_{p \ m}^* \otimes I_v \otimes I_v)$$

análogamente

$$(3.2.2.5) \quad P_2 = E_2' (E_2 E_2')^{-1} E_2 = F_{m \ q} () \frac{1}{m q^2} F_{q \ q} () F_{q \ m} ()$$

y de nuevo aplicando las propiedades de la función F se tiene que

$$P_2 = \frac{1}{m} (\bar{\psi}_m \bar{\psi}_q^* (\bar{\psi}_q \bar{\psi}_q^*)^{-1} \bar{\psi}_q \bar{\psi}_m^* \otimes I_v \otimes I_v) =$$

$$= m^{-1} (\bar{\psi}_{q \ m} \bar{\psi}_{q \ m}^* \otimes I_v \otimes I_v)$$

ahora como

$$(3.2.2.6) \quad I_{v^2 m} - E' (E E')^{-1} E = I_{v^2 m} - E_1' \{E_1 (I - P_2) E_1'\}^{-1} E_1 (I - P_2) +$$

$$+ E_2' \{E_2 (I - P_1) E_2'\}^{-1} E_2 (I - P_1) = I_{v^2 m} - E_1' (E_1 E_1' - E_1 P_2 E_1')^{-1}$$

$$(E_1 - E_1 P_2) + E_2' (E_2 E_2' - E_2 P_1 E_2')^{-1} (E_2 - E_2 P_1)$$

Si tenemos en cuenta los resultados anteriores (3.2.2.2) , (3.2.2.3)

(3.2.2.4) y (3.2.2.5), se verifica que

(3.2.2.7)

$$P_{1/2} = E_1' (E_1 (I - P_2) E_1')^{-1} E_1 (I - P_2) = m^{-1} (\vec{\varphi}_{p,m} \vec{\varphi}_{p,m}^* \otimes I_V \otimes I_V)$$

y análogamente

(3.2.2.8)

$$P_{2/1} = E_2' (E_2 (I - P_1) E_2')^{-1} E_2 (I - P_1) = m^{-1} (\vec{\varphi}_{q,m} \vec{\varphi}_{q,m}^* \otimes I_V \otimes I_V)$$

con las propiedades siguientes:

$$P_{1/2}^2 = P_{1/2} \quad , \quad P_{1/2} E_1' = E_1' \quad , \quad P_{1/2} E_2' = \underline{0} \quad , \quad P_{1/2} P_1 = P_1$$

$$P_{1/2} P_2 = \underline{0} \quad , \quad P_{1/2} P_{2/1} = \underline{0} \quad , \quad P_{1/2} = P_1 - P_1 P_{2/1} \quad \text{etc....}$$

y sustituyendo en (3.2.2.6) se tiene que

$$(3.2.2.9) \quad I - P = I - P_{1/2} - P_{2/1} = \frac{m}{m-2} (\vec{\varphi}_m \vec{\varphi}_m^* \otimes I_V \otimes I_V)$$

Por otra parte si desarrollamos la matriz de covarianzas de $n^{1/2} \hat{r}$, esto es, $D^{-1} \Omega_{\mu_e \mu_e} (D^{-1})'$, se tiene que como

$$\begin{aligned} D^{-1} \Omega_{\mu_e \mu_e} (D^{-1})' &= (\varphi_m \vec{\varphi}_m^* \otimes z^q V^{1/2} \otimes (\Theta^{-1})' (V^{-1/2})^{-1})^{-1} \Omega_{\mu_e \mu_e} \\ ((\vec{\varphi}_m \vec{\varphi}_m^*)' \otimes (z^*)^q (V^{1/2})' \otimes V^{-1/2} \Theta^{-1}) &= \text{Cte} (\vec{\varphi}_m \vec{\varphi}_m^* \otimes z^{-q} V^{-1/2} \otimes (V^{1/2})' \Theta) \\ \text{cte} (\vec{\varphi}_{qm} \vec{\varphi}_{qm}^* \otimes V \otimes (\Theta V \Theta')^{-1}) &\text{cte} (\vec{\varphi}_m \vec{\varphi}_m^* \otimes (z^*)^{-q} (V^{-1/2})' \otimes \Theta V^{1/2}) \end{aligned}$$

ya que la submatriz $\Omega_{\mu_e \mu_e}$ es una submatriz de

$$\vec{\varphi}_q \vec{\varphi}_q^* \otimes V \otimes (\Theta V \Theta')^{-1}$$

y

$$D^{-1} \Omega_{\mu_e \mu_e} (D^{-1})'$$

se puede poner en la forma

$$(3.2.2.10) \quad \text{Cte} (\vec{\varphi}_m \vec{\varphi}_m^* \otimes I_v \otimes I_v)$$

En todo lo anterior se hace uso de las propiedades del producto tensorial de matrices y de la aplicación $F_{p,q}$, y se obtiene que

$$D^{-1} \Omega_{\mu_e \mu_e} (D^{-1})' \approx I_{v^2 m} - E' (E E')^{-1} E$$

con lo que queda demostrada la proposición

Proposición 2.-

La matriz $I_{v^2 m} - E' (E E')^{-1} E$ es de rango $v^2 (m-p-q)$.

Demostración.-

Sea

$$E = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

luego

$$E' = \left[\begin{array}{c|c} p v^2 & q v^2 \\ \hline E_1' & E_2' \end{array} \right] m v^2$$

es una matriz particionada con dimensión $(\text{Ker } E') = 0$, al ser la aplicación lineal asociada, inyectiva, y tal que

$$\text{Imagen } E' = \text{Imagen } E_1' \oplus \text{Imagen } E_2'$$

y como

$$P = E' (E E')^{-1} E \quad \text{y} \quad P_2 = E_2' (E_2 E_2')^{-1} E_2$$

entonces

$$P_{1/2} = E_1' \{ E_1 (I - P_2) E_1' \}^{-1} E_1 (I - P_2)$$

es un proyector de imagen (E') sobre imagen (E_1') a lo largo de imagen (E_2') y si existe matriz inversa de $E_1 (I - P_2) E_1'$ entonces el proyector P_2 es ortogonal.

Lo anterior significa que $\forall \bar{p}_1 \in \text{Imagen } (E'_1)$ se verifica que

$$E'_1 \{ E_1 (I - P_2) E'_1 \}^{-1} E_1 (I - P_2) \bar{p}_1 = \bar{p}_1 \quad \text{y}$$

$$E'_1 \{ E_1 (I - P_2) E'_1 \}^{-1} E_1 (I - P_2) \bar{p}_2 = \bar{o}$$

Analogamente

$$P_{2/1} = E'_2 \{ E_2 (I - P_1) E'_2 \}^{-1} E_2 (I - P_1)$$

es un proyector de Imagen (E'_1) sobre Imagen (E'_2) a lo largo de Imagen (E'_1) luego $\forall \bar{p}_2 \in \text{Imagen } (E'_2)$ se verifica que

$$E'_2 \{ E_2 (I - P_1) E'_2 \}^{-1} E_2 (I - P_1) \bar{p}_2 = \bar{p}_2$$

mientras que

$$E'_2 \{ E_2 (I - P_1) E'_2 \}^{-1} E_2 (I - P_1) \bar{p}_1 = \bar{o}$$

por lo anterior se tiene que

$$\forall \bar{p} \in \text{Imagen } (E') = \text{Imagen } (E'_1) + \text{Imagen } (E'_2) = \bar{p}_1 + \bar{p}_2$$

$$(3.2.2.11) \{ I - E' (E E')^{-1} E \} \bar{p}_1 = \bar{p}_1 - \{ E'_1 (E_1 (I - P_2) E'_1)^{-1} E_1 (I - P_2) \} \bar{p}_1 - \\ - \{ E'_2 (E_2 (I - P_1) E'_2)^{-1} E_2 (I - P_1) \} \bar{p}_1 = \bar{p}_1 - \bar{p}_1 - \bar{o} = \bar{o}$$

$$\text{análogamente } (3.2.2.12) \{ I - E' (E E')^{-1} E \} \bar{p}_2 = \bar{p}_2 - \bar{o} - \bar{p}_2 = \bar{o}$$

En la ecuación vectorial (3.2.2.11) hay $p v^2$, la dimensión de E'_1 es $m v^2 \times p v^2$, autovectores asociados a $p v^2$ autovalores nulos.

En la ecuación vectorial (3.2.2.12) hay $q v^2$ autovectores asociados a otros tantos autovalores nulos, ya que la dimensión de E'_2 es $m v^2 \times q v^2$.

De ambas ecuaciones vectoriales se deduce que

$$\text{rango}(I - E' (E E')^{-1} E) = v^2 m - p v^2 - q v^2$$

como queríamos probar.

-Consecuencia de las proposiciones 1 y 2.

Teniendo en cuenta que

$$D^{-1} \Omega_{\mu_e \mu_e} (D')^{-1} \simeq I_{v^2 m} - E' (E E')^{-1} E$$

siendo el rango de las matrices igual a

$$v^2 m - v^2 p - v^2 q$$

y de que

$$n^{1/2} \hat{r} \xrightarrow{L} N(0, D^{-1} \Omega_{\mu_e \mu_e} (D'')^{-1})$$

como también

$$n^{1/2} \hat{r} \xrightarrow[n \rightarrow \infty]{D} n^{-1/2} \hat{s}$$

se tiene que

$$(3.2.2.13) \quad S = n^{-1} \hat{s}' \Omega^{-1} \hat{s} \xrightarrow[n \rightarrow \infty]{nr'} n \hat{r}' (I_{v2m} - E' (EE')^{-1} E) \hat{r}$$

luego se distribuye asintóticamente según una $\chi^2_{v2(m-p-q)}$

3.2.3.- Equivalencia asintótica de los estadísticos S_c^* y Q .

Proposición.-

Como consecuencia de la idempotencia de la matriz

$$I_{v2m} - E' (EE')^{-1} E$$

se deduce que una elección válida de la matriz generalizada inversa de esta matriz vendrá dada por la matriz identidad.

Demostración.-

Sea

$$P = I_{v2m} - E' (EE')^{-1} E$$

de forma tal que

$$P = P' = P^2$$

como además $E' (EE')^{-1} = E^+$

es una matriz inversa Moore-Penrose (Pollock, 1980), de forma tal que

$$EE^+ = I$$

y que

$$E^+ E = E' (EE')^{-1} E$$

es simétrica.

Si definimos $P^- = (I - E^+ E)^-$ de forma tal que $PP^-P = P$, luego una elección válida para P^- es la matriz identidad, pues $PIP = P^2 = P$, como queríamos probar,

Consecuencias.-

1) Al considerar

$$(I_{V2m} - E' (EE')^{-1} E)^- = I_{V2m}$$

se tiene que cuando $n \rightarrow \infty$ el "estadístico portmanteau"

$$S_c^* \xrightarrow{n \rightarrow \infty} n \hat{r}' I_{V2m}^{-1} \hat{r} = n \hat{r}' \hat{r} = Q$$

con lo que se obtiene una versión multivariante de éste estadístico

$$Q = n \sum_{k=1}^m \sum_{i=1}^y \sum_{j=1}^y \{ \hat{r}_{ij}(k) \}^2$$

2) Otra consecuencia de lo anterior, es que para $m \geq 1$ y finito un test "portmanteau" basado en las autocorrelaciones y correlaciones cruzadas residuales será igual a un "score test" con los mismos grados de libertad, cuando la hipótesis alternativa sea $MA(q+m)$, ya que en este caso $D = I_{mV2}$ con lo que

$$Q = S_c^* = n \hat{r}' \hat{\Omega}^{-1} \mu_e \mu_e' \hat{r}$$

3.3.- CALCULO PRACTICO DEL ESTADISTICO S_c^* DEL " SCORE TEST"

3.3.1.- Construcción de la regresión auxiliar.

Con vistas a sistematizar el proceso de cálculo, enunciamos la siguiente proposición:

Proposición.-

El estadístico S_c^* es la suma de cuadrados explicada por una regresión generalizada mínimo cuadrática de los residuos sobre las derivadas estimadas.

Demostración.-

Sea

$$\vec{w}' = (\vec{a}_1' : \dots : \vec{a}_n')_{1 \times nv} \quad \text{y} \quad D_t' = \frac{\partial \vec{a}_t'}{\partial \theta} \quad \text{la}$$

matriz de derivadas respecto a los parámetros del modelo donde se puede particionar dicha matriz en la forma

$$D_t = (D_{\alpha_t} : D_{\mu_t})_{v \times m}$$

donde las matrices

$$D_{\alpha_t} = \frac{\partial \vec{a}_t}{\partial \alpha} \quad \text{y} \quad D_{\mu_t} = \frac{\partial \vec{a}_t}{\partial \mu}$$

son de dimensiones

$$v \times p' v^2 \quad \text{y} \quad v \times q' v^2$$

Partimos entonces de las ecuaciones simultáneas de la regresión particular

$$(3.3.1.1) \quad \hat{a}_t = \hat{V}^{-1} \hat{D}_t \bar{\theta} + \bar{\epsilon}_t \quad \forall t=1, \dots, n$$

que puestas en forma matricial se tiene

$$(3.3.1.2) \quad \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_n \end{pmatrix}_{nv \times 1} = \begin{pmatrix} \hat{V}^{-1} \hat{D}_1 \\ \vdots \\ \hat{V}^{-1} \hat{D}_n \end{pmatrix} \bar{\theta} + \begin{pmatrix} \bar{\epsilon}_1 \\ \vdots \\ \bar{\epsilon}_n \end{pmatrix}_{nv \times 1}$$

o en forma vectorial compacta

$$(3.3.1.3) \quad \hat{w} = \hat{X} \bar{\theta} + \bar{e} = (I_n \otimes \hat{V}^{-1}) \hat{M} \bar{\theta} + \bar{e}$$

Ahora consideramos que

$$E(\bar{e}) = \bar{0}, \quad \text{que } \text{Var}(\bar{e}) = E(\bar{e} \bar{e}') = I_n \otimes \hat{V}^{-1}$$

y que la matriz de sensibilidad del modelo, $M' (I_n \otimes \hat{V}^{-1}) M$, es regular, para lo que elegimos adecuadamente las matrices D_t de la matriz $M_{nv \times m}$,

Un estimador eficiente de $\bar{\theta}$ viene dado por

$$(3.3.1.4) \quad \hat{\bar{\theta}} = \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{w} = \{ \hat{X}' (\text{Var}(\bar{e}))^{-1} \hat{X} \}^{-1} \hat{X}' \hat{w}$$

y la parte de varianza explicada por la regresión generalizada mínimo cuadrática viene dada por

$$S = \hat{\theta}' \hat{X}' (\text{Var}(\hat{e}))^{-1} (\text{Var}(\hat{e})) \hat{X} \hat{\theta} = \hat{w}' (I_n \otimes \hat{V}^{-1}) \hat{M} \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}'$$

$$(I_n \otimes \hat{V}^{-1}) (I_n \otimes \hat{V}) (I_n \otimes \hat{V}) (I_n \otimes \hat{V}^{-1}) \hat{M} \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{w} =$$

$$= \hat{w}' (I_n \otimes \hat{V}^{-1}) \hat{M} \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}' \hat{M} \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{w} =$$

$$= \hat{w}' (I_n \otimes \hat{V}^{-1}) \hat{M} \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{w}$$

puesto que la matriz

$$\hat{M} \{ \hat{M}' (I_n \otimes \hat{V}^{-1}) \hat{M} \}^{-1} \hat{M}'$$

es una matriz idempotente

Si definimos a continuación

$$\hat{w} = \bar{\psi}_n \otimes \hat{a}_t$$

luego

$$\hat{w}' = \bar{\psi}_n^* \otimes \hat{a}_t'$$

$$\hat{M} = \bar{\psi}_n \otimes \hat{D}_t$$

y entonces

$$M' = \bar{\psi}_n^* \otimes \hat{D}_t'$$

podemos desarrollar tensorialmente S, en la forma siguiente:

$$S = (\bar{\psi}_n^* \otimes \hat{a}_t') (I_n \otimes \hat{V}^{-1}) (\bar{\psi}_n \otimes \hat{D}_t) \{ (\bar{\psi}_n^* \otimes \hat{D}_t') (I_n \otimes \hat{V}^{-1}) (\bar{\psi}_n \otimes \hat{D}_t) \}^{-1}$$

$$\cdot (\bar{\psi}_n^* \otimes \hat{D}_t') (I_n \otimes \hat{V}^{-1}) (\bar{\psi}_n \otimes \hat{a}_t) = (\bar{\psi}_n^* \bar{\psi}_n \otimes \hat{a}_t' \hat{V}^{-1} \hat{D}_t)$$

$$(\bar{\psi}_n^* \bar{\psi}_n \otimes \hat{D}_t' \hat{V}^{-1} \hat{D}_t)^{-1} (\bar{\psi}_n^* \bar{\psi}_n \otimes \hat{D}_t' \hat{V}^{-1} \hat{a}_t) =$$

$$= \sum_{t=1}^n \hat{a}_t' \hat{V}^{-1} \hat{D}_t \left(\frac{1}{n} \sum_{t=1}^n \hat{D}_t' \hat{V}^{-1} \hat{D}_t \right)^{-1} \cdot n^{-1} \sum_{t=1}^n \hat{D}_t' \hat{V}^{-1} \hat{a}_t =$$

$$= \bar{s}' (\hat{\theta})_n^{-1} \hat{\Omega}^{-1} \quad \bar{s} (\hat{\theta}) = S_c^*$$

c. q. d

APENDICE

3.4.- APENDICE.

A.1.- Transformación F.

Sea $M(z)$ el cuerpo de las series de potencias, con coeficientes matrices reales de dimensión $u \times v$, esto es

$$M(z) = \left\{ \Gamma_{\eta\nu}(z) = \sum_{k=-\infty}^{\infty} \Gamma_k z^k \mid \Gamma_k \in \mathcal{L}_{u \times v} \right\}$$

con
$$\left| \sum_{k=-\infty}^{\infty} \det \Gamma_k \right| < \infty$$

y con
$$\Gamma_{-k} = \Gamma_k' \quad \text{si } u=v$$

y sea
$$M^+(z) = \left\{ \Gamma(z) \in M(z) \mid \Gamma_k = \underline{0} \quad \forall k < 0 \right\}$$

A.1.1.- Definición.-

La transformación $F_{m \times n}$ es una aplicación de la forma

$$F_{m,n}: M(z) \xrightarrow{\Gamma_{\eta\nu}(z)} \beta$$

$$F_{m,n}(\Gamma_{\eta\nu}(z)) = P$$

donde $P_{m \times n}$ es una matriz cuyo elemento i, j viene dado por la matriz que es el coeficiente de z^{i-j} y que pertenece a un conjunto de matrices banda.

En nuestro caso teniendo en cuenta la definición de $\Gamma_{\eta\nu}(z)$

de (3.1.6.1) se sigue que

$$F_{p,q}(\Gamma_{\eta\nu}(z)) = F_{p,q} \left(\sum_{k=-\infty}^{\infty} \Gamma_k z^k \right) = F_{p,q} \left(\sum_{k=-\infty}^{\infty} \delta_{\eta\nu}(k) z^k \right) =$$

$$= \begin{bmatrix} \delta_{\eta\nu}(0) & \dots & \delta_{\eta\nu}(q-1) \\ \vdots & & \vdots \\ \delta_{\eta\nu}(p-1) & \dots & \delta_{\eta\nu}(p-q) \end{bmatrix} = (\delta_{\eta\nu l}(i-j)) = \bar{\varphi}_p(z) \bar{\varphi}_q^*(z) \otimes \delta_{\eta\nu}(k)$$

donde

$$\delta_{\eta\nu l}(i-j)$$

es el elemento de la matriz que esta en la fila $(i-1)u+h$ y la columna

$(j-1)v+1$, para $i = 1, \dots, p$; $j = 1, \dots, q$; $h = 1, \dots, u$ y $l = 1, \dots, v$

siendo $\bar{\varphi}_q^*(z)$ la conjugada traspuesta de $\bar{\varphi}_q(z)$, donde $z = e^{i\omega}$,

luego

$$F_{p,q}(\Gamma_{\eta\nu}(z)) = \bar{\Psi}_p(z) \bar{\Psi}_q^*(z) \otimes \Gamma_{\eta\nu}(z)$$

A.2.-Propiedades de la transformación F.-

A.2.1.-Propiedad 1ª :

$$F_{p,q}(\Gamma_{\nu\nu}(z^{-1})) = F'_{q,p}(\Gamma_{\nu\nu}(z)) \quad \forall \Gamma_{\nu\nu}(z) \in M(z)$$

Demostración:

$$\begin{aligned} F_{p,q}(\Gamma_{\nu\nu}(z^{-1})) &= F_{p,q}\left(\sum_{k=-\infty}^{\infty} \Gamma_k z^{-k}\right) = F_{p,q}\left(\sum_{k=-\infty}^{\infty} \delta_{\nu\nu}(k) z^{-k}\right) = \\ &= F_{p,q}\left(\sum_{k=-\infty}^{\infty} \delta_{\nu\nu}(-k) z^k\right) = \bar{\Psi}_p \bar{\Psi}_q^* \otimes \delta_{\nu\nu}(-k) \end{aligned}$$

Ahora teniendo en cuenta que

$$\delta_{\nu\nu}(-k) = \delta'_{\nu\nu}(k)$$

se tiene que

$$F_{p,q}(\Gamma_{\nu\nu}(z^{-1})) = (\bar{\Psi}_q \bar{\Psi}_p^*)' \otimes \delta'_{\nu\nu}(k)$$

$$= (\bar{\Psi}_q \bar{\Psi}_p^* \otimes \delta_{\nu\nu}(k))' = F'_{q,p}(\Gamma_{\nu\nu}(z)) \quad \text{con lo que queda probada}$$

A.2.2.-Propiedad 2ª

Sean

$$\Gamma_{\eta\nu}(z) \quad \text{y} \quad \Lambda_{\nu\beta}(z) \in M^+(z)$$

si

$$F_{pp}(\Gamma_{\eta\nu}(z)) = \bar{\Psi}_p \bar{\Psi}_p^* \otimes \delta_{\eta\nu}(k)$$

y

$$F_{pq}(\Lambda_{\nu\beta}(z)) = \bar{\Psi}_p \bar{\Psi}_q^* \otimes \tau_{\nu\beta}(k)$$

se tiene que

$$F_{pq}(\Gamma_{\eta\nu}(z) \Lambda_{\nu\beta}(z)) = \frac{1}{p} F_{pp}(\Gamma_{\eta\nu}(z)) F_{pq}(\Lambda_{\nu\beta}(z))$$

Demostración:

$$\begin{aligned} \mathbb{F}_{pq}(\Gamma_{\eta\nu}(z) \wedge \nu_{\beta}(z)) &= \mathbb{F}_{pq} \left(\sum_{t=0}^{\infty} \Gamma_t z^t \sum_{s=0}^{\infty} \Lambda_s z^s \right) = \\ &= \mathbb{F}_{pq} \left(\sum_{t+s=0}^{\infty} \sum_{t=0}^{t+s} \Gamma_t \Lambda_s z^{s+t} \right) = \mathbb{F}_{pq} \left(\sum_{t+s} \sum_t \delta_{\eta\nu}(t) \nu_{\beta}(s) z^{t+s} \right) = \\ &= \mathbb{F}_{pq} \left(\sum_{\omega=0}^{\infty} \sum_{t=0}^{\omega} \delta_{\eta\nu}(t) \nu_{\beta}(\omega-t) z^{\omega} \right) = \mathbb{F}_{pq} \left(\sum_{\omega=0}^{\infty} \psi_{\eta\beta}(\omega) z^{\omega} \right) = \\ &= \mathbb{F}_{pq} \left(\sum_{\omega=0}^{\infty} \Delta_{\omega} z^{\omega} \right) = \mathbb{F}_{pq}(\Delta_{\eta\beta}(z)) \end{aligned}$$

donde

$$w = t+s$$

y

$$\psi_{\eta\beta}(w) = \sum_{t=0}^{\infty} \delta_{\eta\nu}(t) \nu_{\beta}(w-t)$$

luego

$$\begin{aligned} \mathbb{F}_{pq}(\Gamma_{\eta\nu}(z) \wedge \nu_{\beta}(z)) &= \vec{\psi}_p \vec{\psi}_q^* \otimes \delta_{\eta\nu}(k) \tau_{\nu\beta}(k) = \\ &= \frac{1}{p} \vec{\psi}_p \vec{\psi}_p^* \vec{\psi}_p \vec{\psi}_q^* \otimes \delta_{\eta\nu}(k) \tau_{\nu\beta}(k) = \frac{1}{p} (\vec{\psi}_p \vec{\psi}_p^* \otimes \delta_{\eta\nu}(k)) \\ &\quad (\vec{\psi}_p \vec{\psi}_q^* \otimes \tau_{\nu\beta}(k)) = \frac{1}{p} \mathbb{F}_{pp}(\Gamma_{\eta\nu}(z)) \mathbb{F}_{pq}(\Lambda_{\nu\beta}(z)) \end{aligned}$$

puesto que

$$\begin{aligned} \vec{\psi}_p^* \vec{\psi}_p &= (\bar{z}, \bar{z}^2, \dots, \bar{z}^p) (z, z^2, \dots, z^p)' = \\ &= z\bar{z} + \dots + \bar{z}^p z^p = p \end{aligned}$$

ya que

$$z\bar{z} = 1 \quad \text{como queriamos demostrar}$$

A.2.3.- Propiedad 3a :

Para todo $\Gamma_{\eta\eta}(z) \in M^+(z)$ y tal que $\Gamma_{\eta\eta}^{-1}(z) \in M^+(z)$

se verifica para cada p , que

$$\mathbb{F}_{pp}(\Gamma_{\eta\eta}^{-1}(z)) = p^2 \mathbb{F}_{pp}^{-1}(\Gamma_{\eta\eta}(z))$$

Demostración:

$$\begin{aligned} F_{PP}(\Gamma_{\eta\eta}^{-1}(z)) &= \bar{\psi}_P \bar{\psi}_P^* \otimes \delta_{\eta\eta}^{-1}(K) = p^2 (\bar{\psi}_P \bar{\psi}_P^*)^{-} \otimes \delta_{\eta\eta}^{-1}(K) = \\ &= p^2 (\bar{\psi}_P \bar{\psi}_P^* \otimes \delta_{\eta\eta}(K))^{-} = p^2 F_{PP}^{-}(\Gamma_{\eta\eta}(z)) \end{aligned}$$

Lo anterior es debido a que la matriz $(\bar{\psi}_P \bar{\psi}_P^*)$ es de la forma

$$(\bar{\psi}_P \bar{\psi}_P^*) = (z, z^2, \dots, z^P)' (\bar{z}, \bar{z}^2, \dots, \bar{z}^P) = \begin{bmatrix} 1 & \bar{z} & \bar{z}^2 & \dots & \bar{z}^{P-1} \\ z & 1 & & & \\ & & & & \\ & & & & \\ z^{P-1} & & & & z & 1 \end{bmatrix}$$

matriz no invertible y tal que

$$(\bar{\psi}_P \bar{\psi}_P^*)^2 = p (\bar{\psi}_P \bar{\psi}_P^*)$$

y que

$$(\bar{\psi}_P \bar{\psi}_P^*)^* = \bar{\psi}_P \bar{\psi}_P^*$$

siendo

$$(\bar{\psi}_P \bar{\psi}_P^*)^*$$

la matriz conjugada traspuesta de $(\bar{\psi}_P \bar{\psi}_P^*)$

luego

$$(\bar{\psi}_P \bar{\psi}_P^*)$$

es una matriz hermitica.

Si llamamos $A = (\bar{\psi}_P \bar{\psi}_P^*)$, se tiene que $A^2 = pA$ y además $A = A^*$, luego $A^* A = pA$ y multiplicando a la izquierda por A se tiene que

$$A A^* A = p A^2 = p^2 A$$

luego

$$A \frac{A^*}{p^2} A = A$$

luego

$$\frac{A^*}{p^2}$$

es una matriz generalizada inversa de A , luego queda demostrada la propiedad.

A.2.4-Propiedad 4a :

Para todo $\Gamma_{\alpha\beta}(z)$ y $\Lambda_{\alpha\beta}(z) \in M^+(z)$

si definimos

$$F_{q\infty}(\Gamma_{\alpha\beta}(z)) = \lim_{r \rightarrow \infty} F_{qr}(\Gamma_{\alpha\beta}(z))$$

y

$$F_{\infty p}(\Lambda_{\alpha\beta}(z)) = \lim_{r \rightarrow \infty} F_{rp}(\Lambda_{\alpha\beta}(z))$$

entonces

$$\begin{aligned} F_{qp}(\Gamma_{\alpha\beta}(z^{-1}) \Lambda_{\alpha\beta}(z)) &= F_{q\infty}(\Gamma_{\alpha\beta}(z^{-1})) F_{\infty p}(\Lambda_{\alpha\beta}(z)) = \\ &= \lim_{r \rightarrow \infty} \frac{1}{r} (F_{qr}(\Gamma_{\alpha\beta}(z^{-1})) F_{rp}(\Lambda_{\alpha\beta}(z))) \end{aligned}$$

Demostración:

Teniendo en cuenta la propiedad 2^a se verifica que:

$$\begin{aligned} F_{qp}(\Gamma_{\alpha\beta}(z^{-1}) \Lambda_{\alpha\beta}(z)) &= \frac{1}{q} F_{qq}(\Gamma_{\alpha\beta}(z^{-1})) F_{qp}(\Lambda_{\alpha\beta}(z)) = \\ &= \frac{1}{q} (\vec{\Psi}_q \vec{\Psi}_q^* \otimes \gamma'_{\alpha\beta}(k)) (\vec{\Psi}_q \vec{\Psi}_p^* \otimes \tau_{\alpha\beta}(k)) = \\ &= \frac{1}{q} \vec{\Psi}_q \vec{\Psi}_q^* \vec{\Psi}_p \vec{\Psi}_p^* \otimes \gamma'_{\alpha\beta}(k) \tau_{\alpha\beta}(k) = \vec{\Psi}_q \vec{\Psi}_p^* \otimes \gamma'_{\alpha\beta}(k) \tau_{\alpha\beta}(k) = \\ &= \lim_{r \rightarrow \infty} \frac{1}{r} \vec{\Psi}_q \vec{\Psi}_r^* \vec{\Psi}_r \vec{\Psi}_p^* \otimes \gamma'_{\alpha\beta}(k) \tau_{\alpha\beta}(k) = \\ &= \lim_{r \rightarrow \infty} \frac{1}{r} (\vec{\Psi}_q \vec{\Psi}_r^* \otimes \gamma'_{\alpha\beta}(k)) (\vec{\Psi}_r \vec{\Psi}_p^* \otimes \tau_{\alpha\beta}(k)) = \\ &= \lim_{r \rightarrow \infty} \frac{1}{r} (F'_{rq}(\Gamma_{\alpha\beta}(z)) F_{rp}(\Lambda_{\alpha\beta}(z))) = \\ &= \lim_{r \rightarrow \infty} \frac{1}{r} (F_{qr}(\Gamma_{\alpha\beta}(z^{-1})) F_{rp}(\Lambda_{\alpha\beta}(z))) \end{aligned}$$

Donde hemos aplicado el resultado obtenido en la propiedad 1^a, con lo cual queda terminada la demostración.

A.3.- Algunos resultados de álgebra tensorial

A.3.1.-

Sea $V = \frac{1}{n} \sum_{t=1}^n \vec{a}_t \vec{a}_t'$ la matriz definida en (3.2.1), entonces

$$\begin{aligned} \frac{\partial V^c}{\partial \theta} &= \frac{1}{n} \frac{\partial \sum (\vec{a}_t' \vec{a}_t)'}{\partial \theta} = \frac{1}{n} \sum ((\vec{a}_t')' \otimes I_V) \frac{\partial \vec{a}_t}{\partial \theta} + \frac{1}{n} \sum (\vec{a}_t \otimes \vec{a}_t) \frac{\partial 1}{\partial \theta} + \\ &+ \frac{1}{n} \sum (I \otimes \vec{a}_t) \frac{\partial \vec{a}_t}{\partial \theta} = \frac{1}{n} \sum (\vec{a}_t \otimes I_V) D_t + \frac{1}{n} \sum (I_V \otimes \vec{a}_t) D_t \end{aligned}$$

puesto que el segundo sumando es el vector nulo.

A.3.2.-

Si partimos de que $\sum_{t=1}^n \vec{a}_t' \vec{a}_t$ es una constante, derivando

esta igualdad respecto de θ se obtiene:

$$\begin{aligned} \frac{\partial \sum_{t=1}^n \vec{a}_t' \vec{a}_t}{\partial \theta} &= \frac{\partial \sum_{t=1}^n \vec{a}_t' I_V \vec{a}_t}{\partial \theta} = \sum ((I_V \vec{a}_t)' \otimes 1) \frac{\partial (\vec{a}_t')^c}{\partial \theta} + \\ &+ \sum (\vec{a}_t' \otimes \vec{a}_t') \frac{\partial I_V^c}{\partial \theta} + \sum (I_V \otimes \vec{a}_t') \frac{\partial \vec{a}_t}{\partial \theta} = \vec{0}_{1 \times m_1} \end{aligned}$$

pero como el segundo sumando es la matriz nula, se tiene que :

$$\sum (\vec{a}_t \otimes 1) \frac{\partial \vec{a}_t}{\partial \theta} + \sum (1 \otimes \vec{a}_t') \frac{\partial \vec{a}_t}{\partial \theta} = \vec{0}$$

luego

$$\sum \vec{a}_t' D_t + \sum \vec{a}_t' D_t = \vec{0}, \text{ luego } \sum \vec{a}_t' D_t = \vec{0}_{1 \times m}$$

y de aquí se deduce que

$$\sum_{t=1}^n D_t' \vec{a}_t = \vec{0}_{m_1 \times 1}$$

BIBLIOGRAFIA

ANDERSON, T.W. (1971). " The Statistical Analysis of Time Series"

Wiley: New York.

AITCHISON, J. and SILVEY, S.D. (1960). " Maximum likelihood estimation procedures and associated tests of significance". J. Royal Statistics. Soc. B, 22, 154-171.

BASILEVSKY, A. (1983). " Applied Matrix Algebra in the Statistical Sciences".

North-Holland; New York - Amsterdam.

BILLINGSLEY, P. (1961). " The Levy-Lindeberg Theorem for Martingales".

Proc. Amer. Math. Soc. 12, 788-792.

(1974) " Statistical Inference for Markov Processes".

Statistical Research Monographs, Vol. II. University of Chicago.

(1968). " Convergence of Probability Measure". Wiley:

New York.

BIRKHOFF, G., and MAC LANE, S. (1977). " A Survey of Modern Algebra". 4 th.

ed. Macmillan : New York.

BOX, G.E.P. and PIERCE, D.A. (1970). " Distribution of residual autocorre-

lations in ARIMA time series models". J. Amer. Statist.

Assoc. 65, 1509-26.

BOX, G.E.P. and JENKINS, G.M. (1976). " Time Series Analysis". (Revised. ed.).

Holden Day: San Francisco.

CLEVELAND, W.P. and TIAO, G.C. (1979). " Modeling Seasonal Time Series".

Revue Economic Appliquée. 32, 107-29.

COX, D.R. (1983). " Asymptotic Theory : Some Recent Developments".

Qüestió, Vol 7, nº 4, 527-529.

COX, D.R. and HINKLEY, D.V. (1974). " Theoretical Statistics". Chapman
and Hall : London.

CRAMER, H. (1946). " Mathematical Methods of Statistics". Prince ton Uni-
versity Press.

DAVIES, N. , TRIGGS, C.M. and NEWBOLD, P. (1977). " Significance Levels
of the Box-Pierce Portmanteau Statistic in Finite
Samples". Biometrika. 64, 517-522.

DAVIES, N. and NEWBOLD, P. (1979). " Some Power Studies of a Portmanteau
Test of Time Series Model Specification". BiométriKa.
66, 153-155.

DEISTLER, M. (1983). " The Properties of the Parameterization of ARMAX.
Systems and Their Relevance for Structural Estimation
and Dynamic Specification". Econométrica. 51, 1187-1207.

DUNSMUIR, W. and HANNAN, E.J. (1976). " Vector Linear Time Series Models"
Adv. in Appl. Probab. 8, 339-364.

FISHER, R.A. (1946). " A System of Scoring Linkage Data with Special
Reference to Pied Factors in Mice ". Amer. Nat. 80,
568-78.

GERALD, S.R. (1980). " Matrix Derivatives Lecture Notes in Statistics.

Vol. 2. Marcel Dekker. Inc: New York and Basel.

GODFREY, L.G. (1979). " Testing the Adequacy of a Time Series Model"

BiométriKa. 66, 67-72.

GODOLPHIN, E.J. (1980). " A Method for Testing the Order of an Autoregressive-Moving Average Process". BiométriKa. 67, 699-703.

HAUGH, L.D. and BOX, G.E.P. (1977). " Identification of Dynamic Regression. (Distributed Lag.) Models Connecting two. Time Series". J. Amer. Statist. Assoc. 72, 121-130.

HANNAN, E.J. (1970). " Multiple Time Series". Wiley: New York.

HILLER, S.C. and TIAO, G.C. (1979). " Likelihood Function of Stationary Multiple Autoregressive Moving Average Models". J. Amer. Statist. Assoc. 74, 652-660.

HOSKING, J.R.M. (1980, a) " The Multivariate Portmanteau Statistic ". J. Amer. Statist Assoc. 75, 602-608.

(1980, b). " Lagrange Multiplier Test of Time Series Models". J.R. Statist. Soc. B. 40, 170-181.

IBRAHIMOV, I. and ROZANOV, Y. (1978). " Gaussian Random Processes".

Springer-Verlag: Berlin.

KOHN, R. (1978). "Asymptotic Properties of Time Domain Gaussian Estimators". Adv. in Appl. Probab. 10, 339-359.

(1979). "Asymptotic Estimation and Hypothesis Results for Vector Linear Time Series Models". Econometrica, 47, 1005-1030.

LJUNG, G.M. and BOX, G.E.P. (1978). "On a Measure of Lack of Fit in Time Series Models". Biométrie, 65, 297-303.

MAC LEOD, A.I. (1975). "Derivation of the theoretical Autocovariance Function of ARMA Time Series". Applied Statistics, 24, 255-256.

(1978). "On the Distribution of Residual Autocorrelations in Box-Jenkins Models". J.R. Statist. Soc. B. 40, 296-302.

(1979). "Distribution of Residual Cross-Correlation in Univariate ARMA Time Series Models". J. Amer. Statist. Assoc, 74., 849-855

MAC LEOD, A.I. and LI, W.K. (1983). "Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations". Journal of Time Series Analysis, 4, 269-273.

MOORE, D.S. (1977). "Generalized Inverses, Wald's Method and the Construction of Chi-Squared Test of Fit". Journal of the American Statistical Association

MORAN, P.A.P. (1970). "On Asymptotically Optimal Test of Composite Hypotheses". Biometrika. 57, 47-55.

NEUDECKER, H. (1969). " Some Theorems on Matrix Differentiation with Special Reference to Kronecker Matrix Products". J. Amer. Stat. Ass. 69, 953-963.

NEWBOLD, P. (1980). " The Equivalence of Two Tests of Time Series Model Adequacy". Biométri-
ka, 67. 463-465.

NICHOLLS, D.F. (1976). " The Efficient Estimation of Vector Linear Time Series Models". Biométri-
ka, 63. 381-390.

POLLOCK, D.S.G. (1979). " The Algebra of Econometrics". Wiley: New York.

POSKITT, D.S. and TREMAYNE, A.R. (1980). " Testing the Specification of a Fitted Autoregressive Moving-Average Model". Biométri-
ca, 67. 359-363.

(1981). " An Approach to Testing Linear Time Series Models". Ann. Statist. 9, 974-986.

(1982). " Diagnostic Tests for Multiple Time Series Models". Ann. Statist. 10, 114-120.

RAO, C.R. (1947). " Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation".

RAO, C.R. and MITRA, S.K. (1971). " Generalized Inverse of Matrices and its Applications". Wiley; New York.

RAO, C. R. (1973). "Linear Statistical Inference and its Applications".
Wiley: New York.

RAO, M.B. ; SUBBA RAO, T. and WALKER, A.M. (1983). "On the Existence
of Some Bilinear Time Series Models". Journal of Time
Series Analysis. 4, 95-110.

ROGERS, G.S. (1980). "Matrix Derivatives" Marcel Dekker, inc. New York and
Basel.

ROZANOV, Y.A. (1967). "Stationary Random Processes. Series in Time
Series Analysis ". Holden Day.

SAIKKONEN, P. (1983). "Asymptotic Relative Efficiency of Some Tests of
Fit in Time Series Models". Journal of Time Series Ana-
lysis, 4. 69-78.

SILVEY, S.D. (1959). "The Lagrangian Multiplier Test". Ann. Math. Statist.
30, 389-407.

(1970). "Statistical Inference" . Chapman and Hall: London.

SPIEGELHALTER, D.J. (1983). "Diagnostic Tests of Distributional Shape"
Biométrie, 70. 401-409.

TIAO, G. C. and GRUPE, M.R. (1980). "Hidden Periodic Autoregressive-Mo-
ving Average Models in Time Series Data". Biométrie,
67. 365-73.

TRUONG-VAN, B. (1983). " Generalized Seasonal ARIMA processes: Regularity/Singularity and Linear Prediction". Journal of Time Series Analysis, 4. 111-126.

WHITTLE, P. (1952). " Tests of Fit in Time Series". Biométrie, 39. 309-318.

WALD, A. (1943). " Test of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is Large". Transac. Amer. Math. Soc. ,54. 426-482.

UNIVERSIDAD DE SEVILLA

FACULTAD DE CIENCIAS MATEMÁTICAS

Trasido el Tribunal integrado por ...
a ... de la fecha, para juzgar la Tesis ...

Beser Heras Martínez

... "Nuevos Tests para contrastar la
bondad de modelos aritméticos y memoria de
series temporales"

... otorgarle la calificación de Sobresaliente "cum laude"

Sevilla, 17 de Abril

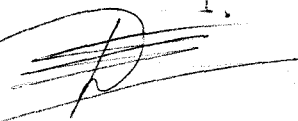
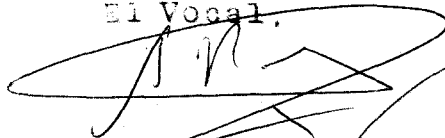
85

El Vocal,

El Vocal,

1.

H. Castro



El Secretario,

El Doctorado,

M. J. ...

